МИНИСТЕРСТВО ОБРАЗОВАНИЯ И НАУКИ
РОССИЙСКОЙ ФЕДЕРАЦИИ
ТОМСКИЙ ГОСУДАРСТВЕННЫЙ УНИВЕРСИТЕТ

МАТЕРИАЛЫ

VII Международной молодежной научной конференции

«МАТЕМАТИЧЕСКОЕ И ПРОГРАММНОЕ ОБЕСПЕЧЕНИЕ ИНФОРМАЦИОННЫХ, ТЕХНИЧЕСКИХ И ЭКОНОМИЧЕСКИХ СИСТЕМ»

Томск, 23–25 мая 2019 г.

Под общей редакцией
кандидата технических наук И.С. Шмырина

Томск
Издательский Дом Томского государственного университета
2019

ББК 22.17–22.19
УДК 519.2, 519.7, 519.8
Т78

ЧЛЕНЫ КОЛЛЕГИИ, РУКОВОДИТЕЛИ НАУЧНЫХ РЕДАКЦИЙ
ПО НАПРАВЛЕНИЯМ:

д-р техн. наук, проф. **А.А. Глазунов** – научная редакция «Механика, математика»; д-р техн. наук, проф. **Э.Р. Шрагер** – научная редакция «Механика, математика»; д-р техн. наук, проф. **А.М. Горцев** – научная редакция «Информатика и кибернетика»; д-р техн. наук, проф. **С.П. Сущенко** – научная редакция «Информатика и кибернетика»; д-р физ.-мат. наук, проф. **В.Г. Багров** – научная редакция «Физика»; д-р физ.-мат. наук, проф. **А.И. Потекаев** – научная редакция «Физика»; д-р биол. наук, проф. **С.П. Кулижский** – научная редакция «Биология»; д-р геол.-минер. наук, проф. **В.П. Парначев** – научная редакция «Науки о Земле, химия»; канд. хим. наук, доц. **Ю.Г. Слижов** – научная редакция «Науки о Земле, химия»; д-р филол. наук, проф. **Т.А. Демешкина** – научная редакция «История, филология»; д-р ист. наук, проф. **В.П. Зиновьев** – научная редакция «История, филология»; д-р экон. наук, проф. **В.И. Канов** – научная редакция «Юридические и экономические науки»; д-р юрид. наук, проф. **В.А. Уткин** – научная редакция «Юридические и экономические науки»; д-р ист. наук, проф. **Э.И. Черняк** – научная редакция «Философия, социология, психология, педагогика, искусствознание»; д-р психол. наук, проф. **Э.В. Галажинский** – научная редакция «Философия, социология, психология, педагогика, искусствознание»

НАУЧНАЯ РЕДАКЦИЯ ТОМА:

д-р техн. наук, проф. **А.М. Горцев**, д-р техн. наук, проф. **С.П. Сущенко**, д-р физ.-мат. наук, доц. **Ю.Г. Дмитриев**, д-р физ.-мат. наук, доц. **С.П. Моисеева**, д-р физ.-мат. наук, проф. **В.В. Конев**, д-р техн. наук, проф. **А.Ю. Матросова**, д-р техн. наук, проф. **А.А. Назаров**, д-р техн. наук, проф. **К.И. Лившиц**, канд. техн. наук **С.А. Останин**, канд. техн. наук **И.С. Шмырин**.

Т78 Труды Томского государственного университета. – Т. 303. Серия физико-математическая: Математическое и программное обеспечение информационных, технических и экономических систем : материалы VII Международной молодежной научной конференции. Томск, 23–25 мая 2019 г. / под общ. ред. И.С. Шмырина. – Томск : Издательский Дом Томского государственного университета, 2019. – 368 с.

ISBN 978-5-94621-851-1

Сборник содержит материалы VII Международной молодёжной научной конференции «Математическое и программное обеспечение информационных, технических и экономических систем», проводившейся 23–25 мая 2019 г. на базе Института прикладной математики и компьютерных наук Томского государственного университета. Материалы сгруппированы в соответствии с работавшими на конференции секциями.

Для научных работников, преподавателей, аспирантов, магистрантов и студентов.

УДК 539.3.004
ББК 22,25.22.251.22.62

ISBN 978-5-94621-851-1 © Томский государственный университет, 2019

I. УПРАВЛЕНИЕ И ОБРАБОТКА ИНФОРМАЦИИ

АППРОКСИМАЦИЯ РЕАЛЬНОГО ТРАФИКА ОДНИМ ИЗ ДВАЖДЫ СТОХАСТИЧЕСКИХ ПОТОКОВ СОБЫТИЙ

Л.А. Нежельская, Е.Ф. Сидорова
Томский государственный университет

Введение

Интенсивное развитие вычислительной техники, спутниковых, компьютерных, мобильных сетей связи выявили неадекватность пуассоновского потока [1] реальным информационным потокам и непригодность соответствующих моделей функционирования современных телекоммуникационных систем для анализа протекающих в них процессов. Таким образом, требования практики послужили стимулом к созданию новых математических моделей в виде дважды стохастических потоков с интенсивностью, представляющей собой либо непрерывный [2,3], либо кусочно-постоянный [4,5] случайный процесс с конечным (произвольным) числом состояний. Построенные модели находят применение во многих отраслях науки и техники: теория сетей связи и радиосвязи [6], исследование биофизических процессов [7], статистическая обработка данных [8], финансово-экономическое моделирование и управление запасами [9] и т.д.

При изучении реального трафика первоначально возникает проблема выбора математической модели потока, адекватно его описывающей. Согласно достаточно большому числу исследований потоки с переключениями (потоки со ступенчатой функцией интенсивности), чаще называемые MC-потоками (Markov chain) [4] либо MAP-потоками (Markovian Arrival Process) [5], являются наиболее характерными и подходящими математическими моделями потоков в реальных телекоммуникационных сетях [10–12]. К числу таких потоков, в частности, относится рассматриваемый в рамках настоящей статьи обобщенный синхронный поток событий второго порядка [13,14].

При реализации адаптивного управления математической моделью реальной системы выделяют два основных класса задач: оценивание в произвольный момент времени состояний входящего потока (фильтрация его интенсивности) [13–15] и оценивание его параметров [16,17] по наблюдениям за ним. В настоящей работе имеет место сопоставление численных результатов оценивания параметров распределения длительности интервала между событиями коррелированного и рекуррентного потоков, используя соответствующие аналитических формулы. Кроме того, приводятся полученные на их основе результаты аппроксимации реальных данных трафика, собранных Леландом и Уилсоном [18] по нескольким локальным сетям Ethernet в Bellcore Morristown Research and Engineering Center, исследуемым дважды стохастическим потоком.

1. Математическая модель потока

Рассматривается функционирующий в стационарном режиме обобщенный синхронный поток событий второго порядка, сопровождающий случайный процесс которого $\lambda(t)$ является кусочно-постоянным с двумя состояниями S_1 и S_2; здесь и далее S_i понимается как i-е состояние $\lambda(t)$ и имеет место при $\lambda(t)=\lambda_i$, $i=1,\ 2$, $\lambda_1>\lambda_2\geq 0$.

Длительность интервала между событиями потока в i-м состоянии определяется случайной величиной $\eta_i=\min\left(\xi_i^{(1)},\ \xi_i^{(2)}\right)$, где случайные величины $\xi_i^{(1)}$ и $\xi_i^{(2)}$ независимы и распределены по законам $F_i^{(1)}(t)=1-e^{-\lambda_i t}$ и $F_i^{(2)}(t)=1-e^{-\alpha_i t}$, соответственно. В момент наступления события потока процесс $\lambda(t)$ либо переходит из i-го состояния в j-е, $i\neq j$, либо остается в i-м состоянии, $i=j$, с вероятностью $P_1^{(1)}\left(\lambda_j\,|\,\lambda_i\right)$ или

$P_1^{(2)}\left(\lambda_j\,|\,\lambda_i\right)$, $i,j=1,\ 2$, в зависимости от значения, принимаемого η_i, $i=1,\ 2$. При этом $P_1^{(1)}\left(\lambda_j\,|\,\lambda_i\right)+P_1^{(1)}\left(\lambda_i\,|\,\lambda_i\right)=1$, $P_1^{(2)}\left(\lambda_j\,|\,\lambda_i\right)+P_1^{(2)}\left(\lambda_i\,|\,\lambda_i\right)=1$, $i,j=1,2$, $i\neq j$.

В сделанных предположениях процесс $\lambda(t)$ является скрытым марковским [13].

Вариант возникающей ситуации приведен на рис. 1, где S_1, S_2 – состояния $\lambda(t)$; $t_1,t_2,...,t_k,...$ – доступные наблюдению моменты наступления событий в потоке.

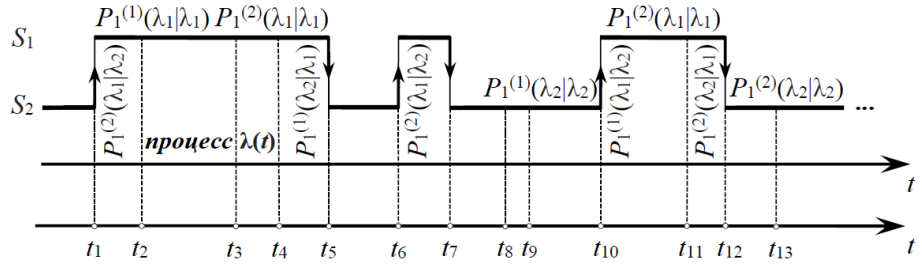


Рис. 1. Формирование наблюдаемого потока событий

Таким образом, последовательность наблюдаемых на интервале (t_0,t), где t_0 – начало наблюдений, t – окончание наблюдений, моментов наступления событий $t_1,t_2,...,t_k,...$ порождает вложенную цепь Маркова $\{\lambda(t_k)\}$ в силу сформулированных предпосылок.

2. Оценивание параметров распределения методом моментов

Определим значение длительности k-го интервала между событиями наблюдаемого потока t_k и t_{k+1}, $k=1,2,...$, как $\tau_k=t_{k+1}-t_k$, $\tau_k\geq 0$. Для плотности вероятности τ_k вследствие функционирования потока в стационарном режиме справедливо $p(\tau_k)=p(\tau)$, $\tau\geq 0$, для $k\geq 1$, что позволяет без ограничения общности положить момент наступления события t_k равным нулю или, что то же самое, момент наступления события $\tau=0$.

Расположение на временной оси двух смежных интервалов (t_k,t_{k+1}) и (t_{k+1},t_{k+2}) длительностей $\tau_k=t_{k+1}-t_k$ и $\tau_{k+1}=t_{k+2}-t_{k+1}$, $\tau_k\geq 0$, $\tau_{k+1}\geq 0$, ввиду стационарности потока произвольно, что дает возможность, положив $k=1$, рассматривать (t_1,t_2) и (t_2,t_3) со значениями длительностей $\tau_1=t_2-t_1$ и $\tau_2=t_3-t_2$. При этом $\tau_1=0$ соответствует моменту t_1 наступления события потока, $\tau_2=0$ – моменту t_2 наступления следующего его события, а совместная плотность вероятности τ_1 и τ_2 есть $p(\tau_1,\tau_2)$, $\tau_1\geq 0$, $\tau_2\geq 0$.

Теорема 1. Одномерная плотность вероятности значений длительности интервала между событиями обобщенного синхронного потока второго порядка имеет вид

$$p(\tau)=\gamma z_1 e^{-z_1\tau}+(1-\gamma)z_2 e^{-z_2\tau},\ \tau\geq 0,\ z_1=\lambda_1+\alpha_1,\ z_2=\lambda_2+\alpha_2,$$

$$\gamma=\frac{z_1\left[\lambda_2 P_1^{(1)}\left(\lambda_1\,|\,\lambda_2\right)+\alpha_2 P_1^{(2)}\left(\lambda_1\,|\,\lambda_2\right)\right]}{z_2\left[\lambda_1 P_1^{(1)}\left(\lambda_2\,|\,\lambda_1\right)+\alpha_1 P_1^{(2)}\left(\lambda_2\,|\,\lambda_1\right)\right]+z_1\left[\lambda_2 P_1^{(1)}\left(\lambda_1\,|\,\lambda_2\right)+\alpha_2 P_1^{(2)}\left(\lambda_1\,|\,\lambda_2\right)\right]}. \qquad (1)$$

Отметим, что (1) есть плотность гиперэкспоненциального распределения с соответствующими параметрами; в дальнейшем принимается $(\lambda_1+\alpha_1)\neq(\lambda_2+\alpha_2)$.

Теорема 2. Обобщенный синхронный поток событий второго порядка в общем случае является *коррелированным* с совместной плотностью вероятности $p(\tau_1,\tau_2)$ вида

$$p(\tau_1,\tau_2)=\left[1-\frac{\lambda_1 P_1^{(1)}(\lambda_2\,|\,\lambda_1)+\alpha_1 P_1^{(2)}(\lambda_2\,|\,\lambda_1)}{\lambda_1+\alpha_1}-\frac{\lambda_2 P_1^{(1)}(\lambda_1\,|\,\lambda_2)+\alpha_2 P_1^{(2)}(\lambda_1\,|\,\lambda_2)}{\lambda_2+\alpha_2}\right]\times$$

$$\times\gamma(1-\gamma)\left[(\lambda_1+\alpha_1)e^{-(\lambda_1+\alpha_1)\tau_1}-(\lambda_2+\alpha_2)e^{-(\lambda_2+\alpha_2)\tau_1}\right]\times \qquad (2)$$

$$\times\left[(\lambda_1+\alpha_1)e^{-(\lambda_1+\alpha_1)\tau_2}-(\lambda_2+\alpha_2)e^{-(\lambda_2+\alpha_2)\tau_2}\right]+p(\tau_1)\,p(\tau_2),\ \tau_1\ge 0,\ \tau_2\ge 0,$$

$$\gamma=\frac{(\lambda_1+\alpha_1)\left[\lambda_2 P_1^{(1)}(\lambda_1\,|\,\lambda_2)+\alpha_2 P_1^{(2)}(\lambda_1\,|\,\lambda_2)\right]}{(\lambda_2+\alpha_2)\left[\lambda_1 P_1^{(1)}(\lambda_2\,|\,\lambda_1)+\alpha_1 P_1^{(2)}(\lambda_2\,|\,\lambda_1)\right]+(\lambda_1+\alpha_1)\left[\lambda_2 P_1^{(1)}(\lambda_1\,|\,\lambda_2)+\alpha_2 P_1^{(2)}(\lambda_1\,|\,\lambda_2)\right]};$$

плотности $p(\tau_k)$ определены в (1) для $\tau=\tau_k$, $k=1,2$.

Установим случаи, при которых исследуемый поток событий становится *рекуррентным*: из формулы (2) для совместной плотности вероятности $p(\tau_1,\tau_2)$ определим условия, при которых она факторизуется, т.е. $p(\tau_1,\tau_2)=p(\tau_1)\,p(\tau_2)$.

1. Если $\dfrac{\lambda_1 P_1^{(1)}(\lambda_2\,|\,\lambda_1)+\alpha_1 P_1^{(2)}(\lambda_2\,|\,\lambda_1)}{\lambda_1+\alpha_1}+\dfrac{\lambda_2 P_1^{(1)}(\lambda_1\,|\,\lambda_2)+\alpha_2 P_1^{(2)}(\lambda_1\,|\,\lambda_2)}{\lambda_2+\alpha_2}=1$, то

$$p(\tau)=\left(\lambda_1 P_1^{(1)}(\lambda_1\,|\,\lambda_1)+\alpha_1 P_1^{(2)}(\lambda_1\,|\,\lambda_1)\right)\ e^{-(\lambda_1+\alpha_1)\tau}+$$
$$+\left(\lambda_2 P_1^{(1)}(\lambda_2\,|\,\lambda_2)+\alpha_2 P_1^{(2)}(\lambda_2\,|\,\lambda_2)\right)\ e^{-(\lambda_2+\alpha_2)\tau},\ \tau\ge 0. \qquad (3)$$

Следующие условия факторизации $p(\tau_1,\tau_2)$ получим, анализируя

$$\gamma(1-\gamma)=$$
$$=\frac{(\lambda_1+\alpha_1)\left[\lambda_2 P_1^{(1)}(\lambda_1\,|\,\lambda_2)+\alpha_2 P_1^{(2)}(\lambda_1\,|\,\lambda_2)\right]\ (\lambda_2+\alpha_2)\left[\lambda_1 P_1^{(1)}(\lambda_2\,|\,\lambda_1)+\alpha_1 P_1^{(2)}(\lambda_2\,|\,\lambda_1)\right]}{\left\{(\lambda_1+\alpha_1)\left[\lambda_2 P_1^{(1)}(\lambda_1\,|\,\lambda_2)+\alpha_2 P_1^{(2)}(\lambda_1\,|\,\lambda_2)\right]+(\lambda_2+\alpha_2)\left[\lambda_1 P_1^{(1)}(\lambda_2\,|\,\lambda_1)+\alpha_1 P_1^{(2)}(\lambda_2\,|\,\lambda_1)\right]\right\}^2}.$$

2. Если $\lambda_2 P_1^{(1)}(\lambda_1\,|\,\lambda_2)+\alpha_2 P_1^{(2)}(\lambda_1\,|\,\lambda_2)=0$, то

$$p(\tau)=(\lambda_2+\alpha_2)e^{-(\lambda_2+\alpha_2)\tau},\ \tau\ge 0. \qquad (4)$$

3. Если $\lambda_1 P_1^{(1)}(\lambda_2\,|\,\lambda_1)+\alpha_1 P_1^{(2)}(\lambda_2\,|\,\lambda_1)=0$, то

$$p(\tau)=(\lambda_1+\alpha_1)e^{-(\lambda_1+\alpha_1)\tau},\ \tau\ge 0. \qquad (5)$$

Если выполняется одно из перечисленных условий и плотность вероятности принимает вид (3), (4) или (5), то имеет место *рекуррентный* поток событий.

Рассмотрим выборку $\tau_1,\tau_2,...,\tau_n$ из распределения (1), зависящего от трех неизвестных параметров z_1, z_2, γ; при этом без ограничения общности доопределим $z_1>z_2$. Согласно методу моментов, в силу близости теоретической и эмпирической функций распределения при достаточно больших n, теоретический начальный момент l-го порядка $M\tau^l=\int_0^\infty \tau^l p(\tau\,|\,z_1,z_2,\gamma)d\tau$ близок к соответствующему выборочному моменту – статистике $C_l=\dfrac{1}{n}\sum_{k=1}^{n}\tau_k^{\,l}$, где $\tau_k=t_{k+1}-t_k$. Таким образом, оценки параметров плотности вероятности длительности интервала между событиями *коррелированного* потока z_1, z_2, γ определяются как решение системы уравнений моментов $M\tau^l=C_l$, $l=\overline{1,3}$, и имеют вид

$$\hat{z}_{1,2}=\frac{1}{2}\left\{\frac{6C_1C_2-2C_3}{3C_2^{\,2}-2C_1C_3}\pm\sqrt{\left(\frac{6C_1C_2-2C_3}{3C_2^{\,2}-2C_1C_3}\right)^2-4\frac{12C_1^{\,2}-6C_2}{3C_2^{\,2}-2C_1C_3}}\right\},\qquad(6)$$

$$\hat{\gamma}=\hat{z}_1\left(1-C_1\hat{z}_2\right)\left(\hat{z}_1-\hat{z}_2\right)^{-1}.\qquad(7)$$

Рассмотрим далее $\tau_1,\tau_2,...,\tau_n$ из распределения $p\left(\tau\,|\,z_1,z_2,\beta_1,\beta_2\right)=\beta_1e^{-z_1\tau}+\beta_2e^{-z_2\tau}$, зависящего от $z_1=\lambda_1+\alpha_1$, $z_2=\lambda_2+\alpha_2$, $\beta_1=\lambda_1P_1^{(1)}\left(\lambda_1\,|\,\lambda_1\right)+\alpha_1P_1^{(2)}\left(\lambda_1\,|\,\lambda_1\right)$, $\beta_2=\lambda_2P_1^{(1)}\left(\lambda_2\,|\,\lambda_2\right)+\alpha_2P_1^{(2)}\left(\lambda_2\,|\,\lambda_2\right)$; $z_1>z_2$. Согласно методу моментов, оценки параметров плотности вероятности длительности интервала между событиями *рекуррентного* потока вида (3) находятся однозначно из системы $M\tau^l=C_l$, $l=\overline{1,\,4}$, как

$$\hat{z}_{1,2}=\frac{1}{2}\left\{\frac{12C_2C_3-6C_1C_4}{4C_3^{\,2}-3C_2C_4}\pm\sqrt{\left(\frac{12C_2C_3-6C_1C_4}{4C_3^{\,2}-3C_2C_4}\right)^2-4\frac{36C_2^{\,2}-24C_1C_3}{4C_3^{\,2}-3C_2C_4}}\right\},\qquad(8)$$

$$\hat{\beta}_1=\left(\hat{z}_1^{\,3}C_1-\hat{z}_1^{\,3}\hat{z}_2C_2/2\right)\left(\hat{z}_1-\hat{z}_2\right)^{-1},\ \hat{\beta}_2=\left(\hat{z}_1\hat{z}_2^{\,3}C_2/2-\hat{z}_2^{\,3}C_1\right)\left(\hat{z}_1-\hat{z}_2\right)^{-1}.\qquad(9)$$

Подчеркнем, что решения (6), (7) и (8), (9) единственны, однако только (8), (9) обеспечивает состоятельность получаемых оценок параметров $\hat{z}_1$, $\hat{z}_2$, $\hat{\beta}_1$, $\hat{\beta}_2$ [19].

3. Результаты численных расчетов

Получение численных результатов оценивания параметров распределений опосредовано работой программы, реализованной средствами языка программирования C# в среде разработки Microsoft Visual Studio в виде графического пользовательского интерфейса. В основу реализации положено имитационное моделирование потока с тем, чтобы получить статистики C_l, $l=\overline{1,4}$, с последующим вычислением соответствующих модельным данным оценок $\hat{z}_1$, $\hat{z}_2$, $\hat{\gamma}$ по формулам (6), (7) или $\hat{z}_1$, $\hat{z}_2$, $\hat{\beta}_1$, $\hat{\beta}_2$ – по (8), (9).

Для определения качества оценок поставлены эксперименты. В первом отслеживается временной интервал установления стационарного режима функционирования потоков и обнаруживается их соотношение между собой. Результаты работы модели получены при значениях параметров, зафиксированных в табл. 1, для каждого значения времени моделирования T_m в $N=100$ независимых реализациях; для каждой определялись оценки $\hat{\theta}^{(k)}$, $k=\overline{1,\ N}$, соответствующих θ (согласно (6), (7) для *коррелированного* потока и (8), (9) – для *рекуррентного*), на основе которых вычислялись выборочное среднее оценки $\hat{M}\left\{\hat{\theta}\right\}=\frac{1}{N}\sum_{k=1}^{N}\hat{\theta}^{(k)}$ и выборочная квадратическая ошибка $\sqrt{\hat{V}\left\{\hat{\theta}\right\}}$, $\hat{V}\left\{\hat{\theta}\right\}=\frac{1}{N}\sum_{k=1}^{N}\left(\hat{\theta}^{(k)}-\theta\right)^2$, $\theta\in\left\{z_1,z_2,\gamma\right\}$, $\hat{\theta}\in\left\{\hat{z}_1,\hat{z}_2,\hat{\gamma}\right\}$ или $\theta\in\left\{z_1,z_2,\beta_1,\beta_2\right\}$, $\hat{\theta}\in\left\{\hat{z}_1,\hat{z}_2,\hat{\beta}_1,\hat{\beta}_2\right\}$.

Таблица 1

Модельные данные

$\lambda_1=5.6$	*коррелированный обобщенный синхронный поток второго порядка*
$\lambda_2=2.4$	$P_1^{(l)}\left(\lambda_i\,|\,\lambda_i\right)=0.15$, $P_1^{(l)}\left(\lambda_j\,|\,\lambda_i\right)=0.85$, $l=1,2$, $i,j=1,2$, $i\neq j$
$\alpha_1=3.1$	*рекуррентный обобщенный синхронный поток второго порядка*
$\alpha_2=0.7$	$P_1^{(l)}\left(\lambda_1\,|\,\lambda_1\right)=0.85$, $P_1^{(l)}\left(\lambda_2\,|\,\lambda_1\right)=0.15$, $P_1^{(l)}\left(\lambda_2\,|\,\lambda_2\right)=0.15$, $P_1^{(l)}\left(\lambda_1\,|\,\lambda_2\right)=0.85$, $l=1,2$

Таблица 2

Результаты первого статистического эксперимента

θ	T_m	100	200	300	400	500	600	700	800	900	1000
коррелированный обобщенный синхронный поток событий второго порядка											
$z_1 = 8.7$	$\hat{M}\{\hat{z}_1\}$	9.1869	9.4047	9.4732	8.8685	8.9751	8.5591	8.7149	8.9234	8.6733	8.8603
	$\sqrt{\hat{V}\{\hat{z}_1\}}$	1.5295	1.4679	1.4775	1.0442	0.8134	0.7732	0.6568	0.7197	0.6233	0.5226
$z_2 = 3.1$	$\hat{M}\{\hat{z}_2\}$	2.4258	2.4420	2.4875	2.4071	2.4079	2.3681	2.3786	2.4300	2.3924	2.4417
	$\sqrt{\hat{V}\{\hat{z}_2\}}$	0.8110	0.7661	0.6667	0.7018	0.6907	0.6774	0.6247	0.6336	0.6133	0.5598
$\gamma = 0.5$	$\hat{M}\{\hat{\gamma}\}$	0.4928	0.4774	0.4705	0.4999	0.4916	0.5140	0.5069	0.4916	0.5041	0.4884
	$\sqrt{\hat{V}\{\hat{\gamma}\}}$	0.0895	0.0657	0.0638	0.0522	0.0394	0.0357	0.0312	0.0351	0.0357	0.0301
рекуррентный обобщенный синхронный поток событий второго порядка											
$z_1 = 8.7$	$\hat{M}\{\hat{z}_1\}$	9.1943	9.6051	9.8515	9.2953	9.4632	8.3462	9.2001	9.2796	9.0358	8.7579
	$\sqrt{\hat{V}\{\hat{z}_1\}}$	2.0252	2.1310	1.8032	1.5967	1.5852	0.9274	0.8660	0.9118	0.8835	0.8242
$z_2 = 3.1$	$\hat{M}\{\hat{z}_2\}$	2.2765	4.4376	2.5015	2.3687	2.4505	2.4174	2.4805	2.4270	2.4425	2.3238
	$\sqrt{\hat{V}\{\hat{z}_2\}}$	0.9690	0.8324	0.7558	0.7868	0.7094	0.7126	0.6915	0.6862	0.6652	0.5963
$\beta_1 = 7.395$	$\hat{M}\{\hat{\beta}_1\}$	7.9872	8.3992	8.5468	8.0760	8.1565	7.0727	7.8463	8.0241	7.7247	7.5386
	$\sqrt{\hat{V}\{\hat{\beta}_1\}}$	1.9233	2.0061	1.8118	1.5869	1.4402	0.9049	0.8750	0.8886	0.8510	0.8348
$\beta_2 = 0.465$	$\hat{M}\{\hat{\beta}_2\}$	0.3799	0.4234	0.4670	0.3903	0.4283	0.3509	0.4192	0.3982	0.3940	0.3604
	$\sqrt{\hat{V}\{\hat{\beta}_2\}}$	0.2375	0.1952	0.1867	0.1757	0.1732	0.1747	0.1417	0.1376	0.1316	0.1354

Анализ представленных в табл. 2 результатов текущего эксперимента приводит к утверждению о сильной зависимости получаемых оценок от времени моделирования, а именно: с увеличением значения T_m выборочные средние стабилизируются, а качество самих оценок улучшается в смысле уменьшения значений выборочной квадратической ошибки. Последнее вполне естественно и объясняется в первую очередь самой концепцией метода моментов, основывающегося на необходимых для проведения вычислений значениях статистик, взятых из соответствующих реализаций исследуемого потока событий. В статистиках заключена информация об интервалах между наблюдаемыми моментами наступления событий в *коррелированном* или *рекуррентном* потоке, следовательно, чем больше число интервалов в реализации, тем больше информации будет накоплено в статистиках и, как следствие, точность оценивания параметров $p(\tau)$ будет выше.

По сопоставимым параметрам ($\hat{z}_1$, $\hat{z}_2$) сравним качество получаемых оценок и скорость установления стационарного режима в *обоих* случаях моделирования потока.

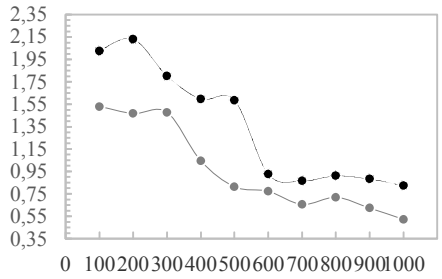


Рис. 2. График зависимости $\sqrt{\hat{V}\{\hat{z}_1\}}$ от времени T_m
(светлым маркером – для коррелированного потока, темным маркером – для рекуррентного)

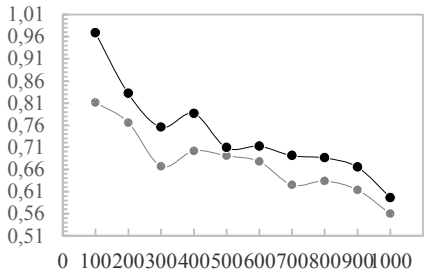


Рис.3. График зависимости $\sqrt{\hat{V}\{\hat{z}_2\}}$ от времени T_m
(светлым маркером – для коррелированного потока, темным маркером – для рекуррентного)

Последние графические результаты (рис. 2,3) позволяют сделать вывод о лучшем в смысле малости значений выборочной квадратической ошибки оценивании параметров $\hat{z}_1$, $\hat{z}_2$ *коррелированного* потока (светлым маркером на иллюстрациях), нежели $\hat{z}_1$, $\hat{z}_2$ рекуррентного потока (темным маркером). Такая закономерность объясняется, во-первых, синхронностью потока и, следовательно, трудоемкостью распознавания его состояния для вычисления соответствующего значения оценки интенсивности, во-вторых, ограниченностью содержащейся в $p(\tau)$ информации и, в-третьих, большим количеством необходимых для оценивания параметров (в *рекуррентном* потоке оцениваются четыре параметра (3), в *коррелированном* – три (1)). Выборочные квадратические ошибки $\sqrt{\hat{V}\{\hat{\theta}\}}$ для всех вариантов расчета являются достаточно стабильными при $T_m \geq 700$ ед. времени как для *коррелированного* ($\hat{\theta} \in \{\hat{z}_1, \hat{z}_2, \hat{\gamma}\}$), так и для *рекуррентного* ($\hat{\theta} \in \{\hat{z}_1, \hat{z}_2, \hat{\beta}_1, \hat{\beta}_2\}$) потоков. Как следствие, в дальнейшем принимается $T_m = 1000$ ед. времени.

Предметом исследования во втором статистическом эксперименте выступают зависимости $\hat{M}\{\hat{\theta}\}$ и $\sqrt{\hat{V}\{\hat{\theta}\}}$, $\hat{\theta} \in \{\hat{z}_1, \hat{z}_2, \hat{\gamma}\}$ или $\hat{\theta} \in \{\hat{z}_1, \hat{z}_2, \hat{\beta}_1, \hat{\beta}_2\}$, от изменений параметров λ_i, α_i, $i = 1,2$, при фиксированных $T_m = 1000$, $N = 100$ и исходных данных табл. 1.

Результаты табл. 3,4, демонстрируют, что при увеличении значений параметров λ_i, α_i, $i = 1,2$, при фиксированных остальных качество оценок $\hat{z}_1$, $\hat{z}_2$, $\hat{\gamma}$ для *коррелированного* и $\hat{z}_1$, $\hat{z}_2$, $\hat{\beta}_1$, $\hat{\beta}_2$ – для *рекуррентного* потоков улучшается в смысле уменьшения значений выборочной квадратической ошибки. Это объясняется тем, что увеличение каждого из λ_i, α_i, $i = 1,2$, влечет за собой более частое наступление событий, что увеличивает число наблюдаемых интервалов, на которых основываются статистики.

Таблица 3

Результаты второго статистического эксперимента

λ_1	α_1	5.6	3.1	5.8	3.3	6.0	3.5	6.2	3.7
коррелированный поток событий									
z_1		8.7		9.1		9.5		9.9	
$M\{\hat{z}_1\}$		9.0338		9.0830		9.5753		10.0315	
$\sqrt{\hat{V}\{\hat{z}_1\}}$		0.5439		0.5408		0.5009		0.4735	
z_2		3.1		3.1		3.1		3.1	
$\hat{M}\{\hat{z}_2\}$		2.4160		2.3879		2.3959		2.4118	
$\sqrt{\hat{V}\{\hat{z}_2\}}$		0.6852		0.7132		0.7069		0.6818	
γ		0.5		0.5		0.5		0.5	
$\hat{M}\{\hat{\gamma}\}$		0.4895		0.5021		0.4989		0.4916	
$\sqrt{\hat{V}\{\hat{\gamma}\}}$		0.0224		0.0225		0.0204		0.0201	
рекуррентный поток событий									
z_1		8.7		9.1		9.5		9.9	
$M\{\hat{z}_1\}$		8.8180		9.0938		10.2305		10.3914	
$\sqrt{\hat{V}\{\hat{z}_1\}}$		0.9322		0.8021		0.9655		0.8963	
z_2		3.1		3.1		3.1		3.1	
$\hat{M}\{\hat{z}_2\}$		2.3743		2.3440		2.4831		2.4400	
$\sqrt{\hat{V}\{\hat{z}_2\}}$		0.7385		0.7726		0.6259		0.6668	
β_1		7.395		7.735		8.075		8.415	
$\hat{M}\{\hat{\beta}_1\}$		7.5638		7.7930		8.8275		8.9843	
$\sqrt{\hat{V}\{\hat{\beta}_1\}}$		0.9122		0.7659		0.9697		0.8900	
β_2		0.465		0.465		0.465		0.465	
$\hat{M}\{\hat{\beta}_2\}$		0.3570		0.3435		0.4183		0.3880	
$\sqrt{\hat{V}\{\hat{\beta}_2\}}$		0.1409		0.1495		0.0829		0.0943	

Таблица 4

Результаты второго статистического эксперимента

λ_2	α_2	2.4	0.7	2.5	0.8	2.6	0.9	2.7	1.0
коррелированный поток событий									
z_1		8.7		8.7		8.7		8.7	
$M\{\hat{z}_1\}$		8.7112		9.0730		8.6678		8.8717	
$\sqrt{\hat{V}\{\hat{z}_1\}}$		0.6440		0.5144		0.4602		0.4203	
z_2		3.1		3.3		3.5		3.7	
$\hat{M}\{\hat{z}_2\}$		2.4023		2.5431		2.6778		2.8877	
$\sqrt{\hat{V}\{\hat{z}_2\}}$		0.7028		0.6578		0.6249		0.6244	
γ		0.5		0.5		0.5		0.5	
$\hat{M}\{\hat{\gamma}\}$		0.5002		0.4828		0.5053		0.4997	
$\sqrt{\hat{V}\{\hat{\gamma}\}}$		0.0251		0.0258		0.0249		0.0187	
рекуррентный поток событий									
z_1		8.7		8.7		8.7		8.7	
$M\{\hat{z}_1\}$		8.6262		8.6853		9.1458		9.0902	
$\sqrt{\hat{V}\{\hat{z}_1\}}$		0.7374		0.6997		0.6766		0.6351	
z_2		3.1		3.3		3.5		3.7	
$\hat{M}\{\hat{z}_2\}$		2.3131		2.4277		2.6953		2.7640	
$\sqrt{\hat{V}\{\hat{z}_2\}}$		0.9007		0.8941		0.8122		0.8431	
β_1		7.395		7.395		7.395		7.395	
$\hat{M}\{\hat{\beta}_1\}$		7.4181		7.4450		7.7588		7.6180	
$\sqrt{\hat{V}\{\hat{\beta}_1\}}$		0.6562		0.6084		0.5396		0.4903	
β_2		0.465		0.495		0.525		0.555	
$\hat{M}\{\hat{\beta}_2\}$		0.3270		0.3546		0.4562		0.4311	
$\sqrt{\hat{V}\{\hat{\beta}_2\}}$		0.1624		0.1696		0.1067		0.1107	

Определим по известной $p(\tau)$ вида (1) вероятностные характеристики случайной величины $\boldsymbol{\tau}$ – длительности интервала между соседними событиями *коррелированного* обобщенного синхронного дважды стохастического потока событий второго порядка.

Нетрудно получить выражения для математического ожидания (первого начального момента) и дисперсии (второго центрального момента) случайной величины $\boldsymbol{\tau}$, относящихся к числу наиболее важных и часто используемых характеристик

$$M\{\boldsymbol{\tau}\} = \gamma z_1^{-1} + (1-\gamma) z_2^{-1},\ D\{\boldsymbol{\tau}\} = 2\left(\gamma z_1^{-2} + (1-\gamma) z_2^{-2}\right) - \left(\gamma z_1^{-1} + (1-\gamma) z_2^{-1}\right)^2,$$

где z_1, z_2, γ определены в (1).

По известной совместной вероятности $p(\tau_1, \tau_2)$ вида (2) найдем числовую характеристику связи случайных величин $\boldsymbol{\tau}_1$ и $\boldsymbol{\tau}_2$ – длительностей двух смежных интервалов, а именно ковариацию $\boldsymbol{\tau}_1$ и $\boldsymbol{\tau}_2$ (второй смешанный центральный момент)

$$\operatorname{cov}\{\boldsymbol{\tau}_1, \boldsymbol{\tau}_2\} = \left[1 - \left(\lambda_1 P_1^{(1)}(\lambda_2 \mid \lambda_1) + \alpha_1 P_1^{(2)}(\lambda_2 \mid \lambda_1)\right) z_1^{-1} - \right.$$
$$\left. - \left(\lambda_2 P_1^{(1)}(\lambda_1 \mid \lambda_2) + \alpha_2 P_1^{(2)}(\lambda_1 \mid \lambda_2)\right) z_2^{-1}\right] \times$$
$$\times \gamma(1-\gamma)\left(z_1^{-1} - z_2^{-1}\right)^2,$$

тогда коэффициент корреляции случайных величин $\boldsymbol{\tau}_1$ и $\boldsymbol{\tau}_2$ запишется как

$$r_{\tau_1\tau_2} = \left[1-(\lambda_1 P_1^{(1)}(\lambda_2\,|\,\lambda_1)+\alpha_1 P_1^{(2)}(\lambda_2\,|\,\lambda_1))z_1^{-1}-\left(\lambda_2 P_1^{(1)}(\lambda_1\,|\,\lambda_2)+\alpha_2 P_1^{(2)}(\lambda_1\,|\,\lambda_2)\right)z_2^{-1}\right]\times$$
$$\times\gamma(1-\gamma)\left(z_1^{-1}-z_2^{-1}\right)^2\left[2\left(\gamma z_1^{-2}+(1-\gamma)z_2^{-2}\right)-\left(\gamma z_1^{-1}+(1-\gamma)z_2^{-1}\right)^2\right]^{-1}.$$

Согласно условию рекуррентности 1, введенные числовые характеристики длительности интервала между соседними событиями *рекуррентного* потока будут иметь вид

$$M\{\boldsymbol{\tau}\}=\beta_1 z_1^{-2}+\beta_2 z_2^{-2},\ D\{\boldsymbol{\tau}\}=2\left(\beta_1 z_1^{-3}+\beta_2 z_2^{-3}\right)-\left(\beta_1 z_1^{-2}+\beta_2 z_2^{-2}\right)^2,$$
$$\beta_1=\lambda_1 P_1^{(1)}(\lambda_1\,|\,\lambda_1)+\alpha_1 P_1^{(2)}(\lambda_1\,|\,\lambda_1),\ \beta_2=\lambda_2 P_1^{(1)}(\lambda_2\,|\,\lambda_2)+\alpha_2 P_1^{(2)}(\lambda_2\,|\,\lambda_2),$$

в то время как ковариация $\boldsymbol{\tau}_1$, $\boldsymbol{\tau}_2$ и коэффициент корреляции тождественно равны нулю.

В третьем эксперименте на имитационной модели были получены статистические оценки рассмотренных выше вероятностных характеристик: выборочное среднее $\hat{M}\{\boldsymbol{\tau}\}=\frac{1}{n}\sum_{k=1}^{n}\tau_k$, выборочная дисперсия $\hat{D}\{\boldsymbol{\tau}\}=\frac{1}{n}\sum_{k=1}^{n}\left(\tau_k-\hat{M}\{\boldsymbol{\tau}\}\right)^2$, выборочная ковариация $\operatorname{cov}\left\{\boldsymbol{\tau}_1^{(n)},\boldsymbol{\tau}_2^{(n)}\right\}=\frac{1}{n}\sum_{k=1}^{n}\boldsymbol{\tau}_{1_k}\boldsymbol{\tau}_{2_k}-\left(\frac{1}{n}\sum_{k=1}^{n}\boldsymbol{\tau}_{1_k}\right)\left(\frac{1}{n}\sum_{k=1}^{n}\boldsymbol{\tau}_{2_k}\right)$ и выборочный коэффициент корреляции $\hat{r}_{\tau_1\tau_2}=\operatorname{cov}\left\{\boldsymbol{\tau}_1^{(n)},\boldsymbol{\tau}_2^{(n)}\right\}\left(\hat{D}\{\boldsymbol{\tau}_1\}\hat{D}\{\boldsymbol{\tau}_2\}\right)^{-1/2}$. Для числовых характеристик $M\{\boldsymbol{\tau}\}$, $D\{\boldsymbol{\tau}\}$ *рекуррентного* потока, кроме того, применялась описанная в [19] процедура построения двусторонних доверительных интервалов $I_\beta(\hat{\theta})$, $\hat{\theta}\in\left\{\hat{M}\{\boldsymbol{\tau}\},\hat{D}\{\boldsymbol{\tau}\}\right\}$, при доверительной вероятности $\beta=0.9$, правомерная в данном случае в силу центральной предельной теоремы (ЦПТ) при достаточно больших *n*, $n\to\infty$, т.к. $\tau_1,\tau_2,...,\tau_n$ есть выборка взаимно независимых случайных величин из распределения с конечными математическим ожиданием и дисперсией (подчеркнем, что для *коррелированного* потока условие взаимной независимости $\tau_1,\tau_2,...,\tau_n$ не удовлетворяется, вследствие чего доверительные интервалы определить по предложенному алгоритму не удается). Для фиксированных реализаций потока, заданных наборами параметров в табл. 1, получены результаты, приведенные в табл. 5.

Таблица 5

Результаты третьего статистического эксперимента

$M\{\boldsymbol{\tau}\}$	$\hat{M}\{\boldsymbol{\tau}\}$	$I_\beta(\hat{M}\{\boldsymbol{\tau}\})$	$D\{\boldsymbol{\tau}\}$	$\hat{D}\{\boldsymbol{\tau}\}$	$I_\beta(\hat{D}\{\boldsymbol{\tau}\})$	$\operatorname{cov}\{\boldsymbol{\tau}_1,\boldsymbol{\tau}_2\}$	$\operatorname{cov}\left\{\boldsymbol{\tau}_1^{(n)},\boldsymbol{\tau}_2^{(n)}\right\}$	$r_{\tau_1\tau_2}$	$\hat{r}_{\tau_1\tau_2}$
коррелированный обобщенный синхронный поток событий второго порядка									
0.2188	0.2194	–	0.0694	0.0701	–	-0.0075	-0.0076	-0.1087	-0.1088
рекуррентный обобщенный синхронный поток событий второго порядка									
0.1461	0.1457	(0.1442;0.1471)	0.0323	0.0323	(0.0310;0.0336)	0	$3\cdot10^{-6}$	0	$9\cdot10^{-5}$

Настоящий эксперимент поставлен при различных модельных параметрах. Анализ табл. 5 позволяет говорить о применимости статистик для решения задачи оценивания параметров распределений методом моментов, а также обоснованности использования аппарата имитационного моделирования для получения численных результатов. Например, статистика C_1 близка к $M\{\boldsymbol{\tau}\}$, $|M\{\boldsymbol{\tau}\}-\hat{M}\{\boldsymbol{\tau}\}|/M\{\boldsymbol{\tau}\}<0.003$, что приводит к утверждению о корректности и непротиворечивости построенной модели, имитирующей функционирование и *коррелированного*, и *рекуррентного* потоков.

Наиболее интересными с практической точки зрения являются результаты, полученные с использованием имитационной модели и возможностей языка программирования R в среде RStudio для реальных данных трафика, собранных Леландом и Уилсо-

ном [18] по нескольким локальным сетям Ethernet в и сформированных в виде записей, каждая из которых содержит по миллиону поступлений пакетов. Рассматривались записи **BC-pAug89**, начатая в 11:25 9 августа 1989 года и длящаяся около 3142,82 секунды (до тех пор, пока не был зарегистрирован миллион пакетов), и **BC-pOct89**, начатая в 11:00 5 октября 1989 года и длящаяся около 1759,62 секунды, в рамках задачи оценки параметров распределения длительности интервала между моментами поступления пакетов (событий) реального потока. Размеры пакетов в настоящем исследовании не учитывались; определяющими являлись моменты их поступлений, образующие реальный *коррелированный* информационный поток. Предполагается, что поток есть композиция (суперпозиция) нескольких дважды стохастических потоков событий.

Для полученных из **BC-pAug89** и **BC-pOct89** последовательностей длительностей интервалов, следуя методу скользящего окна, через шаг фиксированной длины 1000 было выделено по 981 подпоследовательности объема 10000, для каждой из которых вычислялись $\hat{z}_1^{\,real(k)}$, $\hat{z}_2^{\,real(k)}$, $\hat{\gamma}^{real(k)}$ с помощью $C_l^{(k)}$, $l=\overline{1,\,3}$, и с использованием (6), (7). Далее среди всех подпоследовательностей выбиралось по одной, для которой качество аппроксимации *коррелированным* обобщенным синхронным потоком второго порядка было наилучшим в смысле минимаксного критерия: $\delta_{rm}^0=\min\limits_k \delta_{rm_k}\left\{\hat{\theta}\right\}$, где $\delta_{rm_k}\left\{\hat{\theta}\right\}$ имеет смысл относительной погрешности и определяется как $\delta_{rm_k}\left\{\hat{\theta}\right\}=\max\limits_{\theta}\left[\,|\,\hat{\theta}^{real(k)}-\theta^{model}\,|\,/\,\hat{\theta}^{real(k)}\right]$, $k=\overline{1,N}$, $N=981$, $\hat{\theta}^{real(k)}\in\left\{\hat{z}_1^{\,real(k)},\hat{z}_2^{\,real(k)},\hat{\gamma}^{real(k)}\right\}$, $\theta^{model}\in\left\{z_1^{\,model},z_2^{\,model},\gamma^{model}\right\}$ – «истинные» значения параметров распределения, задаваемые исследователем в соответствии с оценками, полученными методом моментов по формулам (6), (7). Полученные результаты, имеющие место для подпоследовательности из **BC-pAug89**, отражены в табл. 6, в которой, кроме того, для выявления целесообразности и правомерности применения имитационного моделирования дополнительно приводятся $\hat{z}_1^{\,simulation}$, $\hat{z}_2^{\,simulation}$, $\hat{\gamma}^{simulation}$, определенные посредством процедуры оценивания соответствующих параметров на основе одной из многочисленных реализаций модели, а именно той (при соответствующих λ_i, α_i, $P_1^{(1)}\left(\lambda_j\,|\,\lambda_i\right)$, $P_1^{(2)}\left(\lambda_j\,|\,\lambda_i\right)$, $i,j=1,2$), для которой выбранный показатель качества $\delta_{rms}^1=\max\left\{\delta_{ms}\left\{\hat{\theta}\right\},\delta_{rs_k}\left\{\hat{\theta}\right\}\right\}$, $\delta_{ms}\left\{\hat{\theta}\right\}=\left|\theta^{model}-\hat{\theta}^{simulation}\right|\Big/\theta^{model}$, $\delta_{rs_k}\left\{\hat{\theta}\right\}=\left|\hat{\theta}^{real(k)}-\hat{\theta}^{simulation}\right|\Big/\hat{\theta}^{real(k)}$, $\hat{\theta}^{simulation}\in\left\{\hat{z}_1^{\,simulation},\hat{z}_2^{\,simulation},\hat{\gamma}^{simulation}\right\}$, k – номер подпоследовательности с минимальным δ_{rm}^0, достигает заданной точности $\delta_{rms}^1<0.009$. Таким образом, в данном рассмотрении имеет место более строгий критерий качества аппроксимации $\delta=\max\left\{\delta_{rm}^0,\,\delta_{rms}^1\right\}<0.008$.

Таблица 6

Аппроксимация реального трафика

$\hat{z}_1^{\,real}$	$z_1^{\,\mathrm{mod}\,el}$	$\hat{z}_1^{\,simulation}$	$\hat{z}_2^{\,real}$	$z_2^{\,\mathrm{mod}\,el}$	$\hat{z}_2^{\,simulation}$	$\hat{\gamma}^{real}$	$\gamma^{\mathrm{mod}\,el}$	$\hat{\gamma}^{simulation}$
390.7382	390.6379	390.1250	44.6496	44.5111	44.3187	0.9757	0.9722	0.9697
$\delta_{rm}\{\hat{z}_1\}$	$\delta_{ms}\{\hat{z}_1\}$	$\delta_{rs}\{\hat{z}_1\}$	$\delta_{rm}\{\hat{z}_2\}$	$\delta_{ms}\{\hat{z}_2\}$	$\delta_{rs}\{\hat{z}_2\}$	$\delta_{rm}\{\hat{\gamma}\}$	$\delta_{ms}\{\hat{\gamma}\}$	$\delta_{rs}\{\hat{\gamma}\}$
0.000257	0.001313	0.001569	0.003102	0.004323	0.007411	0.002972	0.002571	0.006149

Как отмечено выше, на другом участке наличествует удовлетворительное качество аппроксимации в смысле малости δ_{rm}^0 *коррелированным* потоком, определяемым другим набором параметров, например, для следующей исследованной из **BC-pAug89** подпоследовательности $\delta_{rm}^0=0.003981$. Полученные при этом графики для эмпирических плотностей вероятности длительности интервала между моментами поступления

пакетов, отвечающих выбранным выборкам, и соответствующих аналитически заданных плотностей приведены на рис. 4,5; построение осуществлялось с помощью встроенных функций R.

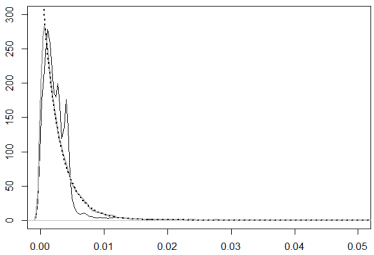


Рис. 4. **BC-pAug89**: графики эмпирических (по реальным данным – темным маркером, по модельным – светлым маркером) и аналитической (пунктиром) плотностей вероятностей

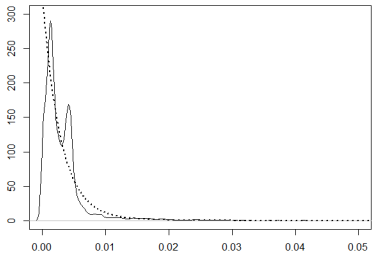


Рис. 5. **BC-pAug89**: графики эмпирической (по реальным данным – темным маркером) и аналитической (пунктиром) плотностей вероятностей

Для соответствующих двум непересекающимся участкам подпоследовательностей из **BC-pOct89** имеют место результаты аппроксимации $\delta_{rm}^{0} = 0.002718$ и $\delta_{rm}^{0} = 0.002031$.

Табл. 7 содержит значения математического ожидания и дисперсии, вычисленных на основании (1), а также выборочного среднего и выборочной дисперсии как их статистических оценок $\boldsymbol{\tau}$ – длительности интервала между моментами прибытия пакетов.

Таблица 7

Результаты работы с реальным трафиком

	$M\{\boldsymbol{\tau}\}$	$\hat{M}\{\boldsymbol{\tau}\}$	$\left\|M\{\boldsymbol{\tau}\} - \hat{M}\{\boldsymbol{\tau}\}\right\| \Big/ M\{\boldsymbol{\tau}\}$	$D\{\boldsymbol{\tau}\}\cdot 10^5$	$\hat{D}\{\boldsymbol{\tau}\}\cdot 10^5$	$\left\|D\{\boldsymbol{\tau}\} - \hat{D}\{\boldsymbol{\tau}\}\right\| \Big/ D\{\boldsymbol{\tau}\}$
BC-pAug89	0.003113	0.003041	0.023129	3.111125	3.050189	0.019586
	0.003521	0.003548	0.007668	2.747071	2.865587	0.043142
BC-pOct89	0.001806	0.001836	0.016611	0.113963	0.121034	0.062046
	0.001966	0.001959	0.003561	0.818885	0.836732	0.021794

По текущим данным отмечается соответствие *коррелированного* потока событий потоку поступлений, имеющему место на некотором интервале времени – наблюдается адекватность математической модели реальному информационному потоку.

С целью получения *рекуррентного* потока, отвечающего [18], использовалась методика перетасовывания, описанная в [20] и примененная в [21]. Рандомизация набора длительностей приводит к разрушению корреляции между ними, сохраняя при этом исходное распределение. По сформированным по каждой из **BC-pAug89** и **BC-pOct89** таким образом последовательностям, как и ранее, согласно методу скользящего окна, было выделено по 981 подпоследовательности; для каждой получены $\hat{z}_1^{\,sh(k)}$, $\hat{z}_2^{\,sh(k)}$, $\hat{\beta}_1^{\,sh(k)}$, $\hat{\beta}_2^{\,sh(k)}$, вычисленные на основе $C_l^{(k)}$, $l = \overline{1,\,4}$, по формулам (8), (9). Среди всех выделенных по записям подпоследовательностей с разрушенными корреляциями выбиралось по одной, для которой качество аппроксимации *рекуррентным* потоком было наилучшим в смысле введенного ранее критерия $\delta_{rm}^{0} = \min\limits_{k}\max\limits_{\theta}\left[\left|\hat{\theta}^{real(k)} - \theta^{model}\right| \Big/ \hat{\theta}^{real(k)}\right]$,

$k = \overline{1,N}$, $N = 981$, $\hat{\theta}^{real(k)} \in \left\{\hat{z}_1^{\,real(k)}, \hat{z}_2^{\,real(k)}, \hat{\beta}_1^{\,real(k)}, \hat{\beta}_2^{\,real(k)}\right\}$,

$\theta^{model} \in \left\{z_1^{\,model}, z_2^{\,model}, \beta_1^{\,model}, \beta_2^{\,model}\right\}$.

Аналогично предыдущему рассмотренному случаю в табл. 8 приведены результаты аппроксимации при критерии качества $\delta = \max\left\{\delta_{rm}^{0},\ \delta_{rms}^{1}\right\}$.

Таблица 8

Аппроксимация реального трафика

$\hat{z}_1^{\,real}$	$z_1^{\,model}$	$\hat{z}_1^{\,simulation}$	$\hat{z}_2^{\,real}$	$z_2^{\,model}$	$\hat{z}_2^{\,simulation}$	$\hat{\beta}_1^{\,real}$	$\beta_1^{\,model}$	$\hat{\beta}_1^{\,simulation}$	$\hat{\beta}_2^{\,real}$	$\beta_2^{\,model}$	$\hat{\beta}_2^{\,simulation}$
388.3658	389.3541	393.3584	49.2536	48.7826	48.0699	378.5851	378.6359	376.1599	1.3239	1.3429	1.3512
$\delta_{rm}\{\hat{z}_1\}$	$\delta_{ms}\{\hat{z}_1\}$	$\delta_{rs}\{\hat{z}_1\}$	$\delta_{rm}\{\hat{z}_2\}$	$\delta_{ms}\{\hat{z}_2\}$	$\delta_{rs}\{\hat{z}_2\}$	$\delta_{rm}\{\hat{\beta}_1\}$	$\delta_{ms}\{\hat{\beta}_1\}$	$\delta_{rs}\{\hat{\beta}_1\}$	$\delta_{rm}\{\hat{\beta}_2\}$	$\delta_{ms}\{\hat{\beta}_2\}$	$\delta_{rs}\{\hat{\beta}_2\}$
0.0025	0.0103	0.0129	0.0096	0.0146	0.0240	0.0001	0.0065	0.0064	0.0144	0.0062	0.0206

Анализ данных табл. 6, 7 приводит к обнаружению ухудшения аппроксимации в смысле увеличения δ. Такая закономерность обусловлена тем, что процедура перетасовывания способствует существенному сокращению корреляционной зависимости между длительностями, однако не позволяет полностью ее разрушить для реального потока.

Показатель качества в текущем определении достиг точности $\delta_{rm}^0 = 0.014612$ для **BC-pAug89** и $\delta_{rm}^0 = 0.004976$ для **BC-pOct89**. В табл. 9 приводятся числовые характеристики длительности интервала между событиями *рекуррентного* потока, а также их оценки.

Таблица 9

Результаты работы с реальным трафиком

	$M\{\tau\}$	$\hat{M}\{\tau\}$	$\left\|M\{\tau\} - \hat{M}\{\tau\}\right\| / M\{\tau\}$	$D\{\tau\}\cdot 10^5$	$\hat{D}\{\tau\}\cdot 10^5$	$\left\|D\{\tau\} - \hat{D}\{\tau\}\right\| / D\{\tau\}$
BC-pAug89	0.003062	0.003056	0.001960	2.659021	2.575102	0.031560
BC-pOct89	0.001841	0.001821	0.010864	1.255121	1.254798	0.000257

В совокупности полученные результаты свидетельствуют о том, на некоторых участках реальный информационный поток может быть аппроксимирован одним из видов дважды стохастических потоков, на другом – другим или тем же с отличными входными значениями параметров в соответствии с выбранными критериями качества.

Заключение

В данной работе на имитационной модели получены численные результаты оценивания параметров распределения длительности интервала в обобщенном синхронном потоке второго порядка в *коррелированном* и *рекуррентном* случаях по соответствующим аналитическим формулам. Каждый из приведенных экспериментов отражает корректное соотношение результатов с физической интерпретацией и демонстрирует приемлемое качество оценивания в смысле малости выбранного критерия. Проиллюстрированный пример работы с реальными данными трафика в рамках аппроксимации его исследуемым потоком отмечает практическую применимость аналитических результатов оценивания и результатов имитационного моделирования к реальным информационным потокам, математическим моделями которых выступают дважды стохастические потоки.

ЛИТЕРАТУРА

1. *Хинчин А.Я.* Математические методы теории массового обслуживания // Тр. МИАН СССР. – 1995. – Т. 49. – С. 3–122.

2. Cox D.R. The analysis of non-Markovian stochastic processes by the inclusion of supplementary variables // Proceedings of the Cambridge Philosophical Society. – 1955. – V. 51. – No. 3. – P. 433–441.

3. *Kingman J.F.C.* On doubly stochastic Poisson process // Proceedings of the Cambridge Philosophical Society. – 1964. – V. 60. – Is. 4. – P. 923–930.

4. *Башарин Г.П., Кокотушкин В.А., Наумов В.А.* О методе эквивалентных замен расчета фрагментов сетей связи // Изв. АН СССР. Техн. кибернетика. – 1980. – № 1. – С. 55–61.

5. *Lucantoni D.M.* New results on the single server queue with a batch markovian arrival process // Communications in Statistics Stochastic Models. – 1991. – V. 7. – P. 1–46.

6. *Basharin G.P., Gaidamaka Y.V., Samouylov K.E.* Mathematical theory of teletraffic and its application to the analysis of multiservice communication of next generation networks // Automatic Control and Computer Science. – 2013. – V. 47. – Is. 2. – P. 62–69.

7. *Card H.C.* Doubly stochastic Poisson processes in artificial neural learning // Neural Networks, IEEE Transactions. – 1998. – V. 9. – Is. 1. – P. 229–231.

8. *Bouzas P.R., Valderrama M.J., Aguilera A.M., Ruiz-Fuentes N.* Modelling the mean of a doubly stochastic Poisson process by functional data analysis // Computational Statistics and Data Analysis. – 2006. – V. 50 (10). – P. 2655–2667.

9. *Лившиц К.И., Сухотина Л.Ю., Шифердекер И.Ю.* Математическая модель деятельности некоммерческого фонда при дважды стохастическом потоке платежей // Вестник Томского государственного университета. Управление, вычислительная техника и информатика. – 2007. – № 1. – С. 36–43.

10. *Klimenok V., Dudin A., Vishnevsky V.* Tandem queueing system with correlated input and cross-traffic // Communications in Computer and Information Science. – 2013. – V. 370. – P. 416–425.

11. *Ниссенбаум О.В., Пахомов И.П.* Аппроксимация сетевого трафика моделью альтернирующего потока событий // Прикладная дискретная математика. – 2009. – № 1. – С. 78–79.

12. *Пагано М., Рыков В.В., Хохлов Ю.С.* Модели телетрафика. – М.: ИНФРА-М, 2019. – 178 с.

13. *Nezhelskaya L., Sidorova E.* Optimal estimation of the states of synchronous generalized flow of events of the second order under its complete observability // Communications in Computer and Information Science. – 2018. – V. 912. – P. 157–171.

14. *Нежельская Л.А., Сидорова Е.Ф.* Оптимальная оценка состояний обобщенного синхронного потока событий второго порядка в условиях неполной наблюдаемости // Вестник Томского государственного университета. Управление, вычислительная техника и информатика. – 2018. – № 45. – С. 30–41.

15. *Горцев А.М., Зуевич В.Л.* Оптимальная оценка состояний асинхронного дважды стохастического потока событий с произвольным числом состояний // Вестник Томского государственного университета. Управление, вычислительная техника и информатика. – 2010. – № 2. – С. 44–65.

16. *Gortsev A.M., Nezhel'skaya L.A.* Estimate of parameters of synchronously alternating Poisson stream of events by the moment method // Telecommunications and Radio Engineering (English translation of Elektrosvyaz and Radiotekhnika). – 1996. – V. 50. – Is. 1. – P. 56–63.

17. *Okamura H., Dohi T., Trivedi K.S.* Markovian arrival process parameter estimation with group data // IEEE/ACM Transactions on Networking (TON). – 2009. – V. 17. – No. 4. – P. 1326–1339.

18. *Leland W.E., Wilson D.V.* High time-resolution measurement and analysis of LAN traffic: implications for LAN interconnection // Proceedings of the IEEE INFOCOM'91, Bal Harbour, FL. – 1991. – P. 1360–1366.

19. *Малинковский Ю.В.* Теория вероятностей и математическая статистика (часть 2. Математическая статистика). – Гомель: УО «ГТУ им. Ф. Скорины», 2004. – 146 с.

20. *Karagiannis T., Faloutsos M.* SELFIS: A tool for self-similarity and long-range dependence analysis // In: 1st Workshop on Fractals and Self-Similarity in Data Mining: Issues and Approaches (in KDD), Edmonton, Canada. – 2002.

21. *Erramilli A., Narayan O., Willinger W.* Experimental queueing analysis with long-range dependent packet traffic // IEEE/ACM Transactions on Networking. – 1996. – V. 4. – No. 2. – P. 209–223.

ОЦЕНИВАНИЕ МЕТОДОМ МОМЕНТОВ ПАРАМЕТРОВ ПЛОТНОСТИ ВЕРОЯТНОСТИ ДЛИТЕЛЬНОСТИ ИНТЕРВАЛА МЕЖДУ СОБЫТИЯМИ РЕКУРРЕНТНОГО ПОЛУСИНХРОННОГО ПОТОКА ВТОРОГО ПОРЯДКА В ОСОБОМ СЛУЧАЕ

Л.А. Нежельская, Д.А. Тумашкина
Томский государственный университет

Введение

В настоящее время при описании и анализе реальных экономических, технических, физических и других процессов и систем часто возникает необходимость применять математический аппарат теории массового обслуживания (ТМО). В связи с бурным развитием информационных технологий важнейшими сферами приложений ТМО являются проектирование и создание цифровых сетей интегрального обслуживания (ЦСИО). Как правило, на практике довольно часто параметры, определяющие поток событий, неизвестны, либо случайным образом изменяются со временем. В этой связи адекватными математическими моделями информационных потоков сообщений, функционирующих в ЦСИО, являются дважды стохастические потоки событий [1–10]. В данных потоках событий случайны не только моменты наступления событий, но и интенсивность потока представляет собой случайный процесс, т.е. имеет место двойная стохастика, что и приводит к изучению дважды стохастических потоков событий.

Однако условия функционирования реальных процессов и систем таковы, что параметры входящих потоков являются ненаблюдаемыми, наблюдаются лишь моменты наступления событий потока. Поэтому существенный интерес представляет решение задач оценки состояний [11–13] и параметров [14–19] дважды стохастического потока по наблюдениям за моментами наступления событий.

В данной работе рассматривается дважды стохастический поток событий с числом состояний, равным двум. В зависимости от того, каким образом происходит переход из состояния в состояние, данные потоки событий можно разделить на следующие три типа: 1) синхронные потоки, переход из состояния в состояние в которых непосредственно зависит от наступления события [20]; 2) асинхронные потоки, переход из состояния в состояние в которых не зависит от того, наступило событие или нет [21]; 3) полусинхронные потоки, у которых для одного состояния справедливо определение первого типа, а для другого состояния – второго типа [22]. Объектом исследования настоящей работы является полусинхронный поток событий второго порядка.

Для рассматриваемого потока задача оптимальной оценки состояний при полной наблюдаемости потока решена в [12,13] и при его неполной наблюдаемости в [23].

В настоящей работе для особого случая задания параметров полусинхронного потока второго порядка приводится явный вид плотности вероятности значений длительности интервала между моментами наступления событий потока, а также находится явный вид совместной плотности вероятности значений длительностей смежных интервалов. Находятся условия рекуррентности потока в рассматриваемом случае. Для рекуррентного потока методом моментов решается задача оценивания параметров плотности распределения.

1. Постановка задачи

Рассматривается стационарный режим функционирования полусинхронного дважды стохастического потока событий второго порядка (далее поток), сопровождающий случайный процесс которого $\lambda(t)$ является кусочно-постоянным с двумя состояниями S_1 и S_2. Длительность пребывания процесса $\lambda(t)$ в первом состоянии определяется случайной величиной $\eta = \min\left(\xi^{(1)}, \xi^{(2)}\right)$, где случайная величина $\xi^{(1)}$ имеет функцию распределения $F_1^{(1)}(t) = 1 - e^{-\lambda_1 t}$, случайная величина $\xi^{(2)}$ – функцию распределения $F_1^{(2)}(t) = 1 - e^{-\alpha_1 t}$; причем $\xi^{(1)}$ и $\xi^{(2)}$ – независимые случайные величины.

В момент наступления события потока процесс $\lambda(t)$ переходит из первого состояния во второе либо с вероятностью $P_1^{(1)}\left(\lambda_2 \mid \lambda_1\right)$, либо с вероятностью $P_1^{(2)}\left(\lambda_2 \mid \lambda_1\right)$ в зависимости от того, какое значение приняла случайная величина η. В момент наступления события потока процесс $\lambda(t)$ остается в первом состоянии либо с вероятностью $P_1^{(1)}\left(\lambda_1 \mid \lambda_1\right)$, либо с вероятностью $P_1^{(2)}\left(\lambda_1 \mid \lambda_1\right)$ в зависимости от значения случайной величины η. Здесь $P_1^{(1)}\left(\lambda_2 \mid \lambda_1\right) + P_1^{(1)}\left(\lambda_1 \mid \lambda_1\right) = 1$, $P_1^{(2)}\left(\lambda_2 \mid \lambda_1\right) + P_1^{(2)}\left(\lambda_1 \mid \lambda_1\right) = 1$. Длительность интервала между событиями потока в первом состоянии процесса $\lambda(t)$ является случайной величиной с функцией распределения $F_1(t) = 1 - e^{-(\lambda_1 + \alpha_1)t}$.

Длительность пребывания процесса $\lambda(t)$ во втором состоянии есть случайная величина с функцией распределения $F_2(t) = 1 - e^{-\alpha_2 t}$. В течение времени пребывания процесса $\lambda(t)$ во втором состоянии имеет место пуассоновский поток событий с параметром λ_2.

В последующем изложении полагается, что имеет место состояние S_i (i-е состояние) процесса $\lambda(t)$, если $\lambda(t) = \lambda_i$, $i = 1,\ 2$; $\lambda_1 > \lambda_2 \geq 0$.

Матрицы инфинитезимальных характеристик процесса $\lambda(t)$ имеют вид

$$\boldsymbol{D}_0 = \begin{Vmatrix} -(\lambda_1+\alpha_1) & 0 \\ \alpha_2 & -(\lambda_2+\alpha_2) \end{Vmatrix},$$

$$\boldsymbol{D}_1 = \begin{Vmatrix} \lambda_1 P_1^{(1)}(\lambda_1 \mid \lambda_1) + \alpha_1 P_1^{(2)}(\lambda_1 \mid \lambda_1) & \lambda_1 P_1^{(1)}(\lambda_2 \mid \lambda_1) + \alpha_1 P_1^{(2)}(\lambda_2 \mid \lambda_1) \\ 0 & \lambda_2 \end{Vmatrix}.$$

Элементами матрицы $\boldsymbol{D}_1$ являются интенсивности переходов процесса $\lambda(t)$ из состояния в состояние с наступлением события. Недиагональные элементы матрицы $\boldsymbol{D}_0$ являются интенсивностями переходов из состояния в состояние без наступления события; в свою очередь, диагональные элементы матрицы $\boldsymbol{D}_0$ – интенсивности выхода процесса $\lambda(t)$ из своих состояний, взятые с противоположным знаком.

На рис. 1 приведен пример одной из реализаций процесса $\lambda(t)$, где $t_1, t_2, \ldots$ – моменты наступления событий потока.

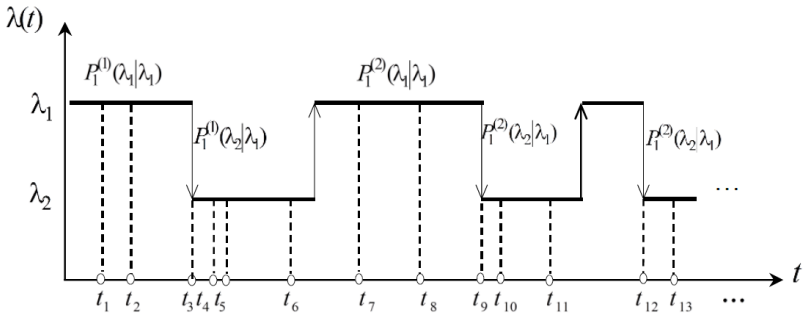


Рис. 1. Полусинхронный поток событий второго порядка

Утверждение. Процесс $\lambda(t)$ является марковским.

Доказательство. Нетрудно показать, что длительность пребывания процесса $\lambda(t)$ в первом состоянии есть случайная величина, распределенная по экспоненциальному закону, с функцией распределения $F_1(t) = 1 - e^{-\left[\lambda_1 P_1^{(1)}(\lambda_2|\lambda_1) + \alpha_1 P_1^{(2)}(\lambda_2|\lambda_1)\right]t}$. В свою очередь, длительность пребывания процесса $\lambda(t)$ во втором состоянии есть случайная величина, распределенная по экспоненциальному закону $F_2(t) = 1 - e^{-\alpha_2 t}$. Отсюда следует формулировка утверждения.

В силу того, что $\lambda(t)$ – принципиально ненаблюдаемый процесс, наблюдаются только моменты времени наступления событий $t_1, t_2, \ldots$, то $\lambda(t)$ является скрытым марковским процессом или ненаблюдаемым сопровождающим марковским процессом.

В моменты времени $t_1, t_2, \ldots, t_k, \ldots$ последовательность $\{\lambda(t_k)\}$ представляет собой вложенную цепь Маркова.

Обозначим $\tau_k = t_{k+1} - t_k$, $k = 1, 2, \ldots$ – значение длительности интервала между соседними событиями потока, $p(\tau)$ – плотность вероятности значений длительности интервала между соседними событиями в потоке. Поскольку рассматривается стационарный режим функционирования потока, $p(\tau_k) = p(\tau)$ для всех $k = 1, 2, \ldots$, $\tau \geq 0$. В этой связи без ограничения общности момент наступления события t_k можно положить равным нулю, т.е. момент наступления события есть $\tau = 0$.

Для особого случаев задания параметров полусинхронного потока второго порядка, а именно, когда коэффициент $(\lambda_1 + \alpha_1) - (\lambda_2 + \alpha_2) = 0$, требуется найти явный вид плотности вероятности значений длительности интервала между моментами наступления событий потока, а также явный вид совместной плотности вероятности значений длительностей смежных интервалов; получить условия рекуррентности потока, для

каждого из выписанных условий найти явный вид плотности вероятности значений длительности интервала между соседними событиями в рекуррентном потоке и оценить методом моментов параметры плотности распределения для рекуррентного потока.

2. Вид плотности вероятности $p(\tau)$ и совместной плотности вероятности $p(\tau_1,\tau_2)$

Рассмотрим особый случай задания параметров рассматриваемого потока, когда коэффициент $(\lambda_1+\alpha_1)-(\lambda_2+\alpha_2)=0$.

Рассмотрим $p_{ij}(\tau)$ – условную вероятность того, что на интервале $(0,\tau)$ нет событий потока и в момент времени τ значение процесса $\lambda(\tau)=\lambda_j$ при условии, что в момент времени $\tau=0$ значение процесса $\lambda(0)=\lambda_i$, $i,j=1,2$ [24].

Лемма 1. Условные вероятности $p_{ij}(\tau)$, $i,j=1,2$, в коррелированном полусинхронном потоке событий второго порядка в случае $(\lambda_1+\alpha_1)-(\lambda_2+\alpha_2)=0$ имеют следующий вид

$$p_{11}(\tau)=e^{-((\lambda_1+\alpha_1)\tau}, \; p_{12}(\tau)=0, \; p_{22}(\tau)=e^{-(\lambda_1+\alpha_1)\tau},$$
$$p_{21}(\tau)=(\lambda_1+\alpha_1-\lambda_2)\tau e^{-(\lambda_1+\alpha_1)\tau}, \; \tau\ge 0. \qquad (1)$$

Лемма 2. Плотности вероятностей $\tilde{p}_{ij}(\tau)$, $i,j=1,2$, в коррелированном полусинхронном потоке событий второго порядка в случае $(\lambda_1+\alpha_1)-(\lambda_2+\alpha_2)=0$ определяются формулами

$$\tilde{p}_{11}(\tau)=\left[\lambda_1 P_1^{(1)}(\lambda_1\mid\lambda_1)+\alpha_1 P_1^{(2)}(\lambda_1\mid\lambda_1)\right]e^{-(\lambda_1+\alpha_1)\tau},$$
$$\tilde{p}_{12}(\tau)=\left[\lambda_1 P_1^{(1)}(\lambda_2\mid\lambda_1)+\alpha_1 P_1^{(2)}(\lambda_2\mid\lambda_1)\right]e^{-(\lambda_1+\alpha_1)\tau},$$
$$\tilde{p}_{21}(\tau)=\alpha_2\left[\lambda_1 P_1^{(1)}(\lambda_1\mid\lambda_1)+\alpha_1 P_1^{(2)}(\lambda_1\mid\lambda_1)\right]\tau e^{-(\lambda_1+\alpha_1)\tau},$$
$$\tilde{p}_{22}(\tau)=\alpha_2\left[\lambda_1 P_1^{(1)}(\lambda_2\mid\lambda_1)+\alpha_1 P_2^{(2)}(\lambda_2\mid\lambda_1)\right]\tau e^{-(\lambda_1+\alpha_1)\tau}+\lambda_2 e^{-(\lambda_1+\alpha_1)\tau}, \; \tau\ge 0. \qquad (2)$$

Лемма 3. Вероятности перехода p_{ij}, $i,j=1,2$, в коррелированном полусинхронном потоке событий второго порядка в случае $(\lambda_1+\alpha_1)-(\lambda_2+\alpha_2)=0$ имеют вид

$$p_{11}=\frac{\lambda_1 P_1^{(1)}(\lambda_1\mid\lambda_1)+\alpha_1 P_1^{(2)}(\lambda_1\mid\lambda_1)}{\lambda_1+\alpha_1}, \; p_{12}=\frac{\lambda_1 P_1^{(1)}(\lambda_2\mid\lambda_1)+\alpha_1 P_1^{(2)}(\lambda_2\mid\lambda_1)}{\lambda_1+\alpha_1},$$
$$p_{21}=\frac{\alpha_2\left[\lambda_1 P_1^{(1)}(\lambda_1\mid\lambda_1)+\alpha_1 P_1^{(2)}(\lambda_1\mid\lambda_1)\right]}{(\lambda_1+\alpha_1)^2},$$
$$p_{22}=\frac{\alpha_2\left[\lambda_1 P_1^{(1)}(\lambda_2\mid\lambda_1)+\alpha_1 P_1^{(2)}(\lambda_2\mid\lambda_1)\right]}{(\lambda_1+\alpha_1)^2}+\frac{\lambda_2}{\lambda_1+\alpha_1}. \qquad (3)$$

Введем в рассмотрение $\pi_i(0)$ – условную стационарную вероятность того, что процесс $\lambda(\tau)$ в момент времени $\tau=0$ находится в состоянии S_i при условии, что в момент времени $\tau=0$ наступило событие потока, $i,j=1,2$, $\pi_1(0)+\pi_2(0)=1$.

Лемма 4. Условные финальные вероятности $\pi_i(0)$, $i=1,2$, в коррелированном полусинхронном потоке событий второго порядка определяются выражениями

$$\pi_1(0)=\frac{\alpha_2\left[\lambda_1 P_1^{(1)}(\lambda_1\mid\lambda_1)+\alpha_1 P_1^{(2)}(\lambda_1\mid\lambda_1)\right]}{(\lambda_1+\alpha_1)^2-\lambda_2\left[\lambda_1 P_1^{(1)}(\lambda_1\mid\lambda_1)+\alpha_1 P_1^{(2)}(\lambda_1\mid\lambda_1)\right]},$$

$$\pi_2(0)=\frac{(\lambda_1+\alpha_1)\left[\lambda_1 P_1^{(1)}(\lambda_2\mid\lambda_1)+\alpha_1 P_1^{(2)}(\lambda_2\mid\lambda_1)\right]}{(\lambda_1+\alpha_1)^2-\lambda_2\left[\lambda_1 P_1^{(1)}(\lambda_1\mid\lambda_1)+\alpha_1 P_1^{(2)}(\lambda_1\mid\lambda_1)\right]},\ \pi_1(0)+\pi_2(0)=1. \qquad (4)$$

Леммы 2 и 4 позволяют сформулировать следующую теорему.

Теорема 1. Плотность вероятности значений длительности интервала между соседними событиями в коррелированном полусинхронном потоке второго порядка для случая $(\lambda_1+\alpha_1)-(\lambda_2+\alpha_2)=0$ имеет вид

$$p(\tau)=\left[(\lambda_1+\alpha_1)-\pi_2(0)(\lambda_1+\alpha_1-\lambda_2)(1-(\lambda_1+\alpha_1)\tau)\right]e^{-(\lambda_1+\alpha_1)\tau},\ \tau\ge 0, \qquad (5)$$

где $\pi_2(0)$ определена в (4).

Доказательство. В силу того, что процесс $\lambda(t)$ обладает марковским свойством, если его эволюцию рассматривать, начиная с момента времени t_k, $k=1,2,\ldots$, наступления события потока, плотность вероятности $p(\tau)$ значений длительности интервала между соседними событиями в рассматриваемом потоке определяется в виде

$$p(\tau)=\sum_{i=1}^{2}\pi_i(0)\sum_{j=1}^{2}\tilde{p}_{ij}(\tau),\ \tau\ge 0. \qquad (6)$$

Подставляя $\tilde{p}_{ij}(\tau)$, определенные формулами (2), в (6), затем, подставляя в (6) выражения (4) для $\pi_i(0)$, $i,j=1,2$, в результате необходимых преобразований приходим к (5). *Теорема 1 доказана.*

Рассмотрим два соседних временных интервала (t_1,t_2), (t_2,t_3) со значениями длительностей $\tau_1=t_2-t_1$ и $\tau_2=t_3-t_2$ соответственно. При этом обозначим совместную плотность вероятности $p(\tau_1,\tau_2)$, $\tau_1\ge 0$, $\tau_2\ge 0$ [24].

Теорема 2. Полусинхронный поток событий второго порядка в общем случае является коррелированным и совместная плотность вероятности значений длительностей смежных интервалов для случая $(\lambda_1+\alpha_1)-(\lambda_2+\alpha_2)=0$ имеет вид

$$p(\tau_1,\tau_2)=p(\tau_1)p(\tau_2)-\frac{\lambda_2\pi_1(0)\pi_2(0)\left(\lambda_1 P_1^{(1)}(\lambda_2\mid\lambda_1)+\alpha_1 P_1^{(2)}(\lambda_2\mid\lambda_1)\right)}{\lambda_1+\alpha_1}\times$$
$$\times(\lambda_1+\alpha_1-\lambda_2)(1-(\lambda_1+\alpha_1)\tau_1)(1-(\lambda_1+\alpha_1)\tau_2)e^{-(\lambda_1+\alpha_1)(\tau_1+\tau_2)},\ \tau_1\ge 0,\ \tau_2\ge 0, \qquad (7)$$

где $p(\tau_k)$ определены в (5) для $\tau=\tau_k$, $k=1,2$; $\pi_1(0)$, $\pi_2(0)$ – в (4).

Доказательство. Последовательность моментов наступления событий t_1, t_2, … образует вложенную цепь Маркова $\{\lambda(t_k)\}$, следовательно, для $p(\tau_1,\tau_2)$ справедливо

$$p(\tau_1,\tau_2)=\sum_{i=1}^{2}\pi_i(0)\sum_{j=1}^{2}\tilde{p}_{ij}(\tau_1)\sum_{k=1}^{2}\tilde{p}_{jk}(\tau_2), \qquad (8)$$

где $\tilde{p}_{ij}(\tau_1)$, $\tilde{p}_{ij}(\tau_2)$ – плотности вероятности, которые соответствуют переходным вероятностям $p_{ij}(\tau_1)$, $p_{ij}(\tau_2)$ и вычисляются по формулам (2) при $\tau=\tau_1$ и $\tau=\tau_2$. Подставляя в (8) выражения для $\tilde{p}_{ij}(\tau_1)$, $\tilde{p}_{ij}(\tau_2)$, затем $\pi_i(0)$, $i=1,2$, определенные в (4), проделывая достаточно трудоемкие преобразования, получим (7). *Теорема 2 доказана.*

3. Условия рекуррентности полусихронного потока событий второго порядка

Рассмотрим случаи, когда поток является рекуррентным. При этом для каждого из случаев запишем вид плотности исходя из выражения (5). Анализируя (7), заметим, что совместная плотность факторизуется $p(\tau_1,\tau_2)=p(\tau_1)p(\tau_2)$, если:

1. $\lambda_1 + \alpha_1 - \lambda_2 = 0$, получим плотность распределения простейшего потока $p(\tau) = \lambda_2 e^{-\lambda_2 \tau}$, $\tau \geq 0$;

2. $-\lambda_2 \pi_1(0)\pi_2(0)\left[\lambda_1 P_1^{(1)}(\lambda_2 \mid \lambda_1) + \alpha_1 P_1^{(2)}(\lambda_2 \mid \lambda_1)\right] = 0$, имеем следующие случаи:

2.1. $\lambda_2 = 0$, $p(\tau) = \left[z - a(1 - z\tau)\right]e^{-z\tau}$, $\tau \geq 0$, где $a = \lambda_1 P_1^{(1)}(\lambda_2 \mid \lambda_1) + \alpha_1 P_1^{(2)}(\lambda_2 \mid \lambda_1)$, $z = \alpha_2$;

2.2. $\pi_1(0) = 0$, учитывая выражение (4) для $\pi_1(0)$, имеем $\lambda_1 P_1^{(1)}(\lambda_1 \mid \lambda_1) + \alpha_1 P_1^{(2)}(\lambda_1 \mid \lambda_1) = 0$, тогда возможны следующие ситуации:

2.2.1. $\alpha_1 = P_1^{(1)}(\lambda_1 \mid \lambda_1) = 0$, $p(\tau) = \left[z - a(1 - z\tau)\right]e^{-z\tau}$, $\tau \geq 0$, где $z = \lambda_1$, $a = \alpha_2$;

2.2.2. $P_1^{(1)}(\lambda_1 \mid \lambda_1) = P_1^{(2)}(\lambda_1 \mid \lambda_1) = 0$, $p(\tau) = \left[z - a(1 - z\tau)\right]e^{-z\tau}$, $\tau \geq 0$, где $z = \lambda_1 + \alpha_1$, $a = \alpha_2$;

2.3. $\pi_2(0) = 0$, учитывая выражение (4) для $\pi_2(0)$, имеем следующие случаи:

2.3.1. $\alpha_1 = P_1^{(1)}(\lambda_2 \mid \lambda_1) = 0$, $p(\tau) = \lambda_1 e^{-\lambda_1 \tau}$, $\tau \geq 0$;

2.3.2. $P_1^{(1)}(\lambda_2 \mid \lambda_1) = P_1^{(2)}(\lambda_2 \mid \lambda_1) = 0$, $p(\tau) = ze^{-z\tau}$, $\tau \geq 0$, где $z = \lambda_1 + \alpha_1$;

Ставится задача оценивания параметров плотности распределения

$$p(\tau) = \left[z - a(1 - z\tau)\right]e^{-z\tau}, \ \tau \geq 0 \tag{9}$$

для рекуррентного потока в случаях потока событий, отличного от простейшего, т.е. для условий 2.1, 2.2.1, 2.2.2.

4. Оценивание параметров распределения методом моментов

Рассмотрим статистики $C_l = 1 \Big/ n \sum_{k=1}^{n} \tau_k^l$, где $\tau_k = t_{k+1} - t_k$ – значение длительности интервала между моментами t_k и t_{k+1} наступления соседних событий в потоке. Пусть имеется выборка $\tau_1, \tau_2, \ldots, \tau_n$ из распределения $p(\tau \mid z, a)$, зависящего от двух неизвестных параметров z, a. Пусть $M(\tau^l) = \int_0^\infty \tau^l p(\tau \mid z, a) d\tau$ – начальный теоретический момент l-го порядка, являющийся функцией от неизвестных параметров. Тогда он близок к соответствующему выборочному моменту $\overline{\tau}^l$ – статистике C_l. Для первых двух начальных моментов вместо приближенных запишем точные равенства, в которых вместо истинных значений параметров подставлены их оценки

$$M(\tau^l) = C_l, \ l = 1, 2. \tag{10}$$

Учитывая вид плотности (9), имеем $M(\tau^l) = \dfrac{l!}{z^l}\left[1 + a\dfrac{l}{z}\right]$, $l = 1, 2$. Тогда система уравнений моментов (10) примет следующий вид

$$(z + a) \Big/ z^2 = C_1, \ 2(z + 2a) \Big/ z^3 = C_2. \tag{11}$$

Полученная система (11) сводится к квадратному уравнению $C_2 z^2 - 4C_1 z + 2 = 0$, решая которое, находим

$$\hat{z}^{(1)} = \left(2C_1 - \sqrt{4C_1^2 - 2C_2}\right) \Big/ C_2, \ \hat{z}^{(2)} = \left(2C_1 + \sqrt{4C_1^2 - 2C_2}\right) \Big/ C_2. \tag{12}$$

При этом должны выполняться следующие условия $\hat{z}^{(1)}\hat{z}^{(2)} = 2/C_2 > 0$, $\hat{z}^{(1)} + \hat{z}^{(2)} = 4C_1/C_2 > 0$, $4C_1^2 - 2C_2 \geq 0$.

Подставляя полученные оценки $\hat{z}^{(1)}$ и $\hat{z}^{(2)}$ в первое уравнение системы (11), находим

$$\hat{a}^{(1)} = C_1\left(\hat{z}^{(1)}\right)^2 - \hat{z}^{(1)},\ \hat{a}^{(2)} = C_1\left(\hat{z}^{(2)}\right)^2 - \hat{z}^{(2)}. \qquad (13)$$

Естественно образом возникает вопрос об осуществлении выбора из пар $\left\{\hat{z}^{(1)}, \hat{a}^{(1)}\right\}$ и $\left\{\hat{z}^{(2)}, \hat{a}^{(2)}\right\}$ в качестве решения задачи оценивания. С учетом явного вида параметров и в силу условия $\left(\lambda_1 + \alpha_1\right) - \left(\lambda_2 + \alpha_2\right) = 0$ нетрудно показать, что оценки $\hat{z}$ и $\hat{a}$ параметров z и a должны удовлетворять условиям $\hat{a} > 0$ и $\hat{z} - \hat{a} > 0$. В ходе аналитической проверки полученных пар $\left\{\hat{z}^{(1)}, \hat{a}^{(1)}\right\}$ и $\left\{\hat{z}^{(2)}, \hat{a}^{(2)}\right\}$ на выполнение указанных условий, сделаны следующие выводы. Если $\hat{a}^{(1)} > 0$, $\hat{a}^{(2)} > 0$ и $C_2/2 \le C_1^2 < 2C_2/3$, то пары равноправны и в качестве оценки выбирается любая; если справедливо $\hat{a}^{(1)} > 0$ и $C_1^2 \ge 2C_2/3$, то в качестве решения задачи оценивания выбирается единственная пара $\left\{\hat{z}^{(1)}, \hat{a}^{(1)}\right\}$.

Замечание. Для особого случая задания параметров потока – $\left(\lambda_1 + \alpha_1\right) - \left(\lambda_2 + \alpha_2\right) = 0$, система из трех и более уравнений моментов является несовместной.

Утверждение. Оценить параметры λ_i, α_i, $P_1^{(1)}\left(\lambda_j \middle| \lambda_1\right)$, $P_1^{(2)}\left(\lambda_j \middle| \lambda_1\right), i, j = 1,2$, полусинхронного потока событий второго порядка, имея информацию только о плотности $p(\tau)$, не представляется возможным.

5. Результаты статистических экспериментов

С целью установления качества оценивания разработан алгоритм вычисления оценок параметров плотности, состоящий из двух этапов. На первом этапе производится имитационное моделирование полусинхронного потока событий второго порядка [25]. На втором этапе реализации алгоритма непосредственно вычисляются оценки параметров $\hat{z}$, $\hat{a}$ по полученным формулам (12), (13), а также находятся выборочные средние значения оценки $\hat{M}(\hat{\theta}) = 1/N \sum_{k=1}^{N} \hat{\theta}^{(k)}$ и выборочные вариации $\hat{V}\left(\hat{\theta}\right) = 1/N \sum_{k=1}^{N} \left(\hat{\theta}^{(k)} - \theta\right)^2$, где $\theta \in \{z, a\}$, $\hat{\theta} \in \{\hat{z}, \hat{a}\}$.

В *первом статистическом эксперименте* рассматривается поток при выполнении условия рекуррентности 2.1. В данном эксперименте иллюстрируется установление стационарного режима. Полагается $\lambda_1 = 5$, $\alpha_1 = 1$, $\alpha_2 = 6$, $P_1^{(1)}\left(\lambda_1 \middle| \lambda_1\right) = P_1^{(2)}\left(\lambda_1 \middle| \lambda_1\right) = 0.6$, $P_1^{(1)}\left(\lambda_2 \middle| \lambda_1\right) = P_1^{(2)}\left(\lambda_2 \middle| \lambda_1\right) = 0.4$, $N = 100$. Рассматривается зависимость $\hat{M}\left(\hat{z}\right)$, $\hat{V}\left(\hat{z}\right)$ от значений времени моделирования $T_m = 100,...,1200$ ед. времени с шагом 100. Результаты отражены в табл. 1, где последняя строка – истинное значение оцениваемого параметра.

Таблица 1

Зависимость $\hat{M}(\hat{z})$, $\hat{V}(\hat{z})$ от значений T_m

T_m	100	200	…	600	700	800	…	1100	1200
$\hat{M}(\hat{z})$	5.3576	5.6249	…	5.8612	5.9657	5.9642	…	5.9656	5.9657
$\hat{V}(\hat{z})$	0.48257	0.32811	…	0.00979	0.00133	0.00128	…	0.00131	0.00132
z	6	6	…	6	6	6	…	6	6

Анализ численных результатов, приведенных в табл. 1, показывает улучшение качества оценивания с увеличением времени моделирования (в смысле уменьшения значения выборочной вариации), а также иллюстрирует установление стационарного режима для $T_m \geq 700$ ед. времени.

Далее в рамках данного эксперимента полагается время моделирования $T_m = 700$ ед. времени и при $N = 100$, параметре $\alpha_1 = 1$, вероятностях $P_1^{(1)}(\lambda_1|\lambda_1) = P_1^{(2)}(\lambda_1|\lambda_1) = 0.6$, $P_1^{(1)}(\lambda_2|\lambda_1) = P_1^{(2)}(\lambda_2|\lambda_1) = 0.4$ устанавливается зависимость $\hat{M}(\hat{\theta})$, $\hat{V}(\hat{\theta})$ от значений параметра $\lambda_1 = 1,...,5$ с шагом 0.5, и, соответственно, параметра α_2, определяемого по формуле $z = \alpha_2 = \lambda_1 + \alpha_1$. Результаты эксперимента отражены в табл. 2, где две нижние строки – истинные значения оцениваемых параметров, и проиллюстрированы на рис. 2.

Таблица 2

Зависимость $\hat{M}(\hat{\theta})$, $\hat{V}(\hat{\theta})$ от значений λ_1

λ_1	1	1.5	2	2.5	3	3.5	4	4.5	5
$\hat{M}(\hat{z})$	2.3828	2.8452	3.2998	3.7595	4.2147	4.6752	4.8778	5.4126	5.9655
$\hat{V}(\hat{z})$	0.1542	0.1214	0.0894	0.0627	0.0531	0.0255	0.0089	0.0029	0.0013
$\hat{M}(\hat{a})$	0.4389	0.6755	0.9228	1.1673	1.4166	1.6965	1.8837	2.2874	2.3713
$\hat{V}(\hat{a})$	0.1424	0.1184	0.0813	0.0602	0.0444	0.0198	0.0082	0.0027	0.0011
z	2	2.5	3	3.5	4	4.5	5	5.5	6
a	0.8	1	1.2	1.4	1.6	1.8	2	2.2	2.4

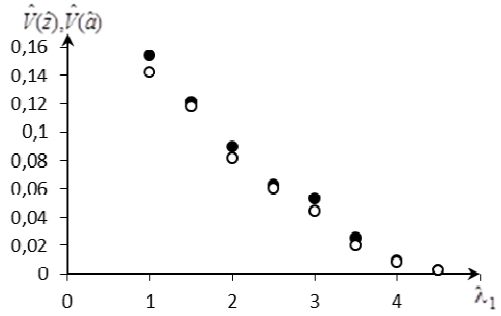


Рис. 2. График зависимости $\hat{V}(\hat{z})$ (черный маркер) и $\hat{V}(\hat{a})$ (белый маркер) от значений параметра λ_1

Проводя анализ приведенных в табл. 2, а также проиллюстрированных на рис. 2, численных результатов, заметим, что с увеличением параметра λ_1, и, соответственно, с увеличением α_2, значение $\hat{V}(\hat{\theta})$ уменьшается, т.к. условия различимости состояний улучшаются. Другими словами, при больших значениях $\lambda_1 - \lambda_2$, имеем меньшие значения $\hat{V}(\hat{\theta})$. А также, анализируя матрицы $\boldsymbol{D}_0$ и $\boldsymbol{D}_1$, заметим, что оцениваемые параметры z, a, являются элементами $\boldsymbol{D}_0$, $\boldsymbol{D}_1$, соответственно. Следовательно, данные параметры являются характеристиками интенсивности перехода процесса $\lambda(t)$ из состояния S_2 в S_1 и наоборот. Следует отметить, что в данном случае процесс $\lambda(t)$ преимущественно находится в состоянии S_1; также отметим отсутствие событий в состоянии S_2, т.к. нет пуассоновского потока ($\lambda_2 = 0$).

Во *втором статистическом эксперименте* рассматривается рекуррентный поток при выполнении условия 2.2.2, т.е. положим $P_1^{(1)}(\lambda_1|\lambda_1) = P_1^{(2)}(\lambda_1|\lambda_1) = 0$. При времени

моделирования $T_m = 700$ ед. времени, $N = 100$, параметрах потока $\lambda_2 = 1$, $\alpha_1 = 1$, вероятностях $P_1^{(1)}(\lambda_2|\lambda_1) = P_1^{(2)}(\lambda_2|\lambda_1) = 1$ рассматривается зависимость $\hat{M}(\hat{\theta})$, $\hat{V}(\hat{\theta})$ от значений параметра $a = \alpha_2 = 2,3,4,5,6$, и, соответственно, параметра λ_1, который определяется по формуле $\lambda_1 = \lambda_2 + \alpha_2 - \alpha_1$. Результаты данного эксперимента отражены в табл. 3, где две нижние строки – истинные значения оцениваемых параметров.

Таблица 3

Зависимость $\hat{M}(\hat{\theta})$, $\hat{V}(\hat{\theta})$ от значений α_2

α_2	**2**	**3**	**4**	**5**	**6**
λ_1	2	3	4	5	6
$\hat{M}(\hat{z})$	3.3274	4.1728	5.0599	6.0396	6.9824
$\hat{V}(\hat{z})$	0.1148	0.0321	0.0042	0.0024	0.0008
$\hat{M}(\hat{a})$	1.6576	3.1635	4.0569	4.9665	5.9817
$\hat{V}(\hat{a})$	0.1032	0.0299	0.0037	0.0022	0.0007
z	3	4	5	6	7
a	2	3	4	5	6

Анализируя приведенные в табл. 3 численные результаты, можно заметить, что увеличение параметра α_2, и, соответственно, увеличение λ_1, положительно сказывается на условиях различимости состояний, потому значение $\hat{V}(\hat{\theta})$ уменьшается.

Заключение

В настоящей работе рассмотрен полусинхронный дважды стохастический поток событий второго порядка. Для особого случая задания параметров потока получен явный вид одномерной плотности вероятности значений длительности интервала между моментами наступления соседних событий потока, а также найден явный вид совместной плотности вероятности значений длительностей смежных интервалов. Найдены условия рекуррентности потока в рассматриваемом случае. Для рекуррентного потока решена задача оценивания параметров плотности распределения методом моментов. Выражения для оценок параметров получены в явном виде, что позволяет производить вычисления без привлечения численных методов. Алгоритм вычисления оценок параметров плотности реализован на языке программирования C# в среде Visual Studio 2013. С целью установления качества оценивания проведены статистические эксперименты, численные результаты которых не противоречат физической интерпретации и иллюстрируют достаточно малую величину выборочной вариации.

ЛИТЕРАТУРА

1. *Башарин Г.П., Кокотушкин В.А., Наумов В.А.* О методе эквивалентных замен расчета фрагментов сетей связи. Ч.1 // Изв. АН СССР. Техническая кибернетика. – 1979. – № 6. – С. 92–99.

2. *Башарин Г.П., Кокотушкин В.А., Наумов В.А.* О методе эквивалентных замен расчета фрагментов сетей связи. Ч.2 // Изв. АН СССР. Техническая кибернетика. – 1980. – № 1. – С. 55–61.

3. *Neuts M.F.* A versatile Markov point process // Journal of Applied Probability. – 1979. – V. 16. – P. 764–779.

4. *Cox D.R.* The analysis of non–Markovian stochastic processes by the inclusion of supplementary variables // Proceedings of the Cambridge Philosophical Society. – 1955. – V. 51. – No. 3. – P. 433–441.

5. *Kingman J.F.C.* On doubly stochastic Poisson process // Proceedings of the Cambridge Philosophical Society. – 1964. – V. 60. – Is. 4. – P. 923–930.

6. *Lucantoni D.M.* New results on the single server queue with a bath markovian arrival process // Communications in Statistics Stochastic Models. – 1991. – V. 7. – P. 1–46.

7. *Дудин А.Н., Клименок В.И.* Системы массового обслуживания с коррелированными потоками. Минск: Изд-во Белорусского государственного университета, 2000. – 175 с.

8. *Basharin G.P., Gaidamaka Yu.V., Samouylov K.E.* Mathematical theory of teletraffic and its application to the analysis of multiservice communication of the next generation networks // Automatic Control and Computer Sciences. – 2013. – V. 47. – No. 2. – P. 62–69.

9. *Klimenok V., Dudin A., Vishnevsky V.* Tandem queueing system with correlated input and cross-traffic // Communications in Computer and Information Science. – 2013. – V. 370. – P. 416–425.

10. *Vishnevsky V.M., Semenova O.V.* Polling systems: theory and applications for broadband wireless networks. – London: Academic Publishing, 2012. – 316 p.

11. *Nezhelskaya L.* Optimal state estimation in modulated MAP event flows with unextendable dead time // Communications in Computer and Information Science. – 2014. – V. 487. – P. 342–350.

12. *Nezhelskaya L., Tumashkina D.* Optimal state estimation of semi-synchronous event flow of the second order under its complete observability // Communications in Computer and Information Science. – 2018. – V. 912. – P. 93–105.

13. *Нежельская Л.А., Тумашкина Д.А.* Оптимальное оценивание состояний полусинхронного потока событий второго порядка // Труды Томского государственного университета. Серия физико-математическая. Томск: Издательский Дом ТГУ. – 2017. – Т. 301. – С. 97–105.

14. *Nezhel'skaya L.* Probability density function for modulated MAP event flows with unextendable dead time // Communications in Computer and Information Science. – 2015. – V. 564. – P. 141–151.

15. *Gortsev A.M., Nezhel'skaya L.A.* Estimate of parameters of synchronously alternating Poisson stream of events by the moment method // Telecommunications and Radio Engineering. – 1996. – V. 50. – Is. 1. – P. 56–63.

16. *Gortsev A.M., Klimov I.S.* Estimation of the parameters of an alternating Poisson stream of events // Telecommunications and Radio Engineering. – 1993. – V. 48. – Is. 10. – P. 40–45.

17. *Горцев А.М., Нежельская Л.А.* Оценивание длительности мертвого времени и параметров синхронного альтерниирующего потока событий // Вестник Томского государственного университета. – 2003. – № S6. – С. 232–239.

18. *Василевская Т.П., Горцев А.М., Нежельская Л.А.* Оценивание длительности мертвого времени и параметров синхронного альтерниирующего потока с проявлением либо непроявлением событий // Вестник Томского государственного университета. – 2004. – № S9–2. – С. 129–138.

19. *Горцев А.М., Нежельская Л.А.* Оценивание параметров асинхронного потока с инициированием лишних событий методом моментов // Вестник Томского государственного университета. – 2006. – № S18. – С. 267–273.

20. *Горцев А.М., Нежельская Л.А.* Оценивание параметров синхронного дважды стохастического потока событий методом моментов // Вестник Томского государственного университета. – 2002. – № S1-1. – С. 24–29.

21. *Леонова М.А., Нежельская Л.А.* Вероятность ошибки при оценивании состояний обобщенного асинхронного потока событий // Вестник Томского государственного университета. Управление, вычислительная техника и информатика. – 2012. – № 2 (19). – С. 88–101.

22. *Калягин А.А., Нежельская Л.А.* Сравнение МП- и ММ- оценок длительности мертвого времени в обобщенном полусихронном потоке событий // Вестник Томского государственного университета. Управление, вычислительная техника и информатика. – 2015. – № 3 (32). – С. 23–32.

23. *Нежельская Л.А., Тумашкина Д.А.* Оптимальное оценивание состояний полусинхронного потока событий второго порядка при непродлевающемся мертвом времени // Вестник Томского государственного университета. Управление, вычислительная техника и информатика. – 2019. – № 46. – С. 73–82.

24. *Нежельская Л.А.* Совместная плотность вероятностей длительности интервалов модулированного MAP-потока событий и условия рекуррентности потока // Вестник Томского государственного университета. Управление, вычислительная техника и информатика. – 2015. – №1 (30). – С. 57–67.

25. *Нежельская Л.А., Тумашкина Д.А.* Имитационная модель полусинхронного потока второго порядка // Труды Томского государственного университета. Серия физико-математическая. Томск: Издательский Дом ТГУ, 2016. – Т. 299. – С. 109–114.

АПОСТЕРИОРНЫЕ ВЕРОЯТНОСТИ СОСТОЯНИЙ ОБОБЩЁННОГО MAP-ПОТОКА СОБЫТИЙ С ПРОИЗВОЛЬНЫМ ЧИСЛОМ СОСТОЯНИЙ

А.В. Кеба, Л.А. Нежельская

Томский государственный университет

Введение

В связи с интенсивным развитием компьютерных систем, расширением информационных сетей связи формируется широкая сфера применения аппарата теории массового обслуживания (ТМО), а именно: в логистической, экономической, финансовой областях, в планировании и организации процессов производства, обслуживания и потребления, в транспортных сетях и других отраслях. ТМО представляет собой раздел теории случайных процессов и занимается исследованием систем массового обслуживания (СМО).

Классическая модель входящего потока событий, являющегося основным элементом СМО, в настоящее время перестала соответствовать информационным потокам сообщений в режиме реального времени в результате стремительного развития техники, телекоммуникационных, спутниковых и компьютерных сетей связи. Тем самым требования практики стали причиной рассмотрения дважды стохастических потоков в виде математической модели реальных информационных потоков сообщений. Такие потоки характеризуются случайностью моментов времени наступления событий в потоке, представлением интенсивности потока как случайного процесса и подразделяются на два класса: 1) интенсивность потоков есть непрерывный случайный процесс [1]; 2) интенсивность потоков представляет собой кучно-постоянный случайный процесс с конечным числом состояний [2,3]. Потоки второго класса, к которым относится обобщённый MAP-поток событий с произвольным числом состояний [4], в работе [2] получили название MC-потоков (Markov chain).

При исследовании дважды стохастических потоков событий выделяют два основных раздела задач, базой для которых служат наблюдения за моментами времени наступления событий: 1) оценка состояний потока событий [5–7]; 2) оценка параметров потока [8–10].

В данной работе находятся в явном виде апостериорные и априорные вероятности состояний обобщённого MAP-потока событий, являющиеся основой оптимального оценивания состояний потока.

1. Постановка задачи

Рассматривается обобщённый MAP-поток событий (далее поток), сопровождающий процесс $\lambda(t)$ которого является случайным кусочно-постоянным ненаблюдаемым процессом с n состояниями: $S_1, S_2,\ldots, S_n$. Если $\lambda(t)=\lambda_i$, $i=\overline{1,n}$, то будем говорить, что имеет место i-е состояние (S_i) процесса $\lambda(t)$, причём $\lambda_1>\lambda_2>\ldots>\lambda_n>0$.

Длительность пребывания процесса $\lambda(t)$ в i-м состоянии ($i=\overline{1,n}$) определяется случайной величиной, распределённой по экспоненциальному закону, с функцией распределения $F_i(t)=1-e^{-\lambda_i t}$, $i=\overline{1,n}$. Если процесс $\lambda(t)$ в момент времени t находится в i-м состоянии, т.е. $\lambda(t)=\lambda_i$, то на полуинтервале $\left[t;t+\Delta t\right)$, где Δt – достаточно малая величина, в момент окончания i-го состояния происходит розыгрыш состояния:

1) наступает событие потока, и процесс $\lambda(t)$ переходит из i-го состояния в j-е состояние с вероятностью $P_1\left(\lambda_j\,|\,\lambda_i\right)$, $i,j=\overline{1,n}$;
2) не наступает событие потока, и процесс $\lambda(t)$ переходит из i-го состояния в j-е состояние с вероятностью $P_0\left(\lambda_j\,|\,\lambda_i\right)$, $i,j=\overline{1,n}$.

При этом выполняется условие нормировки

$$\sum_{j=1}^{n}P_0\left(\lambda_j\,|\,\lambda_i\right)+\sum_{j=1}^{n}P_1\left(\lambda_j\,|\,\lambda_i\right)=1,\ i=\overline{1,n}.$$

Утверждение 1. Процесс $\lambda(t)$ для обобщённого MAP-потока событий с произвольным числом состояний является скрытым марковским процессом.

Замечание. Обобщённый MAP-поток событий с произвольным числом состояний отличается от классического MAP-потока введением вероятности $P_0\left(\lambda_i\,|\,\lambda_i\right)\neq 0$, $i=\overline{1,n}$, перехода процесса $\lambda(t)$ из i-го состояния в i-е без наступления события потока.

Одна из реализаций процесса $\lambda(t)$ и обобщённого MAP-потока событий с n состояниями изображена на рис. 1, где $t_1, t_2,\ldots$ – моменты наступления событий.

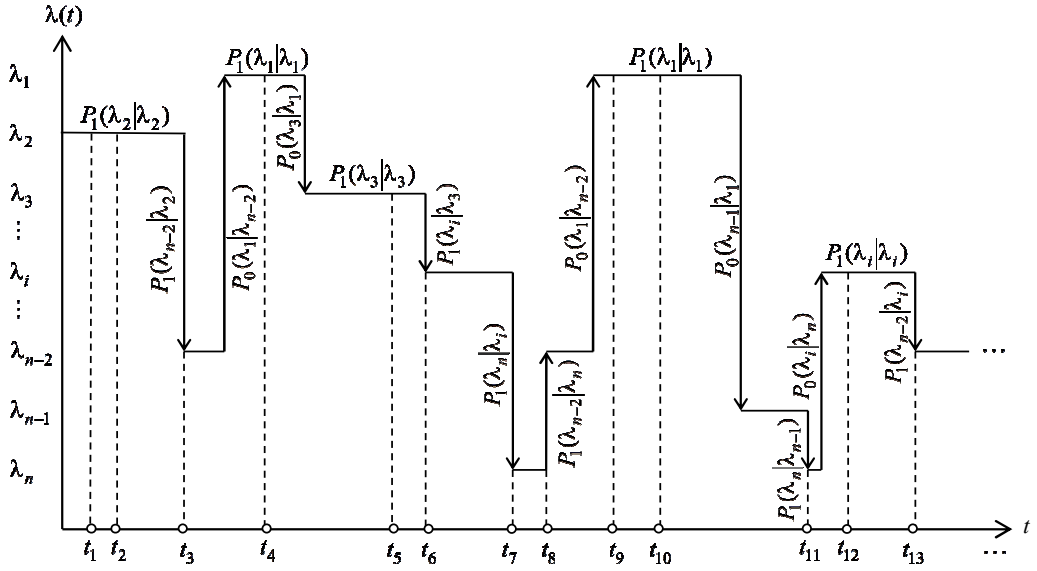


Рис. 1. Реализация обобщённого MAP-потока событий с произвольным числом состояний

Блочная матрица инфинитезимальных характеристик процесса $\lambda(t)$ имеет вид $D = \|D_0 | D_1\|$, где

$$D_0 = \begin{Vmatrix} -\lambda_1 & \lambda_1 P_0(\lambda_2 \mid \lambda_1) & \dots & \lambda_1 P_0(\lambda_n \mid \lambda_1) & \lambda_1 P_0(\lambda_1 \mid \lambda_1) \\ \lambda_2 P_0(\lambda_1 \mid \lambda_2) & -\lambda_2 & \dots & \lambda_2 P_0(\lambda_n \mid \lambda_2) & \lambda_2 P_0(\lambda_2 \mid \lambda_2) \\ \lambda_3 P_0(\lambda_1 \mid \lambda_3) & \lambda_3 P_0(\lambda_2 \mid \lambda_3) & \dots & \lambda_3 P_0(\lambda_n \mid \lambda_3) & \lambda_3 P_0(\lambda_3 \mid \lambda_3) \\ \dots & \dots & \dots & \dots & \dots \\ \lambda_n P_0(\lambda_1 \mid \lambda_n) & \lambda_n P_0(\lambda_2 \mid \lambda_n) & \dots & -\lambda_n & \lambda_n P_0(\lambda_n \mid \lambda_n) \end{Vmatrix}_{n\times(n+1)},$$

$$D_1 = \begin{Vmatrix} \lambda_1 P_1(\lambda_1 \mid \lambda_1) & \lambda_1 P_1(\lambda_2 \mid \lambda_1) & \dots & \lambda_1 P_1(\lambda_{n-1} \mid \lambda_1) & \lambda_1 P_1(\lambda_n \mid \lambda_1) \\ \lambda_2 P_1(\lambda_1 \mid \lambda_2) & \lambda_2 P_1(\lambda_2 \mid \lambda_2) & \dots & \lambda_2 P_1(\lambda_{n-1} \mid \lambda_2) & \lambda_2 P_1(\lambda_n \mid \lambda_2) \\ \lambda_3 P_1(\lambda_1 \mid \lambda_3) & \lambda_3 P_1(\lambda_2 \mid \lambda_3) & \dots & \lambda_3 P_1(\lambda_{n-1} \mid \lambda_3) & \lambda_3 P_1(\lambda_n \mid \lambda_3) \\ \dots & \dots & \dots & \dots & \dots \\ \lambda_n P_1(\lambda_1 \mid \lambda_n) & \lambda_n P_1(\lambda_2 \mid \lambda_n) & \dots & \lambda_n P_1(\lambda_{n-1} \mid \lambda_n) & \lambda_n P_1(\lambda_n \mid \lambda_n) \end{Vmatrix}_{n\times n},$$

$$\sum_{j=1}^{n} d_{ij}^{(1)} + \sum_{j=1}^{n+1} d_{ij}^{(0)} = -\lambda_i + \lambda_i \left[\sum_{j=1}^{n} P_0(\lambda_j \mid \lambda_i) + \sum_{j=1}^{n} P_1(\lambda_j \mid \lambda_i) \right] = -\lambda_i + \lambda_i = 0, \quad i = \overline{1,n}.$$

Элементы $d_{ii}^{(0)}$, $i = \overline{1,n}$ прямоугольной матрицы D_0 – интенсивности выхода процесса $\lambda(t)$ из своих состояний, взятые с противоположным знаком; остальные элементы – интенсивности переходов из состояния i в состояние j без наступления события, $i, j = \overline{1,n}$, $i \neq j$. Элементами матрицы D_1 являются интенсивности переходов процесса $\lambda(t)$ из состояния i в состояние j с наступлением события, $i, j = \overline{1,n}$.

Задача настоящего исследования заключается в получении явного вида апостериорных вероятностей $w(\lambda_i \mid t_1,...,t_m,t) = P(\lambda(t) = \lambda_i \mid t_1,...,t_m,t) = w(\lambda_i \mid t)$ того, что в момент времени t значение процесса $\lambda(t) = \lambda_i$, $i = \overline{1,n}$, при условии, что известна реализа-

ция $t_1,...,t_m$ (наблюдения) потока на интервале (t_0,t), где t_0 – момент начала наблюдения, t – момент окончания наблюдения.

Здесь рассматривается стационарный режим функционирования потока, поэтому переходными процессами пренебрегаем и без ограничения общности полагаем $t_0 = 0$.

Оптимальное оценивание состояния процесса $\lambda(t)$ в смысле минимума полной (безусловной) вероятности ошибки принятия решения [11] осуществляется по критерию максимума апостериорной вероятности на основании сравнения $w(\lambda_i \mid t)$, $i = \overline{1,n}$, в момент времени t: если $w(\lambda_i \mid t) \geq w(\lambda_j \mid t)$, $i, j = \overline{1,n}$, $i \neq j$, то оценка $\hat{\lambda}(t) = \lambda_i$.

2. Вывод рекуррентных соотношений для апостериорных вероятностей состояний

Пусть наблюдения за потоком начинаются в момент времени $t = 0$, и время t изменяется дискретно с шагом Δt: $t^{(k)} = k\Delta t$, $k = 0,1,...$ [12]. Введём двумерный случайный процесс $\left(\lambda^{(k)}, r_k\right)$, где $\lambda^{(k)} = \lambda(k\Delta t)$ – значение процесса $\lambda(t)$ в момент времени $t^{(k)} = k\Delta t$ ($\lambda^{(k)} = \lambda_i$, $i = \overline{1,n}$); $r_k = r_k(\Delta t) = r[k\Delta t] - r[(k-1)\Delta t]$ – число событий потока, наступивших на интервале времени $((k-1)\Delta t, k\Delta t)$ длительности Δt, $r_k = 0,1,...$. Поскольку на интервале $(-\Delta t, 0)$ наблюдение за потоком не производится, то r_0 можем положить произвольным, например, $r_0 = 0$.

В силу конструкции процесса $(\lambda^{(k)}, r_k)$ компонента $\lambda^{(k)}$ (значение процесса $\lambda(t)$ в момент времени $t^{(k)} = k\Delta t$) не влияет на компоненту r_k – число событий на интервале $((k-1)\Delta t, k\Delta t)$, а компонента r_k (в зависимости от типа дважды стохастического потока событий) может как влиять, так и не влиять на $\lambda^{(k)}$.

Лемма 1. Случайный процесс $(\lambda^{(k)}, r_k)$ для обобщённого MAP-потока событий является марковским процессом.

Обозначим $p(\lambda^{(k)}, r_k \mid \lambda^{(k-1)}, r_{k-1})$ – вероятность перехода процесса $(\lambda^{(k)}, r_k)$ за один шаг Δt из состояния $(\lambda^{(k-1)}, r_{k-1})$ в состояние $(\lambda^{(k)}, r_k)$.

Введём $\mathbf{r}_m = (r_0, r_1, ..., r_m)$ – последовательность значений количества наблюдённых событий за время от 0 до $m\Delta t$ на интервалах $((k-1)\Delta t, k\Delta t)$ длительности Δt, $k = \overline{0,m}$; $\boldsymbol{\lambda}^{(m)} = (\lambda^{(0)}, \lambda^{(1)}, ..., \lambda^{(m)})$ – последовательность неизвестных (ненаблюдаемых) значений процесса $\lambda(k\Delta t)$ в момент времени $k\Delta t$, $k = \overline{0,m}$ ($\lambda^{(0)} = \lambda(0) = \lambda_i, i = \overline{1,n}$).

$w(\boldsymbol{\lambda}^{(m)}, \mathbf{r}_m)$ – совместная вероятность значений дискретных случайных величин $\boldsymbol{\Lambda}^{(m)}$, $\mathbf{R}^{(m)}$, где $\boldsymbol{\Lambda}^{(m)} = (\Lambda^{(0)}, \Lambda^{(1)}, ..., \Lambda^{(m)})$ и $\mathbf{R}^{(m)} = (R^{(0)}, R^{(1)}, ..., R^{(m)})$ – векторы случайных величин, $\boldsymbol{\lambda}^{(m)} = (\lambda^{(0)}, \lambda^{(1)}, ..., \lambda^{(m)})$ и $\mathbf{r}_m = (r_0, r_1, ..., r_m)$ – векторы значений случайных величин $\boldsymbol{\Lambda}^{(m)}$, $\mathbf{R}^{(m)}$ соответственно.

Обозначим $w(\boldsymbol{\lambda}^{(m)} \mid \mathbf{r}_m)$ – условная вероятность значений $\boldsymbol{\lambda}^{(m)}$ при условии, что наблюдалась реализация $\mathbf{r}_m$; $w(\lambda^{(m)} \mid \mathbf{r}_m)$ – условная вероятность значения $\lambda^{(m)}$ при условии, что наблюдалась реализация $\mathbf{r}_m$.

Теорема 1. Для дважды стохастических потоков событий с n состояниями справедливо рекуррентное соотношение для апостериорных вероятностей [13,14]

$$w(\lambda^{(m+1)}\left|\mathbf{r}_{m+1}\right.)=\frac{\sum\limits_{\lambda^{(m)}=\lambda_1}^{\lambda_n} w(\lambda^{(m)}\left|\mathbf{r}_m\right.)\,p(\lambda^{(m+1)},r_{m+1}\left|\lambda^{(m)},r_m\right.)}{\sum\limits_{\lambda^{(m)}=\lambda_1}^{\lambda_n}\sum\limits_{\lambda^{(m+1)}=\lambda_1}^{\lambda_n} w(\lambda^{(m)}\left|\mathbf{r}_m\right.)\,p(\lambda^{(m+1)},r_{m+1}\left|\lambda^{(m)},r_m\right.)}. \qquad (1)$$

Следствие теоремы 1. Поскольку $p(\lambda^{(m+1)},r_{m+1}\left|\lambda^{(m)},r_m\right.)=p(\lambda^{(m+1)}\left|\lambda^{(m)},r_m\right.)\times$ $\times\, p(r_{m+1}\left|\lambda^{(m)},r_m,\lambda^{(m+1)}\right.)$, рекуррентное соотношение (1) принимает вид

$$w(\lambda^{(m+1)}\left|\mathbf{r}_{m+1}\right.)=\frac{\sum\limits_{\lambda^{(m)}=\lambda_1}^{\lambda_n} w(\lambda^{(m)}\left|\mathbf{r}_m\right.)\,p(\lambda^{(m+1)}\left|\lambda^{(m)},r_m\right.)\,p(r_{m+1}\left|\lambda^{(m)},r_m,\lambda^{(m+1)}\right.)}{\sum\limits_{\lambda^{(m)}=\lambda_1}^{\lambda_n}\sum\limits_{\lambda^{(m+1)}=\lambda_1}^{\lambda_n} w(\lambda^{(m)}\left|\mathbf{r}_m\right.)\,p(\lambda^{(m+1)}\left|\lambda^{(m)},r_m\right.)\,p(r_{m+1}\left|\lambda^{(m)},r_m,\lambda^{(m+1)}\right.)}.$$

Рассмотрим вероятность $w(\lambda^{(m)}\left|\mathbf{r}_m\right.)=w(\lambda^{(m)}\left|r_0(\Delta t),r_1(\Delta t),...,r_m(\Delta t)\right.)$. Компонента $r_k(\Delta t)=r[k\Delta t]-r[(k-1)\Delta t]$ вектора $\mathbf{r}_m$ зависит от текущего времени $t^{(k)}=k\Delta t$, $k=\overline{0,m}$; при этом $t^{(m)}=t$ – момент окончания наблюдений. Итак, $r_k(\Delta t)=r_k(t^{(k)})$, $k=\overline{0,m}$, $t^{(m)}=m\Delta t=t$. Тогда $\mathbf{r}_m=\left(r_0(t^{(0)}),r_1(t^{(1)}),...,r_m(t^{(m)})\right)=\left(r_0\left(t=0\right),r_1\left(\frac{t}{m}\right),\right.$ $\left.r_2\left(\frac{2t}{m}\right),...,r_m\left(t\right)\right)=\mathbf{r}_m(t)$, то есть вектор наблюдений $\mathbf{r}_m$ является функцией времени t.

В связи с этим $w(\lambda^{(m)}\left|\mathbf{r}_m\right.)=w(\lambda^{(m)}\left|\mathbf{r}_m(t)\right.)=w(\lambda^{(m)}\left|t\right.)$. Аналогично для апостериорной вероятности $w(\lambda^{(m+1)}\left|\mathbf{r}_{m+1}\right.)$ имеет место $w(\lambda^{(m+1)}\left|\mathbf{r}_{m+1}(t+\Delta t)\right.)=w(\lambda^{(m+1)}\left|t+\Delta t\right.)$.

Теорема 2. Для дважды стохастических потоков событий с произвольным числом состояний справедливо рекуррентное соотношение для апостериорных вероятностей $w(\lambda^{(m)}\left|t\right.)$, $w(\lambda^{(m+1)}\left|t+\Delta t\right.)$ [13,14]:

$$w(\lambda^{(m+1)}\left|t+\Delta t\right.)=\frac{\sum\limits_{\lambda^{(m)}=\lambda_1}^{\lambda_n} w(\lambda^{(m)}\left|t\right.)\,p(\lambda^{(m+1)}\left|\lambda^{(m)},r_m\right.)\,p(r_{m+1}\left|\lambda^{(m)},r_m,\lambda^{(m+1)}\right.)}{\sum\limits_{\lambda^{(m)}=\lambda_1}^{\lambda_n}\sum\limits_{\lambda^{(m+1)}=\lambda_1}^{\lambda_n} w(\lambda^{(m)}\left|t\right.)\,p(\lambda^{(m+1)}\left|\lambda^{(m)},r_m\right.)\,p(r_{m+1}\left|\lambda^{(m)},r_m,\lambda^{(m+1)}\right.)}. \qquad (2)$$

Утверждение 2. Переходная вероятность $p(\lambda^{(m+1)}\left|\lambda^{(m)},r_m\right.)\,p(r_{m+1}\left|\lambda^{(m)},r_m,\lambda^{(m+1)}\right.)$ для обобщённого MAP-потока событий с произвольным числом состояний принимает вид $p(\lambda^{(m+1)}\left|\lambda^{(m)},r_m\right.)\,p(r_{m+1}\left|\lambda^{(m)},r_m,\lambda^{(m+1)}\right.)=p(\lambda^{(m+1)}\left|\lambda^{(m)}\right.)\,p(r_{m+1}\left|\lambda^{(m)},\lambda^{(m+1)}\right.)$.

Рассмотрим случай $r_{m+1}=0$ в (2), то есть на интервале $(t,t+\Delta t)$, где $t=m\Delta t$, $t+\Delta t=(m+1)\Delta t$, нет событий обобщённого MAP-потока. Этот случай описывает поведение апостериорной вероятности на интервале времени между моментами t_k и t_{k+1}, $k=1,2,...$, наступления соседних событий потока. Положим в (2) $\lambda^{(m+1)}=\lambda_i$. Тогда, с учётом утверждения 2, рекуррентное соотношение (2) для апостериорных вероятностей $w(\lambda^{(m)}\left|t\right.)$, $w(\lambda^{(m+1)}\left|t+\Delta t\right.)$ состояний обобщённого MAP-потока запишется в виде

$$w(\lambda_i|t+\Delta t)=\frac{\sum\limits_{\lambda^{(m)}=\lambda_1}^{\lambda_n} w(\lambda^{(m)}|t)\,p(\lambda^{(m+1)}=\lambda_i|\lambda^{(m)})\,p(r_{m+1}=0|\lambda^{(m)},\lambda^{(m+1)}=\lambda_i)}{\sum\limits_{\lambda^{(m)}=\lambda_1}^{\lambda_n}\sum\limits_{\lambda^{(m+1)}=\lambda_1}^{\lambda_n} w(\lambda^{(m)}|t)\,p(\lambda^{(m+1)}|\lambda^{(m)})\,p(r_{m+1}=0|\lambda^{(m)},\lambda^{(m+1)})}, \quad (3)$$

$$w(\lambda_i|t+\Delta t)=1-\sum_{\substack{j=1\\ j\neq i}}^{n} w(\lambda_j|t+\Delta t),\ i=\overline{1,n}.$$

3. Явный вид апостериорных вероятностей

Выполняя в соотношении (3) необходимые преобразования, с учётом определения потока, и переходя к пределу при $\Delta t\to 0$, получим систему нелинейных дифференциальных уравнений относительно апостериорных вероятностей

$$w'(\lambda_i|t)=-\lambda_i w(\lambda_i|t)+\sum_{j=1}^{n}\lambda_j P_0(\lambda_i|\lambda_j)w(\lambda_j|t)+w(\lambda_i|t)\sum_{k=1}^{n}\lambda_k\left(1-\sum_{j=1}^{n}P_0(\lambda_j|\lambda_k)\right)w(\lambda_k|t), \quad (4)$$

$$i=\overline{1,n}.$$

Обозначим $u_k=1-\sum\limits_{j=1}^{n}P_0(\lambda_j|\lambda_k)=\sum\limits_{j=1}^{n}P_1(\lambda_j|\lambda_k)$.

С учётом сделанного обозначения система уравнений (4) принимает вид

$$w'(\lambda_i|t)=-\lambda_i w(\lambda_i|t)+\sum_{j=1}^{n}\lambda_j P_0(\lambda_i|\lambda_j)w(\lambda_j|t)+w(\lambda_i|t)\sum_{k=1}^{n}\lambda_k u_k w(\lambda_k|t),\ i=\overline{1,n}. \quad (5)$$

Теорема 3. Апостериорные вероятности состояний обобщённого MAP-потока событий с n состояниями, $n=2,3,\ldots$, на интервалах (t_k,t_{k+1}), $k=1,2,\ldots$, между моментами наступления событий имеют вид

$$w(\lambda_i|t)=\frac{\sum\limits_{s=1}^{n}\sum\limits_{j=1}^{n}\frac{\widetilde{A}_{js}}{\det\widetilde{A}_1}\alpha_i^{(s)}w(\lambda_j|t_k)e^{k_s(t-t_k)}}{\sum\limits_{i=1}^{n}\sum\limits_{s=1}^{n}\sum\limits_{j=1}^{n}\frac{\widetilde{A}_{js}}{\det\widetilde{A}_1}\alpha_i^{(s)}w(\lambda_j|t_k)e^{k_s(t-t_k)}},\ i=\overline{1,n},\ t_k<t<t_{k+1},\ k=1,2,\ldots, \quad (6)$$

где величины $\widetilde{A}_1$, $\widetilde{A}_{js}$, $\widetilde{A}_{js}$, k_s определяются в ходе доказательства.

Доказательство. Обозначим $\psi(t)=\sum\limits_{k=1}^{n}\lambda_k u_k w(\lambda_k|t)$, выполним замену переменных

$$w(\lambda_i|t)=x_i(t)e^{\int_{t_k}^{t}\psi(\tau)d\tau}, \quad (7)$$

где $x_i(t)$ – новые неизвестные функции, $i=\overline{1,n}$, $k=1,2,\ldots$. Тогда (5) примет вид

$$x_i'(t)=-\lambda_i x_i(t)+\sum_{j=1}^{n}\lambda_j P_0(\lambda_i|\lambda_j)x_j(t),\ i=\overline{1,n}.$$

Полученную систему дифференциальных уравнений представим в матричном виде

$$\dot{X}=AX, \quad (8)$$

$$X = \left\| \begin{matrix} x_1(t) \\ x_2(t) \\ \dots \\ x_n(t) \end{matrix} \right\|_{n\times 1}, \quad A = \left\| \begin{matrix} -\lambda_1(1-P_0(\lambda_1|\lambda_1)) & \lambda_2 P_0(\lambda_1|\lambda_2) & \dots & \lambda_n P_0(\lambda_1|\lambda_n) \\ \lambda_1 P_0(\lambda_2|\lambda_1) & -\lambda_2(1-P_0(\lambda_2|\lambda_1)) & \dots & \lambda_n P_0(\lambda_2|\lambda_n) \\ \dots & \dots & \dots & \dots \\ \lambda_1 P_0(\lambda_n|\lambda_1) & \lambda_2 P_0(\lambda_n|\lambda_2) & \dots & -\lambda_n(1-P_0(\lambda_n|\lambda_n)) \end{matrix} \right\|_{n\times n}.$$

Будем искать решение системы (8) в виде $X = \widetilde{A}e^{kt}$, где вектор-столбец $\widetilde{A} = \|\alpha_i\|$, $i = \overline{1,n}$. Тогда, подставляя X в (8), получим систему линейных однородных уравнений относительно α_i

$$(A - kE)\widetilde{A} = 0, \tag{9}$$

где E – единичная матрица.

Для того чтобы уравнению (9) удовлетворял нетривиальный вектор $\widetilde{A}$, необходимо и достаточно, чтобы определитель системы (9) был равен нулю, то есть

$$|A - kE| = 0. \tag{10}$$

Для каждого корня k_s, $s = \overline{1,n}$, характеристического уравнения (10) из (9) определяется вектор $\widetilde{A}^{(s)} \neq 0$. Будем рассматривать только случай, когда все корни k_s характеристического уравнения (10) действительны и различны. Получаем n линейно независимых решений $X_s = \widetilde{A}^{(s)}e^{k_s t}$, где $\widetilde{A}^{(s)} = \|\alpha_i^{(s)}\|$, $i,s = \overline{1,n}$. Постоянные $\alpha_i^{(s)}$, $i,s = \overline{1,n}$, определяются из (9) неоднозначно.

Общее решение системы (8) имеет вид

$$X = \sum_{s=1}^{n} C_s \widetilde{A}^{(s)} e^{k_s t} \quad \text{или} \quad x_i(t) = \sum_{s=1}^{n} C_s \alpha_i^{(s)} e^{k_s t}, \ i = \overline{1,n},$$

где C_s – произвольные постоянные, которые определяются из начальных условий. С учётом замены (7) в момент времени $t = t_k$ находим $x_i(t_k) = w(\lambda_i|t_k)$, $i = \overline{1,n}$, $k = 1,2,\dots$.

Тогда имеем систему n линейных неоднородных уравнений для нахождения произвольных постоянных C_s

$$\sum_{s=1}^{n} C_s \alpha_i^{(s)} e^{k_s t_k} = w(\lambda_i|t_k), \ i = \overline{1,n},$$

с отличным от нуля определителем для n линейно независимых решений соответствующей однородной системы. Решая полученную систему с использованием правила Крамера, находим

$$C_s = e^{-k_s t_k} \sum_{j=1}^{n} \frac{\widetilde{A}_{js}}{\det \widetilde{A}_1} w(\lambda_j|t_k), \ s = \overline{1,n},$$

где $\widetilde{A}_{js}$ – алгебраические дополнения элементов $\widetilde{a}_{js}$ матрицы $\widetilde{A}_1$ ($\widetilde{a}_{js} = \alpha_j^{(s)}$); $\det \widetilde{A}_1$ – определитель матрицы $\widetilde{A}_1$.

Поэтапно возвращаясь к ранее сделанным заменам, получаем (7) в виде

$$w(\lambda_i|t) = e^{\int_{t_k}^{t} \psi(\tau)d\tau} \sum_{s=1}^{n}\sum_{j=1}^{n} \frac{\widetilde{A}_{js}}{\det \widetilde{A}_1} \alpha_i^{(s)} w(\lambda_j|t_k) e^{k_s(t-t_k)}, \ i = \overline{1,n}. \tag{11}$$

С учётом условия нормировки, из (11) находится явный вид функции

$$e^{\int_{t_k}^{t}\psi(\tau)d\tau} = \left\{\sum_{i=1}^{n}\sum_{s=1}^{n}\sum_{j=1}^{n}\frac{\widetilde{A}_{js}}{\det\widetilde{A}_1}\alpha_i^{(s)}w(\lambda_j|t_k)e^{k_s(t-t_k)}\right\}^{-1}. \qquad (12)$$

Подставляя (12) в (11), приходим к (6). Теорема 3 доказана.

Для того чтобы найти значение $w(\lambda_i|t)$ в произвольный момент времени t, необходимо знать $w(\lambda_j|t_k)$, $j=\overline{1,n}$, – значение апостериорной вероятности в момент времени t_k, $k=1,2,...$, наступления события потока.

Апостериорная вероятность $w(\lambda_j|t_k)$ определяется следующей теоремой.

Теорема 4. В момент времени t_k наблюдения события в обобщённом MAP-потоке для апостериорных вероятностей $w(\lambda_i|t_k)$, $i=\overline{1,n}$, имеет место формула пересчёта

$$w(\lambda_i|t_k+0)=\frac{\sum_{j=1}^{n}\lambda_j P_1(\lambda_i|\lambda_j)w(\lambda_j|t_k-0)}{\sum_{j=1}^{n}\sum_{s=1}^{n}\lambda_j P_1(\lambda_s|\lambda_j)w(\lambda_j|t_k-0)}, \qquad (13)$$

$$w(\lambda_i|t_k+0)=1-\sum_{\substack{j=1\\ j\neq i}}^{n}w(\lambda_j|t_k+0),\ i=\overline{1,n}.$$

Доказательство. Рассмотрим случай, когда в (2) $r_{m+1}=1$, то есть на интервале времени $(t,t+\Delta t)$ наступает событие потока, допустим, в момент времени t_k. Положим в (2) $\lambda^{(m+1)}=\lambda_i$. Тогда, с учётом утверждения 2, рекуррентное соотношение (2) для апостериорных вероятностей $w(\lambda^{(m)}|t)$, $w(\lambda^{(m+1)}|t+\Delta t)$ состояний потока запишется в виде

$$w(\lambda_i|t+\Delta t)=\frac{\sum_{\lambda^{(m)}=\lambda_1}^{\lambda_n}w(\lambda^{(m)}|t)p(\lambda^{(m+1)}=\lambda_i|\,\lambda^{(m)})p(r_{m+1}=1|\,\lambda^{(m)},\lambda^{(m+1)}=\lambda_i)}{\sum_{\lambda^{(m)}=\lambda_1}^{\lambda_n}\sum_{\lambda^{(m+1)}=\lambda_1}^{\lambda_n}w(\lambda^{(m)}|t)p(\lambda^{(m+1)}|\,\lambda^{(m)})p(r_{m+1}=1|\,\lambda^{(m)},\lambda^{(m+1)})},\ i=\overline{1,n}.$$

Рассмотрим два смежных интервала (t,t_k) и $(t_k,t+\Delta t)$, длительности которых есть $\Delta t'=t_k-t$ и $\Delta t''=t+\Delta t-t_k$ соответственно. Выполняя в последнем рекуррентном соотношении введённые замены $t=t_k-\Delta t'$, $t+\Delta t=t_k+\Delta t''$, учитывая определение потока и переходя к пределу при $\Delta t\to 0$ ($\Delta t'\to 0$ и $\Delta t''\to 0$ одновременно), приходим к (13). Теорема 4 доказана.

Осталось рассмотреть поведение апостериорных вероятностей $w(\lambda_i|t)$, $i=\overline{1,n}$, на отрезке $[t_0,t_1]$, где t_0 – момент начала наблюдения за потоком, t_1 – момент наступления первого события потока. В момент времени t_0 апостериорную вероятность $w(\lambda_i|t)$ целесообразно положить равной априорной финальной вероятности π_i, $i=\overline{1,n}$.

4. Выражения для априорных финальных вероятностей состояний потока

Обозначим $\pi_i(t|t^0)$ – априорная вероятность того, что в момент времени t значение процесса $\lambda(t)=\lambda_i$, $i=\overline{1,n}$, при условии, что функционирование потока началось в момент времени t^0.

Теорема 5. Априорные финальные вероятности состояний процесса $\lambda(t)$ для обобщённого MAP-потока событий имеют вид

$$\pi_j=\frac{P_{nj}}{\det P},\ j=\overline{1,n}, \tag{14}$$

где P_{nj} – алгебраическое дополнение элемента p_{nj} матрицы $P=\left\|p_{ij}\right\|$, $i,j=\overline{1,n}$,

$$p_{nj}=\begin{cases}-\lambda_i\left(1-P_0(\lambda_i|\lambda_i)-P_1(\lambda_i|\lambda_i)\right), & i=\overline{1,n-1};\\ \lambda_j\left(P_0(\lambda_i|\lambda_j)+P_1(\lambda_i|\lambda_j)\right), & i\neq j, i=\overline{1,n-1}, j=\overline{1,n};\\ 1, & i=n, j=\overline{1,n}.\end{cases}$$

Доказательство. Нетрудно показать, что для априорных вероятностей $\pi_i(t|t^0)$, $i=\overline{1,n}$, справедлива система дифференциальных уравнений

$$\pi_i'(t|t^0)=-\lambda_i\pi_i(t|t^0)+\sum_{j=1}^{n}\lambda_j\left[P_0(\lambda_i|\lambda_j)+P_1(\lambda_i|\lambda_j)\right]\pi_j(t|t^0),\ i=\overline{1,n}, \tag{15}$$

$$\sum_{i=1}^{n}\pi_i(t|t^0)=1.$$

Учитывая тот факт, что рассматривается стационарный режим функционирования потока и $\lambda(t)$ – марковский процесс, имеем [15] $\lim\limits_{t\to\infty}\pi_i(t|t^0)=\pi_i$, $i=\overline{1,n}$. Тогда система (15) принимает вид

$$-\lambda_i\pi_i+\sum_{j=1}^{n}\lambda_j\left[P_0(\lambda_i|\lambda_j)+P_1(\lambda_i|\lambda_j)\right]\pi_j=0,\ i=\overline{1,n}.$$

В силу линейной зависимости уравнений последней системы заменим последнее уравнение при $i=n$ условием нормировки $\sum\limits_{i=1}^{n}\pi_i=1$ и, решая систему с использованием формул Крамера, приходим к (14). Теорема 5 доказана.

Заключение

В данной статье получены выражения для апостериорных вероятностей состояний $w(\lambda_i|t)$, $i=\overline{1,n}$, обобщённого MAP-потока событий с произвольным числом состояний, также получены формулы априорных финальных вероятностей состояний π_i, $i=\overline{1,n}$, необходимых для расчёта апостериорных вероятностей в начальный момент времени t_0 наблюдения за потоком.

ЛИТЕРАТУРА

1. *Cox D.R.* The analysis of non-Markovian stochastic processes // Proceedings of the Cambridge Philosophical Society. – 1955. – 51 (3). – P. 433–441.

2. *Башарин Г.П., Кокотушкин В.А., Наумов В.А.* О методе эквивалентных замен расчёта фрагментов сетей связи. Ч. 1 // Известия АН СССР. Техническая кибернетика. – 1979. – №6. – С. 92–99.

3. *Nuets M.F.* A versatile Markov point process // Journal of Applied Probability. – 1979. – 16. – P. 764–779.

4. *Кеба А.В., Нежельская Л.А.* Статистические эксперименты на имитационной модели обобщённого MAP-потока событий с произвольным числом состояний // Труды Томского государственного университета. Серия физико-математическая. – 2018. – Т. 302. – С. 157–164.

5. *Нежельская Л.А.* Оптимальное оценивание состояний полусинхронного потока событий в условиях его частичной наблюдаемости // Вестник Томского государственного университета. – 2000. – № 269. – С. 95–98.

6. *Леонова М.А., Нежельская Л.А.* Вероятность ошибки при оценивании состояний обобщённого асинхронного потока событий // Вестник Томского государственного университета. Управление, вычислительная техника и информатика. – 2012. – № 2 (19). – С. 88–101.

7. *Nezhelskaya L.* Optimal state estimation in modulated MAP event flows with unextendable dead time // Communications in Computer and Information Science. –2014. – 487. – P. 342–350.

8. *Gortsev A.M., Nezhelskaya L.A.* Estimations of parameters of synchronously alternating Poisson stream of events by the moment method // Telecommunications and Radio Engineering. – 1996. – 50 (1). – P. 56-63.

9. *Горцев А.М., Нежельская Л.А.* Оценивание длительности мёртвого времени и параметров синхронного альтернирующего потока событий // Вестник Томского государственного университета. – 2003. – № S6. – С. 232–239.

10. *Горцев А.М., Нежельская Л.А.* Оценивание параметров синхронного дважды стохастического потока событий методом моментов // Вестник Томского государственного университета. – 2002. – №S1-1. – С. 24–29.

11. *Левин Б.Р.* Теоретические основы статистической радиотехники. – М.: Сов. Радио, 1968. – 504 с.

12. *Хазен Э.М.* Методы оптимальных статистических решений и задачи оптимального управления. – М.: Сов. Радио, 1968. – 256 с.

13. *Горцев А.М., Зуевич В.Л.* Оптимальная оценка состояний асинхронного дважды стохастического потока событий с произвольным числом состояний // Вестник Томского государственного университета. Управление, вычислительная техника и информатика. – 2010. – № 2 (11). – С. 44–65.

14. *Горцев А.М., Нежельская Л.А.* Оптимальная нелинейная фильтрация марковского потока событий с переключениями // Техника средств связи. Серия: Системы связи. – 1989. – Вып. 7. – С. 46–54.

15. *Хинчин А.Я.* Работы по математической теории массового обслуживания. – М.: Физматлит, 1963. – 236 с.

УРАВНЕНИЕ МОМЕНТОВ ДЛЯ ОЦЕНКИ ДЛИТЕЛЬНОСТИ МЁРТВОГО ВРЕМЕНИ В MAP-ПОТОКЕ СОБЫТИЙ

Б.О. Лопсанчап, Л.А. Нежельская

Томский государственный университет

Введение

Развитие информационных технологий порождает новые задачи в теории массового обслуживания (ТМО), решение которых находит своё успешное применение в различных сферах деятельности человека: в телекоммуникационных системах и сетях связи, экономике, финансовом и логистико-транспортном деле, в военной и гражданской авиации, в метеорологической и гидрологической службе.

В этой связи модели систем массового обслуживания (СМО) с входящим простейшим потоком событий перестали адекватно описывать реальные системы, т.к. при моделировании и проектировании компьютерных сетей и систем связи возникла необходимость учитывать тот факт, что интенсивность наступления событий меняется во времени. Это обстоятельство привело к появлению новых математических моделей потоков событий, относящихся к классу дважды стохастических потоков [1,2], в которых моменты времени наступления событий есть случайные величины и интенсивность потока представляет собой случайный процесс.

Условия функционирования реальных процессов и систем таковы, что доступны наблюдению лишь моменты времени наступления событий потока. Поэтому с целью обеспечения адаптации реальной системы к такому входящему потоку существенный интерес представляют задачи, связанные с оценкой состояний [3−5] и параметров [6−8] дважды стохастических потоков по наблюдениям за моментами наступления событий.

Большинство авторов работ по ТМО исследуют СМО в условиях, когда все события входящего потока доступны наблюдению. В реальности же зарегистрированное событие может создать период ненаблюдаемости для регистрирующего прибора, так называемое мёртвое время. Период ненаблюдаемости потока может быть фиксированной величиной [9], а также может быть случайным [10].

В настоящей статье рассматривается MAP-поток событий, функционирующий в условиях продлевающегося мёртвого времени фиксированной длительности так, что общий период ненаблюдаемости потока является случайной величиной.

Подчеркнем, что дважды стохастические потоки событий с двумя состояниями в условиях продлевающегося мёртвого времени рассмотрены в статьях [11−13]: в [11] − полусинхронный поток, в [12] − синхронный поток, в [13] − альтернирующий асинхронный поток событий.

В представленной работе для рекуррентного MAP-потока событий при продлевающемся мёртвом времени выписывается уравнение моментов для решения задача оценивания длительности мёртвого времени на основании последовательности моментов наступления событий в наблюдаемом потоке

1. Постановка задачи

Рассматривается MAP-поток событий с двумя состояниями в условиях продлевающегося мёртвого времени, сопровождающий случайный процесс которого $\lambda(t)$ является кусочно-постоянным ненаблюдаемым процессом с двумя состояниями s_1 и s_2. Далее под i-м состоянием понимается состояние s_i, $i=1,2$.

Длительность пребывания процесса $\lambda(t)$ в i-м состоянии, $i=1,2$, есть случайная величина с функцией распределения $F_i(t)=1-e^{-\lambda_i t}$, $i=1,2$. Если процесс $\lambda(t)$ в момент времени t находится в i-м состоянии ($\lambda(t)=\lambda_i$), то на полуинтервале $[t;t+\Delta t)$, где Δt – достаточно малая величина, в момент окончания i-го состояния происходит розыгрыш пар, при котором возможны 3 ситуации:

1) не наступает событие потока и процесс $\lambda(t)$ переходит из i-го состояния в j-е состояние, $i=1,2$, $j=1,2$, $i\neq j$, с вероятностью $P_0(\lambda_j|\lambda_i)$;
2) наступает событие потока и процесс $\lambda(t)$ остается в i-м состоянии, $i=1,2$, с вероятностью $P_1(\lambda_i|\lambda_i)$;
3) наступает событие потока и процесс $\lambda(t)$ переходит из i-го состояния в j-е состояние, $i=1,2$, $j=1,2$, $i\neq j$, с вероятностью $P_1(\lambda_j|\lambda_i)$.

При этом

$$P_0(\lambda_j|\lambda_i)+P_1(\lambda_i|\lambda_i)+P_1(\lambda_j|\lambda_i)=1,\ i=1,2,\ j=1,2,\ i\neq j.$$

Если $\lambda(t)=\lambda_i$, $i=1,2$, то будем говорить, что имеет место i-е состояние процесса $\lambda(t)$. Полагается, что $\lambda_1>\lambda_2>0$.

Матрицы инфинитезимальных характеристик процесса $\lambda(t)$ имеют вид

$$\boldsymbol{D}_0=\begin{Vmatrix} -\lambda_1 & \lambda_1 P_0(\lambda_2|\lambda_1) \\ \lambda_2 P_0(\lambda_1|\lambda_2) & -\lambda_2 \end{Vmatrix},\quad \boldsymbol{D}_1=\begin{Vmatrix} \lambda_1 P_1(\lambda_1|\lambda_1) & \lambda_1 P_1(\lambda_2|\lambda_1) \\ \lambda_2 P_1(\lambda_1|\lambda_2) & \lambda_2 P_1(\lambda_2|\lambda_2) \end{Vmatrix}.$$

Элементы на главной диагонали матрицы $\boldsymbol{D}_0$ представляют собой интенсивности выхода процесса $\lambda(t)$ из своих состояний, взятые с противоположным знаком; остальные элементы – интенсивности переходов из состояния i в состояние j без наступления события потока, $i=1,2$, $j=1,2$, $i\neq j$. Элементами матрицы $\boldsymbol{D}_1$ являются интенсивности переходов процесса $\lambda(t)$ из состояния i в состояние j с наступлением события потока, $i=1,2$, $j=1,2$.

После каждого зарегистрированного события MAP-потока наступает период ненаблюдаемости событий фиксированной длительности T (мёртвое время), который создается этим событием, так что другие события MAP-потока, наступившие в течение времени T, недоступны наблюдению. Вместе с тем, хотя события и не наблюдаются в те-

чение мёртвого времени, каждое из них вызывает продление периода ненаблюдаемости на ту же величину T, так что наблюдаться будет лишь то событие, которое наступило после окончания последнего периода мёртвого времени. По окончании общего периода ненаблюдаемости первое наступившее событие снова создает период мёртвого времени длительности T и т.д. Таким образом, общий период ненаблюдаемости есть случайная величина.

В силу того, что $\lambda(t)$ – принципиально ненаблюдаемый процесс, наблюдаются только моменты времени наступления событий t_1, t_2, …, $\lambda(t)$ является скрытым марковским процессом или ненаблюдаемым сопровождающим марковским процессом. В моменты времени t_1, t_2, …, t_k, … последовательность $\{\lambda(t_k)\}$ представляет собой вложенную цепь Маркова.

На рис. 1 приведён пример одной из возможных реализаций процесса $\lambda(t)$ и наблюдаемого потока, где t_1, t_2, … − моменты наступления событий в наблюдаемом потоке; чёрными кружками обозначены потерянные события исходного потока; штриховкой ─ длительности мёртвого времени; 1 и 2 ─ состояния случайного процесса.

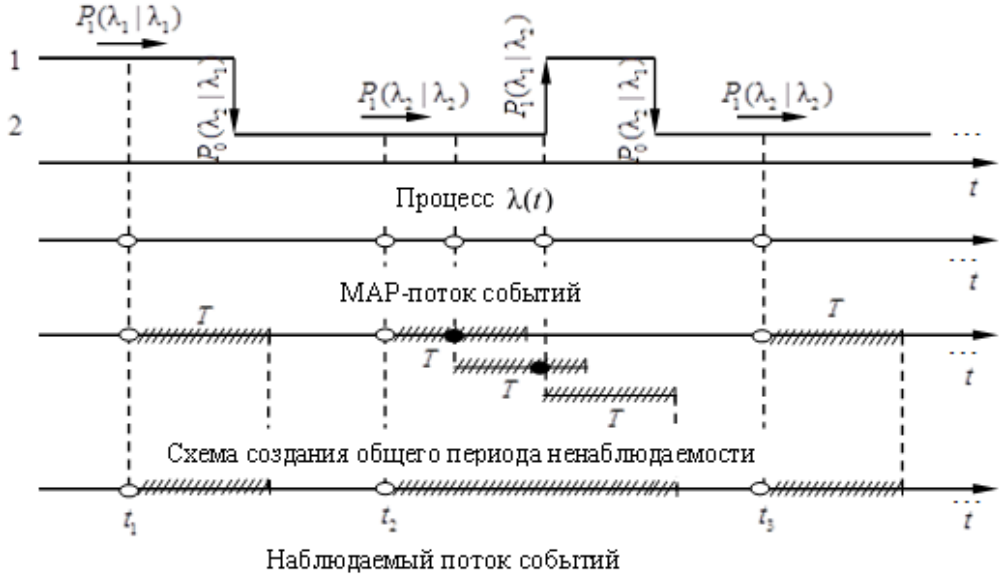


Рис. 1. Формирование наблюдаемого потока событий

Требуется получить преобразование Лапласа плотности вероятности значений длительности интервала между событиями в наблюдаемом потоке и на основании полученной плотности выписать уравнение моментов для определения оценки $\hat{T}$ длительности мёртвого времени T.

2. Плотности вероятностей $p(\tau)$, $p(\tau_1, \tau_2)$ и условия рекуррентности MAP-потока событий

Обозначим $\tau_k = t_{k+1} - t_k$, $k = 1, 2, \ldots$ – значение длительности интервала между соседними событиями потока, $p(\tau)$ – плотность вероятности значений длительности интервала между соседними событиями в потоке. Рассматривается стационарный режим функционирования потока, поэтому $p(\tau_k) = p(\tau)$ для всех $k = 1, 2, \ldots$, $\tau \geq 0$. В этой связи без ограничения общности момент наступления события t_k можно положить равным нулю, т.е. момент наступления события есть $\tau = 0$.

Рассмотрим два соседних временных интервала (t_1,t_2), (t_2,t_3) со значениями длительностей $\tau_1=t_2-t_1$ и $\tau_2=t_3-t_2$ соответственно. Совместная плотность вероятности при этом есть $p(\tau_1,\tau_2)$, $\tau_1\ge 0$, $\tau_2\ge 0$.

Введем в рассмотрение $\pi_i(0)$ – условную стационарную вероятность того, что процесс $\lambda(\tau)$ в момент времени $\tau=0$ находится в *i*-м состоянии при условии, что в момент времени $\tau=0$ наступило событие потока, $i=1,2$, $\pi_1(0)+\pi_2(0)=1$.

Следует отметить, что MAP-поток событий является частным случаем модулированного MAP-потока. В этой связи, полагая в [14] параметры $\alpha_1=\alpha_2=0$, приходим к MAP-потоку событий, для которого справедливы следующие теоремы.

Теорема 1. Плотность вероятности значений длительности интервала между соседними событиями в MAP-потоке имеет вид

$$p(\tau)=\gamma z_1 e^{-z_1\tau}+(1-\gamma)z_2e^{-z_2\tau},\ \tau\ge 0,$$

$$\gamma=\frac{1}{z_2-z_1}\left\{z_2-\lambda_1\pi_1(0)\left[1-P_0(\lambda_2|\lambda_1)\right]-\lambda_2\pi_2(0)\left[1-P_0(\lambda_1|\lambda_2)\right]\right\},$$

$$\pi_1(0)=\frac{P_1(\lambda_1|\lambda_2)+P_1(\lambda_1|\lambda_1)P_0(\lambda_1|\lambda_2)}{P_1(\lambda_1|\lambda_2)+P_1(\lambda_2|\lambda_1)+P_1(\lambda_1|\lambda_1)P_0(\lambda_1|\lambda_2)+P_1(\lambda_2|\lambda_2)P_0(\lambda_2|\lambda_1)},$$

$$\pi_2(0)=\frac{P_1(\lambda_2|\lambda_1)+P_1(\lambda_2|\lambda_2)P_0(\lambda_2|\lambda_1)}{P_1(\lambda_1|\lambda_2)+P_1(\lambda_2|\lambda_1)+P_1(\lambda_1|\lambda_1)P_0(\lambda_1|\lambda_2)+P_1(\lambda_2|\lambda_2)P_0(\lambda_2|\lambda_1)},$$

$$z_{1,2}=\frac{1}{2}\left[(\lambda_1+\lambda_2)\mp\sqrt{(\lambda_1-\lambda_2)^2+4\lambda_1\lambda_2P_0(\lambda_1|\lambda_2)P_0(\lambda_2|\lambda_1)}\right],\ 0<z_1<z_2. \qquad (1)$$

Теорема 2. MAP-поток событий в общем случае является коррелированным потоком и совместная плотность вероятности $p(\tau_1,\tau_2)$ имеет вид

$$p(\tau_1,\tau_2)=p(\tau_1)p(\tau_2)+\gamma(1-\gamma)\frac{\lambda_1\lambda_2}{z_1z_2}\left[P_1(\lambda_1|\lambda_1)P_1(\lambda_2|\lambda_2)-P_1(\lambda_1|\lambda_2)P_1(\lambda_2|\lambda_1)\right]\times$$
$$\times\left(z_1e^{-z_1\tau_1}-z_2e^{-z_2\tau_1}\right)\times\left(z_1e^{-z_1\tau_2}-z_2e^{-z_2\tau_2}\right),\ \tau_1\ge 0,\ \tau_2\ge 0, \qquad (2)$$

где γ, z_1, z_2, $p(\tau_k)$ определены в (1) для $\tau=\tau_k$, $k=1,\ 2$.

Как следует из (2), совместная плотность вероятности $p(\tau_1,\tau_2)$ факторизуется, т.е. $p(\tau_1,\tau_2)=p(\tau_1)p(\tau_2)$, если:

1. $P_1(\lambda_1|\lambda_1)P_1(\lambda_2|\lambda_2)-P_1(\lambda_1|\lambda_2)P_1(\lambda_2|\lambda_1)=0$ (первое условие рекуррентности); в результате необходимых преобразований из (1) находим

$$\pi_1(0)=\frac{P_1(\lambda_1|\lambda_1)}{P_1(\lambda_2|\lambda_1)+P_1(\lambda_1|\lambda_1)},\ \pi_2(0)=\frac{P_1(\lambda_2|\lambda_1)}{P_1(\lambda_2|\lambda_1)+P_1(\lambda_1|\lambda_1)},$$

$$p(\tau)=\gamma z_1e^{-z_1\tau}+(1-\gamma)z_2e^{-z_2\tau},\ \tau\ge 0,\ \gamma=\frac{1}{z_2-z_1}\left\{z_2-\lambda_1P_1(\lambda_1|\lambda_1)-\lambda_2P_1(\lambda_2|\lambda_2)\right\}. \qquad (3)$$

Для получения следующих условий рекуррентности, с учетом явного вида γ, определенного в (1), преобразуем выражение $\gamma(1-\gamma)$ в (2) к виду

$$\gamma(1-\gamma)=\frac{z_1z_2}{\lambda_1\lambda_2(z_1-z_2)^2(a_1+a_2)}\left(\lambda_1\left[1-P_0(\lambda_2|\lambda_1)\right]-\lambda_2\left[1-P_0(\lambda_1|\lambda_2)\right]\right)\times$$
$$\times\left(\lambda_1\pi_1(0)\left[1-P_1(\lambda_1|\lambda_1)\right]-\lambda_2\pi_2(0)\left[1-P_1(\lambda_2|\lambda_2)\right]\right),$$

$$a_1 = P_1\left(\lambda_1|\lambda_2\right) + P_1\left(\lambda_1|\lambda_1\right)P_0\left(\lambda_1|\lambda_2\right),\ a_2 = P_1\left(\lambda_2|\lambda_1\right) + P_1\left(\lambda_2|\lambda_2\right)P_0\left(\lambda_2|\lambda_1\right). \quad (4)$$

2. Если в (4) справедливо $\lambda_1\left[1-P_0\left(\lambda_2|\lambda_1\right)\right]-\lambda_2\left[1-P_0\left(\lambda_1|\lambda_2\right)\right]=0$ (второе условие рекуррентности), то $\gamma\left(1-\gamma\right)=0$ и из (1) находим $\gamma=1$, $p\left(\tau\right)=z_1e^{-z_1\tau}$, $\tau\geq 0$.

3. Если в (4) выполняется условие $\lambda_1\pi_1\left(0\right)\left[1-P_1\left(\lambda_1|\lambda_1\right)\right]-\lambda_2\pi_2\left(0\right)\left[1-P_1\left(\lambda_2|\lambda_2\right)\right]=0$ (третье условие рекуррентности), то, как и в предыдущем случае, $\gamma\left(1-\gamma\right)=0$, из (1) получаем $\gamma=1$ и $p\left(\tau\right)=z_1e^{-z_1\tau}$, $\tau\geq 0$.

3. Преобразование Лапласа $g_\xi\left(s\right)$ плотности вероятности общего периода ненаблюдаемости рекуррентного потока событий

Пусть ξ − значение длительности общего периода ненаблюдаемости в рекуррентном дважды стохастическом MAP-потоке событий. Последовательность t_1, t_2, ... моментов наступления событий в наблюдаемом потоке образует вложенную цепь Маркова и рекуррентность наблюдаемого потока сохранится.

Для нахождения преобразования Лапласа $g_\xi\left(s\right)$ плотности вероятности общего периода ненаблюдаемости для рекуррентного MAP-потока событий сошлёмся на теорему о преобразовании Лапласа плотности вероятности общего периода ненаблюдаемости для произвольного рекуррентного дважды стохастического потока событий [11].

Теорема 3. Преобразование Лапласа плотности вероятности значений длительности общего периода ненаблюдаемости в рекуррентном дважды стохастическом потоке событий, функционирующем в условиях продлевающегося мёртвого времени, имеет вид

$$g_\xi\left(s\right)=\frac{\varphi_0\left(T\right)}{e^{sT}}\left[1-\int_0^T e^{-sx}\tilde{p}\left(x\right)dx\right]^{-1}, \quad (5)$$

где $\varphi_0\left(T\right)=1-\int_0^T\tilde{p}\left(x\right)dx$ − функция Пальма; $\tilde{p}\left(x\right)$ − плотность вероятности длительности интервала между соседними событиями в рекуррентном потоке.

Будем рассматривать первый случай рекуррентности MAP-потока событий с плотностью $p\left(\tau\right)$, представленной формулой (3). Отметим, что второй и третий случаи задания параметров потока событий, когда MAP-поток становится рекуррентным, в данной статье не рассматривается, т.к. в отмеченных случаях MAP-поток событий вырождается в простейший с плотностью $p\left(\tau\right)=z_1e^{-z_1\tau}$, $\tau\geq 0$.

Лемма. Преобразование Лапласа плотности вероятности значений длительности общего периода ненаблюдаемости в рекуррентном MAP-потоке событий, функционирующем в условиях продлевающегося мёртвого времени, имеет вид

$$g_\xi\left(s\right)=\frac{\varphi_0\left(T\right)}{e^{sT}}\left(1-\frac{\gamma z_1}{z_1+s}\left[1-e^{-\left(z_1+s\right)T}\right]-\frac{\left(1-\gamma\right)z_2}{z_2+s}\left[1-e^{-\left(z_2+s\right)T}\right]\right)^{-1},$$

$$\varphi_0\left(T\right)=\int_T^\infty\tilde{p}\left(x\right)dx=\gamma e^{-z_1T}+\left(1-\gamma\right)e^{-z_2T}. \quad (6)$$

Доказательство. Производя в (3) переобозначение $p\left(\tau\right)$ на $\tilde{p}\left(x\right)$ и подставляя (3) в (5), приходим к (6). *Лемма доказана.*

4. Преобразование Лапласа $g_\tau(s)$ плотности вероятности длительности интервала между событиями в наблюдаемом потоке

Рассмотрим временной интервал (t_k, t_{k+1}), длительность которого есть $\tau_k = t_{k+1} - t_k$, или длительность рассматриваемого интервала времени равна $\tau = \xi + \eta$, где η – длительность интервала между моментом окончания общего периода ненаблюдаемости и моментом наступления события t_{k+1}. Следует отметить, что в рекуррентном потоке длительности интервалов между событиями τ_k – независимые случайные величины, поэтому индекс k при τ опущен, и рассматривается произвольный интервал времени между событиями в наблюдаемом потоке. Величины ξ и η являются зависимыми. Тогда плотность вероятности значений длительности интервала между событиями в наблюдаемом потоке $p(\tau)$ примет вид

$$p(\tau) = \int_0^\tau p(\xi)\, p(\eta|\ \xi) d\xi = \int_0^\tau p(\xi)\, p(\tau - \xi|\ \xi)\ \ d\xi\,. \tag{7}$$

Найдем преобразование Лапласа $g_\tau(s)$ плотности вероятности $p(\tau)$.

Теорема 4. Преобразование Лапласа плотности вероятности значений длительности интервала между соседними событиями в рекуррентном MAP-потоке событий с продлевающимся мертвым временем имеет вид

$$g_\tau(s) = \frac{1}{(z_1+s)(z_2+s)}\left\{z_1 z_2 + s\left(\lambda_1\pi_1\left[1 - P_0(\lambda_2|\lambda_1)\right] + \lambda_2\pi_2\left[1 - P_0(\lambda_1|\lambda_2)\right]\right)\right\}\ \ g_\xi(s) +$$

$$+\frac{s}{(z_1+s)(z_2+s)}\left(\lambda_1\left[1 - P_0(\lambda_2|\lambda_1)\right] - \lambda_2\left[1 - P_0(\lambda_1|\lambda_2)\right]\right)\ \ [\pi_2 - p_{12}]\ \ g_\xi(a+s),$$

$$a = \lambda_1\left[1 - P_1(\lambda_1|\lambda_1)\right] + \lambda_2\left[1 - P_1(\lambda_2|\lambda_2)\right],\ \ \pi_1 = \lambda_2\left[1 - P_1(\lambda_2|\lambda_2)\right]\Big/a\,,$$

$$\pi_2 = \lambda_1\left[1 - P_1(\lambda_1|\lambda_1)\right]\Big/a\,,$$

$$p_{12} = \left(P_1(\lambda_2|\lambda_1) + P_0(\lambda_2|\lambda_1)P_1(\lambda_2|\lambda_2)\right)\Big/\left(1 - P_0(\lambda_1|\lambda_2)P_0(\lambda_2|\lambda_1)\right), \tag{8}$$

где z_1, z_2 определены формулой (1); $g_\xi(s)$ – формулой (6); $g_\xi(a+s)$ – формулой (6), в которой вместо s нужно подставить $a+s$.

Доказательство. Пусть момент наступления события в наблюдаемом потоке есть $\tau = 0$. Рассмотрим временной интервал $(0, \tau) = (0, \xi + \eta)$. В формуле (7) условная плотность вероятности $p(\tau - \xi|\ \xi)$, $\tau \ge \xi$, по своему смысловому содержанию совпадает с плотностью $p_T(\tau) = p(\tau|T)$, $\tau \ge T$, для MAP-потока событий, функционирующего в условиях непродлевающегося мёртвого времени фиксированной длительности T, если в соответствующей формуле для $p_T(\tau)$ (в случае модулированного MAP-потока [15]) положить параметры $\alpha_1 = \alpha_2 = 0$, заменить T на ξ и записать указанную формулу с учётом первого условия рекуррентности MAP-потока событий. В силу сказанного

$$p(\tau - \xi|\ \xi) = \begin{cases} 0, & 0 \le \tau < \xi\ , \\ \Gamma(\xi) z_1 e^{-z_1(\tau-\xi)} + \left[1 - \Gamma(\xi)\right] z_2 e^{-z_2(\tau-\xi)}, & \tau \ge \xi\ , \end{cases}$$

$$\Gamma(\xi) = \frac{1}{z_2 - z_1}\left\{z_2 - \lambda_1\pi_1(\xi)\left[1 - P_0(\lambda_2|\lambda_1)\right] - \lambda_2\pi_2(\xi)\left[1 - P_0(\lambda_1|\lambda_2)\right]\right\},$$

$$\pi_1(\xi) = \pi_1 + [\pi_2 - p_{12}]\ \ e^{-a\xi},\ \ \pi_2(\xi) = \pi_2 - [\pi_2 - p_{12}]\ \ e^{-a\xi}, \tag{9}$$

где z_1, z_2 определены формулой (1); π_1, π_2, p_{12}, a – формулой (8).

Преобразование Лапласа плотности $p(\tau)$, с учётом (7), примет вид

$$g_\tau(s)=\int_0^\infty e^{-s\tau}p(\tau)d\tau=\int_0^\infty e^{-s\xi}p(\xi)d\xi\int_0^\infty e^{-s\eta}p(\eta|\ \xi)d\eta, \tag{10}$$

где $\eta=\tau-\xi$. Подставляя в (10) выражение (9), в котором $\tau-\xi=\eta$, и проделывая при этом необходимые преобразования, находим

$$g_\tau(s)=\frac{z_2}{z_2+s}\int_0^\infty e^{-s\xi}p(\xi)d\xi+\left(\frac{z_1}{z_1+s}-\frac{z_2}{z_2+s}\right)\int_0^\infty e^{-s\xi}p(\xi)\Gamma(\xi)d\xi. \tag{11}$$

Подставляя $\Gamma(\xi)$ из (9) в (11), приходим к (8). Теорема доказана.

5. Уравнение моментов для определения оценки $\hat{T}$

Будем решать задачу оценивания длительности мёртвого времени T методом моментов. Введём статистики $C_l=\frac{1}{n}\sum_{k=1}^{n}\tau_k^l$, где $\tau_k=t_{k+1}-t_k$ – значение длительности интервала между моментами t_k и t_{k+1} наступления событий в рекуррентном MAP-потоке, функционирующем в условиях продлевающегося мёртвого времени фиксированной длительности.

Преобразование Лапласа (8) позволяет получить начальные моменты

$$M(\tau^k)=(-1)^k\,g_\tau^{(k)}(s)\big|_{s=0},\ k=1,\ 2,\ \ldots\ . \tag{12}$$

Предположим, что параметры λ_i, $P_0(\lambda_j|\lambda_i)$, $i\neq j$, $P_1(\lambda_j|\lambda_i)$, $i,\ j=1,\ 2$, являются известными. Тогда имеем уравнение моментов, позволяющее оценить длительность мёртвого времени T

$$M(\tau)=C_1, \tag{13}$$

где статистика C_1 при большом n стремится к начальному моменту $M(\tau)$ [16], определённому выражением (12) при $k=1$. Вычисляя производную от (8) по s в точке $s=0$ и проделывая необходимые преобразования в (13), получаем уравнение моментов для оценивания длительности мёртвого времени T

$$\frac{(z_1+z_2)-\left(\lambda_1\pi_1\left[1-P_0(\lambda_2|\lambda_1)\right]+\lambda_2\pi_2\left[1-P_0(\lambda_1|\lambda_2)\right]\right)}{z_1\ \ z_2}+\frac{z_2\gamma\left(1-e^{-z_1T}\right)+z_1(1-\gamma)\left(1-e^{-z_2T}\right)}{z_1\ \ z_2\left(\gamma\ \ e^{-z_1T}+(1-\gamma)e^{-z_2T}\right)}-$$
$$-\frac{\left(\lambda_1\left[1-P_0(\lambda_2|\lambda_1)\right]-\lambda_2\left[1-P_0(\lambda_1|\lambda_2)\right]\right)\ (\pi_2-p_{12})}{z_1\ \ z_2}\times$$
$$\times\frac{\gamma\ \ e^{-(z_1+a)T}+(1-\gamma)e^{-(z_2+a)T}}{1-\frac{z_1\gamma}{z_1+a}\left(1-e^{-(z_1+a)T}\right)-\frac{z_2(1-\gamma)}{z_2+a}\left(1-e^{-(z_2+a)T}\right)}=C_1, \tag{14}$$

где z_1, z_2 определены формулой (1); π_1, π_2, p_{12}, a – формулой (8); $0<T\le\tau_{\min}$, $\tau_{\min}=\min\ \tau_k$, $k=\overline{1,n}$.

Решение уравнения (14) возможно только численно. Обозначив левую часть (14) через $M(T)$, можно показать, что функция $M(T)$ является возрастающей функцией переменной T, $T>0$. Тогда (14) либо имеет единственное решение на полуинтервале $(0,\ \tau_{\min}]$, либо решения не имеет на указанном полуинтервале. Единственный корень уравнения (14) на полуинтервале $(0,\ \tau_{\min}]$ является состоятельной оценкой $\hat{T}$ дли-

тельности мёртвого времени T. Если на $\left(0,\ \tau_{\min}\right]$ уравнение (14) решения не имеет, то возможны два варианта: 1) если $M\left(0\right) \geq C_1$, то в качестве оценки $\hat{T}$ естественно выбрать $\hat{T} = 0$; 2) если $M\left(\tau_{\min}\right) < C_1$, то $\hat{T} = \tau_{\min}$.

Заключение

В настоящей работе рассмотрен MAP-поток событий, функционирующий в условиях продлевающегося мёртвого времени фиксированной длительности T. Выписана плотность вероятности значений длительности интервала между событиями потока и совместная плотность вероятности значений длительностей двух смежных интервалов. На основании формулы для совместной плотности сформулированы условия рекуррентности потока. Для рекуррентного MAP-потока событий получено преобразование Лапласа плотности вероятности общего периода ненаблюдаемости и преобразование Лапласа плотности вероятности значений длительности интервала между соседними событиями в наблюдаемом потоке. С использованием преобразования Лапласа выписано уравнение моментов для определения оценки $\hat{T}$ длительности мёртвого времени T.

ЛИТЕРАТУРА

1. *Башарин Г.П., Кокотушкин В.А., Наумов В.А.* О методе эквивалентных замен расчета фрагментов сетей связи. Ч.1 // Изв. АН СССР. Техническая кибернетика. – 1979. – № 6. – С. 92–99.

2. *Neuts M.F.* A versatile Markov point process // Journal of Applied Probability. – 1979. – V. 16. – P. 764–779.

3. *Леонова М.А., Нежельская Л.А.* Вероятность ошибки при оценивании состояний обобщенного асинхронного потока событий // Вестник Томского государственного университета. Управление, вычислительная техника и информатика. – 2012. – № 2 (19). – С. 88–101.

4. *Nezhelskaya L.* Optimal state estimation in modulated MAP event flows with unextendable dead time // Communications in Computer and Information Science. – 2014. – V. 487. – P. 342–350.

5. *Нежельская Л.А., Сидорова Е.Ф.* Алгоритм оптимального оценивания состояний обобщённого синхронного потока событий второго порядка // Труды Томского государственного университета. Серия физико-математическая. Томск: Издательский Дом ТГУ. – 2017. – Т. 301. – С. 105−113.

6. *Горцев А.М., Нежельская Л.А.* Оценивание длительности мертвого времени и параметров синхронного альтерниирующего потока событий // Вестник Томского государственного университета. – 2003. – № S6. – С. 232–239.

7. *Горцев А.М., Нежельская Л.А.* Оценивание параметров синхронного дважды стохастического потока событий методом моментов // Вестник Томского государственного университета. – 2002. – № S1–1. – С. 24–29.

8. *Василевская Т.П., Горцев А.М., Нежельская Л.А.* Оценивание длительности мертвого времени и параметров синхронного альтерниирующего потока с проявлением либо непроявлением событий // Вестник Томского государственного университета. – 2004. – № S9–2. – С. 129–138.

9. *Nezhel'skaya L.* Probability density function for modulated MAP event flows with unextendable dead time // Communications in Computer and Information Science. – 2015. – V. 564. – P. 141–151.

10. *Горцев А.М., Завгородняя М.Е.* Оценивание параметра непродлевающегося мёртвого времени случайной длительности в пуассоновском потоке событий // Вестник Томского государственного университета. Управление, вычислительная техника и информатика. – 2017. – № 40. – С. 32–40.

11. *Горцев А.М., Нежельская Л.А.* Полусинхронный дважды стохастический поток событий при продлевающемся мертвом времени // Вычислительные технологии. – 2008. – Т. 13. – № 1. – С. 31–41.

12. *Бушланов И.В., Горцев А.М., Нежельская Л.А.* Оценка параметров синхронного дважды стохастического потока событий // Автоматика и телемеханика. – 2008. – № 9. – С. 76–93.

13. *Горцев А.М., Паршина М.Е.* Оценивание параметров альтернирующего потока событий в условиях «мёртвого времени» // Известия высших учебных заведений. Физика. – 1999. – Т. 42. – № 4. – С. 8–13.

14. *Нежельская Л.А.* Совместная плотность вероятностей длительности интервалов модулированного MAP-потока событий и условия рекуррентности потока // Вестник Томского государственного университета. Управление, вычислительная техника и информатика. – 2015. – № 1 (30). – С. 57–67.

15. *Березин Д.В., Нежельская Л.А.* Сравнение МП- и ММ- оценок длительности мертвого времени в модулированном MAP-потоке событий // Вестник Томского государственного университета. Управление, вычислительная техника и информатика. – 2016. – № 3 (36). – С. 26–34.

16. *Малинковский Ю.В.* Теория вероятностей и математическая статистика (часть 2. Математическая статистика). – Гомель: УО «ГТУ им. Ф. Скорины», 2004. – 146 с.

ИМИТАЦИОННОЕ МОДЕЛИРОВАНИЕ ПОЛУСИНХРОННОГО ДВАЖДЫ СТОХАСТИЧЕСКОГО ПОТОКА СОБЫТИЙ С НЕПРОДЛЕВАЮЩИМСЯ СЛУЧАЙНЫМ МЁРТВЫМ ВРЕМЕНЕМ

А.М. Горцев, А.В. Веткина
Томский государственный университет
anyavetkina@gmail.com

Введение

Широкое применение в исследовании реальных физических, технических, экономических и других объектов и систем получили математические модели систем и сетей массового обслуживания (СМО, СеМО). Основными элементами СМО и СеМО при этом являются случайные входящие потоки событий. В подавляющем большинстве работ по исследованию СМО и СеМО до 80-х годов прошлого века в качестве входящих потоков событий рассматривались пуассоновские потоки событий. Однако в связи с бурным развитием вычислительной техники, спутниковых, компьютерных, беспроводных и мобильных сетей связи модель простейшего потока перестала быть адекватной реальным информационным потокам событий. Сама же теория построения математических моделей функционирования информационно-телекоммуникационных систем, существовавшая до середины 80-х годов прошлого века, во многом становится непригодной для анализа информационных процессов, протекающих в современных телекоммуникационных системах. Поэтому в это же время была предпринята успешная попытка создания адекватных математических моделей информационных потоков в телекоммуникационных системах – так называемых дважды стохастических потоков. Дважды стохастические потоки можно разделить на два класса: первый класс составляют потоки, сопровождающий процесс (интенсивность) которых есть непрерывный случайный процесс [1,2]; второй – потоки, сопровождающий процесс (интенсивность) которых есть кусочно-постоянный случайный процесс с конечным (произвольным) числом состояний. Первые результаты исследований потоков второго класса были опубликованы практически в одно и то же время, в 1979 г., в работах [3,4]. В [3] указанные потоки получили название MC (Markov chain)-потоки; в [4] – MVP (Markov versatile processes)-потоки. В статьях [5,6] описанные выше потоки называют также MAP (Markovian Arrival Process)-потоками событий. Подчеркнем, что MC-потоки событий (MAP-потоки) являются наиболее характерной и подходящей моделью потоков в реальных телекоммуникационных сетях, в частности, в широкополосных сетях беспроводной связи вдоль протяженных транспортных магистралей [7–9]. Зарубежными и отечественными авторами работ при описании подобных входящих потоков событий в СМО и СеМО используются термины: дважды стохастические потоки событий, MAP-потоки, MC-потоки и др. В свою очередь, в зависимости от того, каким образом происходит переход интенсивности из состояния в состояние, MC-потоки можно разделить на три типа: 1) синхронные потоки (потоки, у которых состояние интенсивности меняется в случайные моменты времени, являющиеся моментами наступления событий) [10–12]; 2) асинхронные потоки (потоки с интенсивностью, для которой переход из состояния в состояние происходит в случайные моменты времени и не зависит от моментов наступления событий) [13–15]; 3) полусинхронные потоки (потоки, у которых одна часть состояний интенсивности меняется в моменты наступления событий потока, другая часть состояний интенсивности меняется в произвольные моменты времени, не связанные с моментами наступления событий потока) [16–18].

В большинстве публикаций авторы рассматривают математические модели потоков, когда события потоков доступны наблюдению. Однако на практике возникают ситуации, когда наступившее событие влечёт за собой ненаблюдаемость последующих событий. Причиной ненаблюдаемости, как правило, выступает мёртвое время реги-

стрирующих приборов [19], в течение которого зарегистрированное событие обрабатывается, другие же события, наступившие в этот период, теряются. Регистрирующие приборы при этом делятся на два вида: с непродлевающимся мёртвым временем и продлевающимся [19]. Длительность мёртвого времени может быть как детерминированной величиной, так и случайной. Задачи по оценке параметров потока событий в условиях наличия непродлевающегося мёртвого времени фиксированной длительности рассматривались в работах [10,11,16,17].

Однако достаточно открытым остается вопрос изучения потоков событий, когда мёртвое время является случайной величиной с тем или иным законом распределения. В частности, в [20] решена задача оценки параметра распределения непродлевающегося мёртвого времени случайной длительности в пуассоновском потоке событий; в [21] – в асинхронном потоке событий. В настоящей работе рассматривается дважды стохастический полусинхронный поток событий, функционирующий в условиях непродлевающегося случайного мёртвого времени, приводятся результаты статистических экспериментов, реализованных на имитационной модели изучаемого потока.

1. Математическая модель наблюдаемого потока

Рассматривается полусинхронный дважды стохастический поток событий (далее – поток), сопровождающий процесс (интенсивность) которого есть кусочно-постоянный стационарный случайный процесс $\lambda(t)$ с двумя состояниями S_1 и S_2. Будем говорить, что имеет место первое состояние процесса (потока) S_1, если $\lambda(t)=\lambda_1$, и, наоборот, имеет место второе состояние процесса (потока) S_2, если $\lambda(t)=\lambda_2$ ($\lambda_2>\lambda_1\geq 0$). Если имеет место первое состояние процесса S_1, то в течение временного интервала, когда $\lambda(t)=\lambda_1$, поступает пуассоновский поток событий с интенсивностью λ_1. Если имеет место второе состояние процесса S_2, то в течение временного интервала, когда $\lambda(t)=\lambda_2$, поступает пуассоновский поток событий с интенсивностью λ_2. Переход из состояния S_1 процесса $\lambda(t)$ в состояние S_2 возможен только в момент наступления события (свойство синхронных потоков), при этом этот переход осуществляется с вероятностью p (с вероятностью $1-p$ процесс $\lambda(t)$ остается в состоянии S_1). Переход из состояния S_2 процесса $\lambda(t)$ в состояние S_1 может осуществляться в произвольный момент времени, не связанный с моментом наступления события (свойство асинхронных потоков). При этом длительность пребывания процесса $\lambda(t)$ во втором состоянии есть случайная величина, распределенная по экспоненциальному закону $F(t)=1-e^{-\alpha t}$, $t\geq 0$, где α – интенсивность смены состояния S_2 на S_1. Т.к. переход из второго состояния в первое возможен в произвольный момент времени, т.е. не привязан к моменту наступления события во втором состоянии, поток называется полусинхронным дважды стохастическим потоком событий. Матрицы инфинитезимальных коэффициентов процесса $\lambda(t)$ при этом примут вид

$$D_0=\begin{Vmatrix}-\lambda_1 & 0\\ \alpha & -(\lambda_2+\alpha)\end{Vmatrix},\ D_1=\begin{Vmatrix}(1-p)\lambda_1 & p\lambda_1\\ 0 & \lambda_2\end{Vmatrix}.$$

Элементами матрицы D_1 являются интенсивности переходов процесса $\lambda(t)$ из состояния в состояние с наступлением события. Недиагональные элементы матрицы D_0 – интенсивности переходов из состояния в состояние без наступления события. Диагональные элементы матрицы D_0 – интенсивности выхода процесса из своих состояний, взятые с противоположным знаком. В сделанных предположениях $\lambda(t)$ – скрытый марковский процесс ($\lambda(t)$ – принципиально ненаблюдаемый процесс).

После каждого зарегистрированного события в момент времени t_k наступает период мёртвого времени случайной длительности, который порождается этим событием, так что другие события исходного потока, наступившие в течение этого периода мёртвого времени (периода ненаблюдаемости), недоступны наблюдению (теряются) и не вызывают его продления (непродлевающееся мёртвое время). Принимается, что случайная длительность мёртвого времени распределена по равномерному закону с плотностью вероятности $p(T) = 1/T^*$, где T – значение длительности мёртвого времени, $0 \le T \le T^*$.

Вариант возникающей ситуации приведен на рис. 1, где S_1 и S_2 – состояния случайного процесса $\lambda(t)$; временная ось $(0,t)$ – ось моментов наступления наблюдаемых событий в моменты времени t_1, t_2, …; временная ось $\left(0,t^{(1)}\right)$ – ось наступления событий в моменты времени $t_1^{(1)}$, $t_2^{(1)}$, … в первом (S_1) состоянии процесса $\lambda(t)$, на которой также указаны значения длительностей $T_1^{(1)}$, $T_2^{(1)}$, … мёртвых времен, порождаемых наблюдаемыми событиями потока; аналогично для временной оси $\left(0,t^{(2)}\right)$; белыми кружками обозначены наблюдаемые события, чёрными – ненаблюдаемые, штриховкой – периоды продолжительности мёртвого времени; траектория процесса $\lambda(t)$ привязана к временной оси $\left(0,t^{(1)}\right)$.

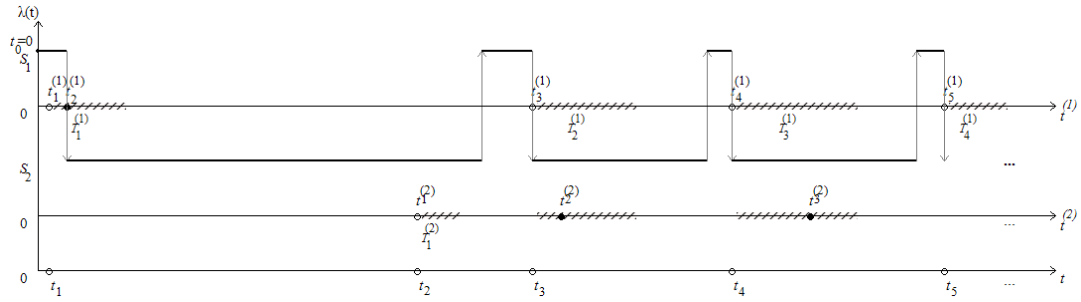


Рис. 1. Формирование наблюдаемого потока событий

Цель настоящей статьи: построение имитационной модели полусинхронного потока событий, функционирующего в условиях непродлевающегося случайного мёртвого времени, и постановка статистических экспериментов (испытаний) на этой имитационной модели для установления её работоспособности.

2. Имитационная модель полусинхронного потока событий при непродлевающемся случайном мёртвом времени

Метод имитационного моделирования широко используется на этапах исследования и проектирования различных систем. Основным средством реализации имитационного моделирования служит компьютер, позволяющий на основе математической модели исследуемой системы осуществить её цифровое моделирование.

Исследование математических моделей систем осуществляется либо аналитическими методами, либо имитационным моделированием. Основные преимущества имитационного моделирования по сравнению с аналитическими методами заключаются в следующем [22]: 1) бóльшая адекватность между сущностью физического явления (объекта) и математической моделью физического явления (объекта); 2) более широкий охват физических процессов и систем по сравнению с аналитическими методами; 3) имитационное моделирование функционирования процессов и систем при разнообраз-

ных законах распределения случайных величин как в установившихся, так и в переходных режимах; 4) получение в результате имитационного моделирования самых разнообразных статистических характеристик исследуемых процессов и систем.

На рис. 2 представлена блок-схема, описывающая формирование полусинхронного дважды стохастического потока событий с непродлевающимся случайным мёртвым временем, где x – равномерно распределенная на $[0,1]$ случайная величина; p – вероятность перехода из первого состояния во второе; τ – интервал времени между двумя соседними событиями; Tm – время моделирования; i – счётчик событий, наступивших в первом состоянии; j – счётчик событий, наступивших во втором состоянии; k – счётчик наблюдаемых событий; T – значение длительности мёртвого времени; $t_i^{(1)}$ – момент наступления i-го события, зафиксированного в первом состоянии; $t_j^{(2)}$ – момент наступления j-го события, зафиксированного во втором состоянии; $t^{(1)}$ – текущее время для первого состояния; $t^{(2)}$ – текущее время для второго состояния; t_k – момент наступления k-го наблюдаемого события; $t_T^{(1)}$ – момент наступления ненаблюдаемого события в первом состоянии; $t_T^{(2)}$ – момент наступления ненаблюдаемого события во втором состоянии; $t_\alpha^{(2)}$ – время, на протяжении которого события наступают во втором состоянии; τ_α, $t_{\lambda 2}^{(2)}$ – вспомогательные величины.

Для получения численных результатов (в соответствии с приведенной на рис. 2 блок-схемой) разработана программа, реализованная на языке C# в среде Microsoft Visual Studio. Результатом работы первой части программы является последовательность значений длительности интервалов между соседними событиями наблюдаемого потока и последовательность значений длительности периодов ненаблюдаемости событий исходного потока. Вторая часть программы – статистическая обработка полученных последовательностей данных.

3. Статистические эксперименты и численные результаты

С целью установления работоспособности имитационной модели полусинхронного потока событий поставлены следующие статистические эксперименты.

Первый статистический эксперимент состоит в построении оценки (выборочного среднего) длительности интервала $\boldsymbol{\tau}$ между двумя соседними событиями в наблюдаемом потоке $\hat{M}(\boldsymbol{\tau}) = \frac{1}{N}\sum_{k=1}^{N}\tau_k$, где $\tau_k = t_{k+1} - t_k$, $k = \overline{1,N}$ – значение длительности k-го интервала между соседними событиями наблюдаемого потока, и в определении зависимости этой оценки от времени моделирования Tm.

Эксперимент проводился при заданных параметрах потока: $\lambda_1 = 2$, $\lambda_2 = 1$, $\alpha = 0.1$, $p = 0.5$; параметр T^* равномерного распределения принимает значения: $T^* = 1, 2, 3$; значение N зависит от заданных значений λ_1, λ_2, α, p, T^*, Tm и конкретной реализации наблюдаемого потока.

В табл. 1 приведены результаты эксперимента. В первой строке табл. 1 указано время моделирования Tm (время наблюдения за потоком) ($Tm = 50, 100, \ldots, 700$ ед. времени); во второй, третьей и четвёртой строках указана оценка математического ожидания $\hat{M}(\boldsymbol{\tau})$ случайной величины τ – длительности интервала между соседними событиями наблюдаемого потока для $T^* = 1, 2, 3$ соответственно.

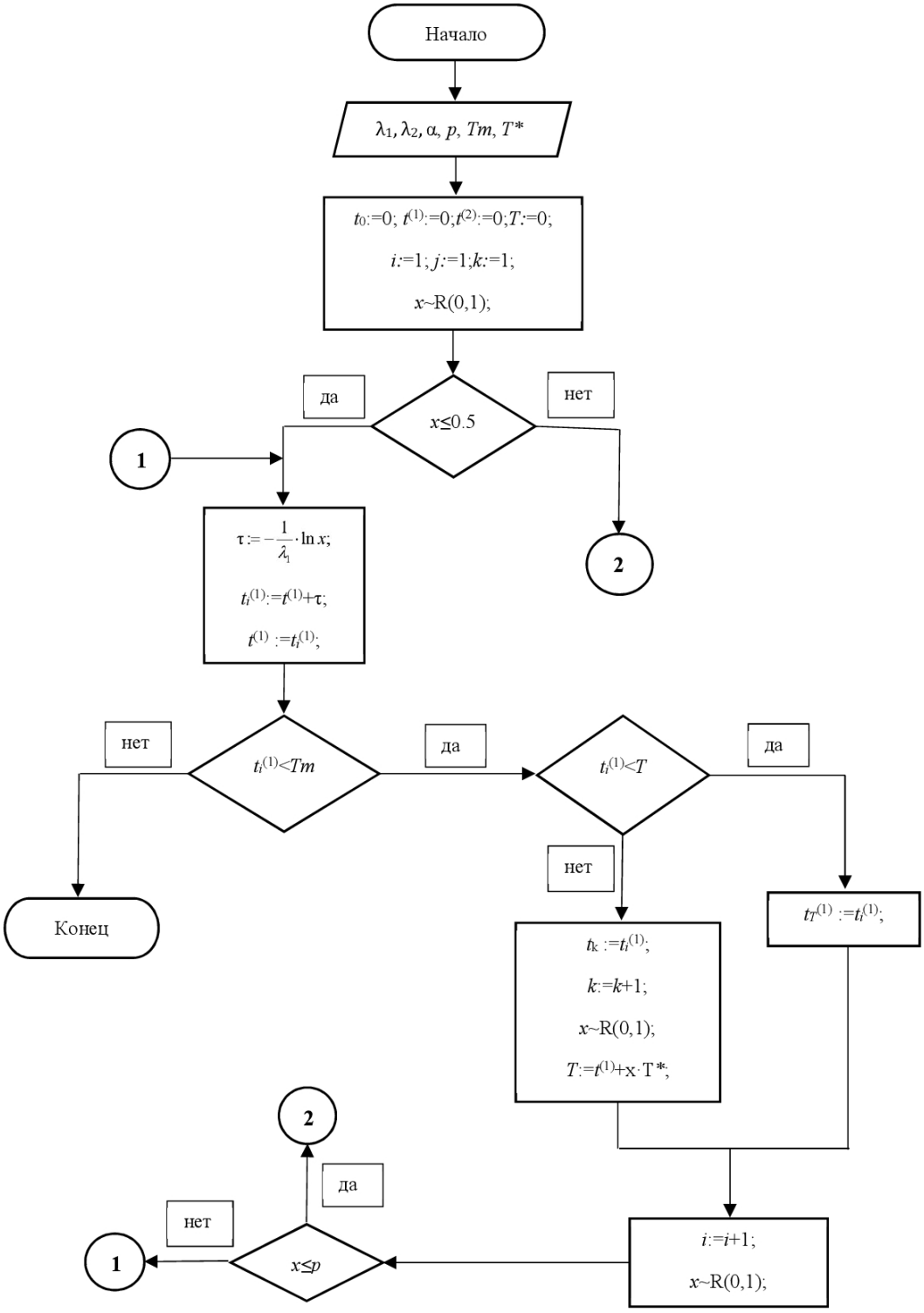


Рис. 2. Блок-схема алгоритма имитационного моделирования полусинхронного дважды стохастического потока событий в условиях непродлевающегося случайного мёртвого времени (часть 1)

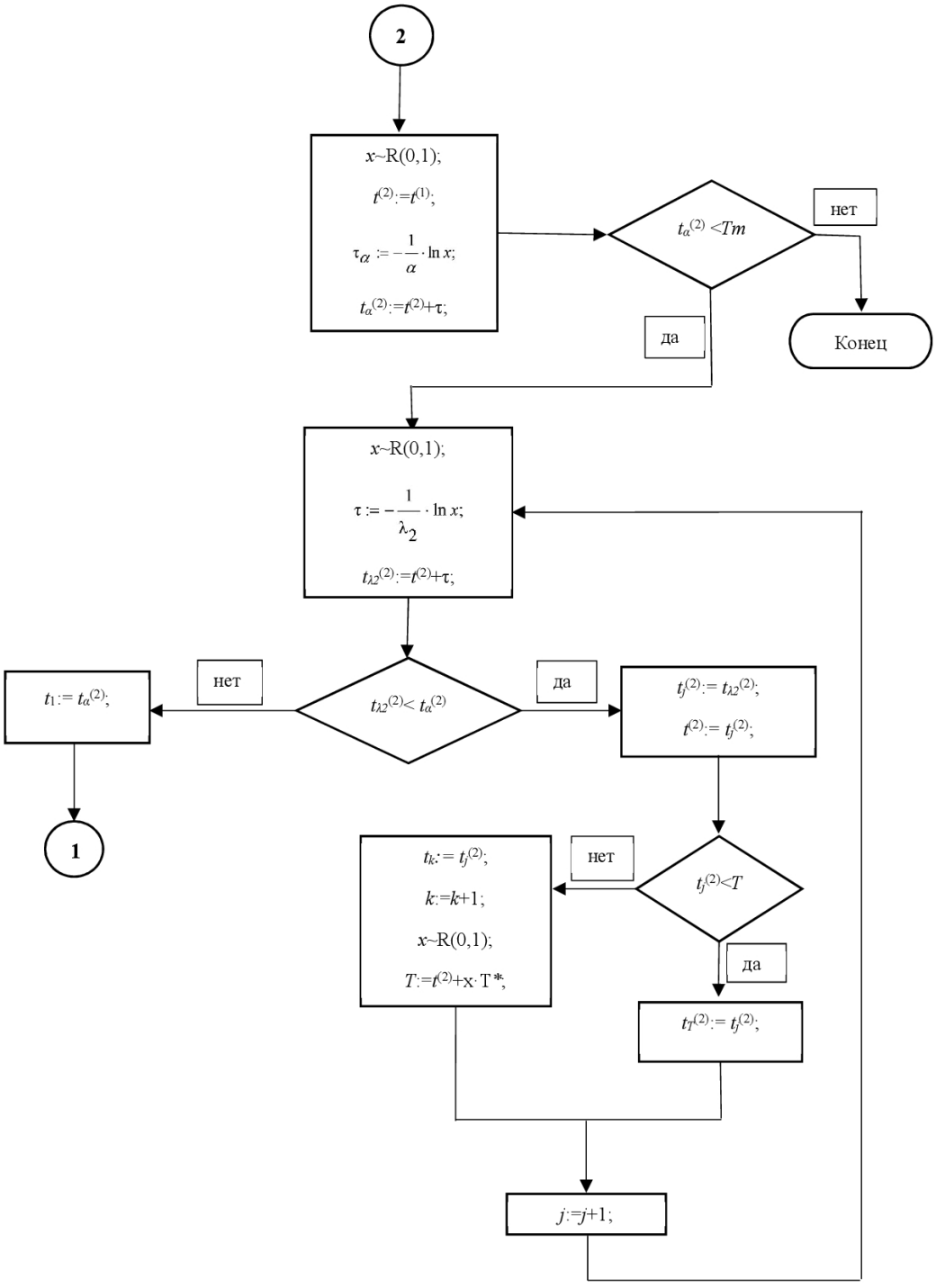


Рис. 2. Блок-схема алгоритма имитационного моделирования полусинхронного дважды стохастического потока событий в условиях непродлевающегося случайного мёртвого времени (часть 2)

Таблица 1

Численные результаты первого эксперимента

Tm		50	100	150	200	250	300	350
$\hat{M}(\tau)$	*T** = 1	1,50887	1,03514	1,16125	1,05862	1,24757	1,09302	1,09394
	*T** = 2	1,28875	1,15514	1,25984	1,28110	1,30888	1,24064	1,20108
	*T** = 3	1,31538	1,34170	1,27192	1,30861	1,45333	1,22709	1,43643
Tm		400	450	500	550	600	650	700
$\hat{M}(\tau)$	*T** = 1	1,21215	1,20782	1,06298	1,15509	1,20958	1,30751	1,11766
	*T** = 2	1,41243	1,39177	1,20390	1,29627	1,32519	1,39132	1,20439
	*T** = 3	1,32106	1,43355	1,35655	1,45906	1,35708	1,42697	1,44772

Для наглядности на рис. 3 приведены графики зависимости $\hat{M}(\tau)$ от времени моделирования *Tm*, построенные по данным табл. 1.

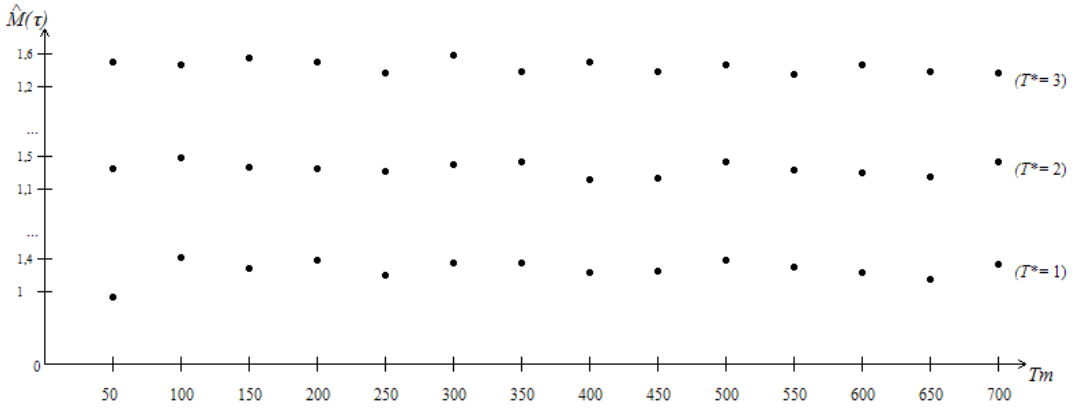


Рис. 3. Графики зависимости $\hat{M}(\tau)$ от значения *Tm*

Из анализа результатов табл. 1 и графиков зависимости $\hat{M}(\tau)$ от *Tm* следует, что стационарный режим функционирования наблюдаемого потока устанавливается при $Tm \geq 300$ ед. времени (оценка стремится к постоянному значению, т.е. при достаточно больших *Tm* оценка устанавливается в пределах своего стационарного значения). Кроме того, заметно увеличение значений оценки $\hat{M}(\tau)$ при увеличении *T**, что естественно, т.к. расширение отрезка $[0, T^*]$ – увеличении параметра распределения длительности мёртвого времени – влечет за собой увеличение значений длительности мёртвого времени, а, следовательно, бóльшее количество событий теряется и длительности интервалов τ между соседними наблюдаемыми событиями увеличиваются.

Второй статистический эксперимент состоит в построении оценки (выборочного среднего) длительности мёртвого времени (случайной величины $\boldsymbol{T}$) $\hat{M}(\boldsymbol{T}) = \frac{1}{N}\sum_{i=1}^{N} T_i$, где T_i – значение длительности *i*-го периода мёртвого времени, и установлении зависимости этой оценки от времени моделирования *Tm*.

Эксперимент проводился при заданных параметрах потока: $\lambda_1 = 2$, $\lambda_2 = 1$, $\alpha = 0.1$, $p = 0.5$; параметр *T** равномерного распределения принимает значения: $T^* = 1, 2, 3$; значение *N* зависит от заданных значений λ_1, λ_2, α, *p*, *T**, *Tm* и конкретной реализации наблюдаемого потока.

В табл. 2 приведены результаты эксперимента. Структура табл. 2 аналогична структуре табл. 1.

Таблица 2

Численные результаты второго эксперимента

Tm		50	100	150	200	250	300	350
$\hat{M}(\boldsymbol{T})$	*T** = 1	0,47327	0,46828	0,46579	0,50032	0,49063	0,46958	0,52671
	*T** = 2	0,93928	1,15202	0,88899	0,84028	0,90908	0,97145	1,19388
	*T** = 3	1,27924	1,24212	1,43369	1,46502	1,46485	1,53622	1,40277
Tm		400	450	500	550	600	650	700
$\hat{M}(\boldsymbol{T})$	*T** = 1	0,48698	0,50775	0,53038	0,48766	0,48897	0,52344	0,50721
	*T** = 2	0,98142	0,94660	0,96821	1,03925	1,00466	1,04405	0,99176
	*T** = 3	1,61113	1,47313	1,59487	1,54814	1,55912	1,54379	1,60621

Для наглядности на рис. 4 приведены графики зависимости $\hat{M}(\boldsymbol{T})$ от времени моделирования *Tm*, построенные по данным табл. 2.

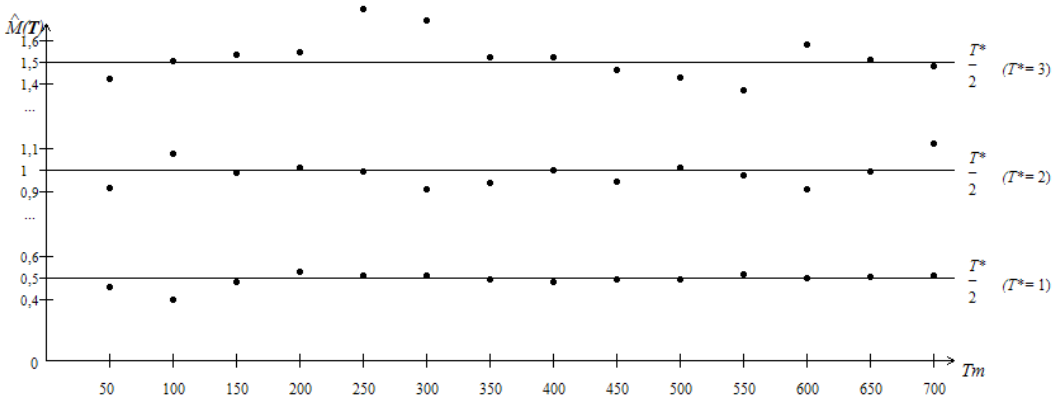


Рис. 4. Графики зависимости $\hat{M}(\boldsymbol{T})$ от значения *Tm*

Из графиков, построенных по данным табл. 2, виден колебательный характер оценки $\hat{M}(\boldsymbol{T})$, которая с увеличением времени моделирования *Tm* стремится к постоянному значению, т.е. при достаточно большом *Tm* оценка $\hat{M}(\boldsymbol{T})$ устанавливается в пределах своего стационарного значения, соответствующего параметру *T**. Например, стационарный режим функционирования наблюдаемого потока устанавливается для $T^*=1$ при $Tm \geq 250$ ед. времени ($\hat{M}(\boldsymbol{T})$ стремится к $M(\boldsymbol{T})=1/2$, начиная с $Tm=250$ ед. времени); для $T^*=2$ – при $Tm \geq 400$ ед. времени ($\hat{M}(\boldsymbol{T})$ стремится к $M(\boldsymbol{T})=1$, начиная с $Tm=400$ ед. времени); для $T^*=3$ – при $Tm \geq 600$ ед. времени ($\hat{M}(\boldsymbol{T})$ стремится к $M(\boldsymbol{T})=3/2$, начиная с $Tm=600$ ед. времени). Кроме того, также, как и в первом эксперименте, видно увеличение значения оценки длительности мёртвого времени $\hat{M}(\boldsymbol{T})$ при увеличении параметра *T**, что вполне естественно.

Заключение

В настоящей работе разработана и исследована имитационная модель полусинхронного дважды стохастического потока событий, функционирующего в условиях непродлевающегося случайного мёртвого времени. Приведенные результаты статистических экспериментов, а также другие многочисленные испытания имитационной модели подтверждают её работоспособность и соответствие результатов моделирования физическому представлению о протекающих процессах формирования наблюдаемого потока событий.

ЛИТЕРАТУРА

1. *Cox D.R.* The analysis of non-Markovian stochastic processes by the inclusion of supplementary variables // Proceedings of the Cambridge Philosophical Society. – 1955. – 51(3). – P. 433–441.

2. *Kingman Y.F.C.* On doubly stochastic Poisson process // Proceedings of the Cambridge Philosophical Society.– 1964. – 60(4). – P. 923–930.

3. *Башарин Г.П., Кокотушкин В.А., Наумов В.А.* О методе эквивалентных замен расчета фрагментов сетей связи. Ч.1 // Изв. АНСССР. Техн. Кибернетика. – 1979. – № 6. – С. 92–99.

4. *Neuts M.F.* A versatile Markovian point process // Journal of Applied Probability. 1979. – 16. – P. 764–779.

5. *Lucantoni D.M.* New results on the single server queue with a batch markovian arrival process // Communication in Statistics Stochastic Models. 1991. – 7. – P. 1–46.

6. *Diamond J.E., Alfa A.S.* The MAP/PH/1 retrial queue // Communication in Statistics Stochastic Models. – 1998. – 14. – P. 1151–1177.

7. *Basharin G.P., Gaidamaka Yu.V., Samouylov K.E.* Mathematical theory of teletraffic and its application to the analysis of multiservice communication of next generation networks // Automatic Control and Computer Science. – 47(2). – P. 62–69.

8. *Вишневский В.М., Ларионов А.А.* Открытая сеть массового обслуживания с коррелированными входными потоками для оценки производительности широкополосных беспроводных сетей // Информационные технологии и математическое моделирование (ИТММ-2016): материалы XV Междунар. конф., Казань, 12–16 сентября 2016. – Томск: Изд-во ТГУ, 2016. – Ч.1. – С. 36–50.

9. *Дудин А.Н., Клименок В.И.* Расчет необходимого числа каналов в современных телекоммуникационных сетях // Информатизация образования. – 2005. – № 4. – С. 56–68.

10. *Горцев А.М., Нежельская Л.А.* Оценивание длительности мёртвого времени и параметров синхронного альтернирующего потока событий // Вестник Томского государственного университета. – 2003. – № S6. – С. 232–239.

11. *Gortsev A.M., Nezhel'skaya L.A.* Estimation of parameters of synchronously alternating Poisson stream of events by the moment method // Telecommunications and Radio Engineering. – 1996. – 50(1). – P. 56–63.

12. *Горцев А.М., Нежельская Л.А.* Оценивание параметров синхронного дважды стохастического потока событий методом моментов // Вестник Томского государственного университета. – 2002. – № S1-1. – С. 24–29.

13. *Горцев А.М., Нежельская Л.А.* Оценивание параметров асинхронного потока с инициированием лишних событий методом моментов // Вестник Томского государственного университета. – 2006. – № S18. – С. 267–273.

14. *Леонова М.А., Нежельская Л.А.* Вероятность ошибки при оценивании состояний обобщенного асинхронного потока событий // Вестник Томского государственного университета. Управление, вычислительная техника и информатика. – 2012. – № 2(19). – С. 88–101.

15. *Горцев А.М., Зуевич В.Л.* Оптимальная оценка состояний асинхронного дважды стохастического потока событий с произвольным числом состояний // Вестник Томского государственного университета. Управление, вычислительная техника и информатика. – 2010. – № 2(11). – С. 44–65.

16. *Калягин А.А., Нежельская Л.А.* Сравнение МП- и ММ-оценок длительности мёртвого времени в обобщенном полусинхронном потоке событий // Вестник Томского государственного университета. Управление, вычислительная техника и информатика. – 2015. – № 3(32). – С. 23–32.

17. *Горцев А.М., Калягин А.А., Нежельская Л.А.* Оценка максимального правдоподобия длительности мёртвого времени в обобщенном полусинхронном потоке // Вестник Томского государственного университета. Управление, вычислительная техника и информатика. – 2015. – № 1(30). – С. 27–37.

18. *Горцев А.М., Нежельская Л.А.* Оценивание параметров полусинхронного дважды стохастического потока событий методом моментов // Вестник Томского государственного университета. – 2002. – № S1-1. – С. 18–23.

19. *Апанасович В.В., Коляда А.А., Чернявский А.Ф.* Статистический анализ случайных потоков в физическом эксперименте. – Минск: Университетское, 1988. – 256 с.

20. *Горцев А.М., Завгородняя М.Е.* Оценивание параметра непродлевающегося мёртвого времени случайной длительности в пуассоновском потоке событий // Вестник Томского государственного университета. Управление, вычислительная техника и информатика. – 2017. – № 40. – С. 32–40.

21. *Васильева Л.А.* Оценивание параметров дважды стохастического потока событий в условиях присутствия мёртвого времени // Вестник Томского государственного университета. – 2002. – № S1-1. – С. 9–13.

22. *Лифшиц А.Л., Мальц Э.А.* Статистическое моделирование систем массового обслуживания. – М.: Сов. Радио, 1978. – 248 с.

СТАТИСТИЧЕСКИЕ ЭКСПЕРИМЕНТЫ НА ИМИТАЦИОННОЙ МОДЕЛИ ОБОБЩЁННОГО АСИНХРОННОГО ПОТОКА СОБЫТИЙ ПРИ НЕПРОДЛЕВАЮЩЕМСЯ СЛУЧАЙНОМ МЁРТВОМ ВРЕМЕНИ

Л.А. Нежельская, А.А. Першина
Томский государственный университет

Введение

Распространёнными математическими моделями физических явлений и процессов являются потоки событий. В частности, такие модели применяются при исследовании информационных потоков сообщений, в телекоммуникационных системах, в спутниковых сетях связи т.п. [1,2]. В большинстве публикаций авторы рассматривают математические модели потоков событий, когда события потока доступны наблюдению. Однако на практике возникают ситуации, когда наступившее событие влечет за собой ненаблюдаемость последующих событий. Причиной ненаблюдаемости, как правило, выступает мёртвое время регистрирующих приборов [3], в течение которого зарегистрированное событие обрабатывается; другие же события, наступившие в этот период, теряются.

Регистрирующие приборы при этом делятся на два вида: с непродлевающимся мёртвым временем и продлевающимся [3]. При этом длительность мёртвого времени может быть как детерминированной величиной, так и случайной. Задачи по оценке параметров и состояний потока событий в условиях наличия мёртвого времени фиксированной длительности, в частности, рассматривались в работах [4–15]. При этом в [4–11] получены результаты для непродлевающегося мёртвого времени, в [12–15] – для продлевающегося.

Однако достаточно открытым остается вопрос изучения потоков событий, когда мёртвое время является случайной величиной с тем или иным законом распределения. Здесь отметим работу [16], в которой решается задача оценки параметра распределения непродлевающегося мёртвого времени случайной длительности в пуассоновском потоке событий, и работу [17], в которой находятся формулы для начальных моментов общего периода ненаблюдаемости в пуассоновском потоке событий при наличии продлевающегося случайного мёртвого времени. В настоящей статье рассматривается обобщённый асинхронный поток событий, функционирующий в условиях непродлевающегося случайного мёртвого времени, приводятся результаты статистических экспериментов, реализованных на имитационной модели изучаемого потока.

1. Математическая модель наблюдаемого потока

Рассматривается обобщенный асинхронный дважды стохастический поток событий, сопровождающий процесс которого есть кусочно-постоянный стационарный случайный процесс $\lambda(t)$ с двумя состояниями λ_1 и λ_2 ($\lambda_1 > \lambda_2 \geq 0$). В течение временного интервала, когда $\lambda(t) = \lambda_i$, имеет место пуассоновский поток событий с интенсивностью λ_i, $i = 1,2$. Переход из первого состояния процесса $\lambda(t)$ во второе (из второго в первое) может осуществляться в произвольный момент времени. При этом длительность пребывания процесса $\lambda(t)$ в i-м состоянии распределена по экспоненциальному закону с параметром α_i, $i = 1,2$. При переходе процесса $\lambda(t)$ из первого состояния во второе инициируется с вероятностью p ($0 \leq p \leq 1$) дополнительное событие во втором состоянии (т.е. сначала осуществляется переход, а затем инициируется дополнительное событие). Наоборот, при переходе процесса $\lambda(t)$ из второго состояния в первое инициируется с вероятностью q ($0 \leq q \leq 1$) дополнительное событие в первом состоянии. При этом матрицы инфинитезимальных коэффициентов примут вид

$$D_0 = \begin{Vmatrix} -(\lambda_1 + \alpha_1) & (1-p)\alpha_1 \\ (1-q)\alpha_2 & -(\lambda_2 + \alpha_2) \end{Vmatrix},\ D_1 = \begin{Vmatrix} \lambda_1 & p\alpha_1 \\ q\alpha_2 & \lambda_2 \end{Vmatrix}.$$

Элементами матрицы D_1 являются интенсивности переходов процесса $\lambda(t)$ из состояния в состояние с наступлением события. Недиагональные элементы матрицы D_0 – интенсивности переходов из состояния в состояние без наступления события. Диагональные элементы матрицы D_0 – интенсивности выхода процесса $\lambda(t)$ из своих состояний, взятые с противоположным знаком. В сделанных предположениях $\lambda(t)$ – скрытый марковский процесс ($\lambda(t)$ – принципиально ненаблюдаемый процесс). После каждого зарегистрированного события в момент времени t_k наступает период мёртвого времени случайной длительности, который порождается этим событием, так что другие события исходного потока, наступившие в течение этого периода мёртвого времени (периода ненаблюдаемости), недоступны наблюдению (теряются) и не вызывают его продления(непродлевающееся мёртвое время). Принимается, что случайная длительность мёртвого времени распределена по равномерному закону с плотностью вероятности $p(T) = 1/T^*$, где T – значение длительности мёртвого времени, $0 \le T \le T^*$. Вариант возникающей ситуации приведен на рис. 1, где 1,2 – состояния случайного процесса $\lambda(t)$; дополнительные события, которые могут наступать в момент перехода процесса $\lambda(t)$ из состояния в состояние, помечены буквами p либо q; на первой временной оси белыми кружками обозначены события исходного обобщённого асинхронного потока; на второй временной оси приведена схема создания мёртвого времени, где серым цветом обозначены значения периодов мёртвого времени, порождаемые каждым зарегистрированным событием исходного потока, черными кружками обозначены ненаблюдаемые события; на третьей временной оси приведены моменты t_1, t_2,… наступления событий наблюдаемого потока, обозначенные белыми кружками.

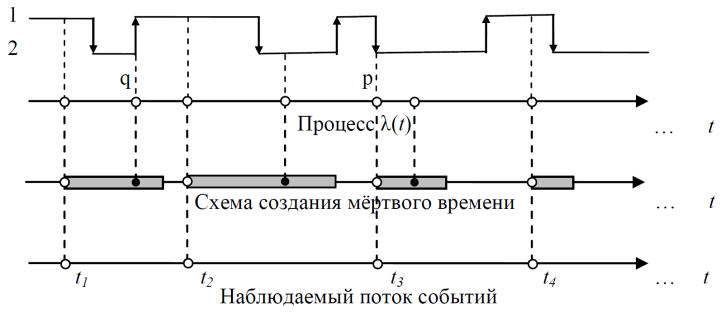


Рис. 1. Обобщённый асинхронный поток событий в условиях непродлевающегося случайного мёртвого времени

Настоящая статья посвящена вопросу построения имитационной модели обобщённного асинхронного потока событий, функционирующего в условиях непродлевающегося случайного мёртвого времени, с целью проведения статистических экспериментов на этой модели для изучения содержательной информации о статистических характеристиках периода ненаблюдаемости.

2. Имитационная модель обобщённого асинхронного потока событий при непродлевающемся случайном мёртвом времени

Метод имитационного моделирования широко применяется на этапе проектирования сложных систем. Основным средством реализации имитационного моделирования служит компьютер, позволяющий осуществлять цифровое моделирование систем и сигналов. При имитационном моделировании используемая математическая модель воспроизводит алгоритм функционирования исследуемой системы во времени при различных сочетаниях значений параметров системы. При исследовании процесса мето-

дом имитационного моделирования должно быть реализовано наблюдение за ним с течением времени.

Основные преимущества имитационного моделирования по сравнению с аналитическими методами заключаются в следующем [18]: 1) большая адекватность между сущностью физического явления (объекта) и математической моделью физического явления (объекта); 2) более широкий охват физических процессов и систем по сравнению с аналитическими методами; 3) имитационное моделирование функционирования процессов и систем при разнообразных законах распределения случайных величин как в установившихся, так и в переходных режимах; 4) получение в результате имитационного моделирования самых разнообразных статистических характеристик исследуемых процессов и систем.

Для получения численных результатов разработана программа. Она реализована на языке программирования C# в среде Microsoft Visual Studio 2017. Первый этап расчета предполагает имитационное моделирование исходного потока событий, схемы создания непродлевающегося случайного мёртвого времени и наблюдаемого потока событий. Результатом работы имитационной модели является последовательность значений длительности периодов ненаблюдаемости событий исходного потока $\{T_1,...,T_n\}$, где T_n – значение длительности мёртвого времени, порождённое последним наблюдаемым событием на интервале моделирования *Tm*.

На рис. 2–4 представлена блок-схема алгоритма имитационного моделирования потока событий в условиях непродлевающегося случайного мёртвого времени. Введены следующие обозначения: *Tm* – время моделирования; *t*1 – момент наступления события в обобщённом асинхронном потоке; *t*2 – момент окончания длительности мёртвого времени; *t*3 – момент наступления события в наблюдаемом потоке; *T* –значение длительности мёртвого времени; γ – значение случайной величины, равномерно распределенной на отрезке [0,1]; τ – интервал времени между соседними событиями обобщенного асинхронного потока событий; *x*, *z* – вспомогательные переменные; *State* – переменная, обозначающая состояние, в котором процесс находится в момент моделирования; *DurationState* – длительность нахождения процесса в состоянии *State*; *EndState* – момент окончания нахождения процесса в состоянии *State*; *Discrete*(*p*) – функция, которая в зависимости от параметра *p* принимает случайное значение 0 или 1 (с вероятностью *p* принимает значение 1, с вероятностью $1-p$ принимает значение 0; *Discrete*(*q*) – функция, которая в зависимости от параметра *q* принимает случайное значение 0 или 1 (с вероятностью *q* принимает значение 1, с вероятностью $1-q$ принимает значение 0; *n*1 – количество событий в обобщённом асинхронном потоке.

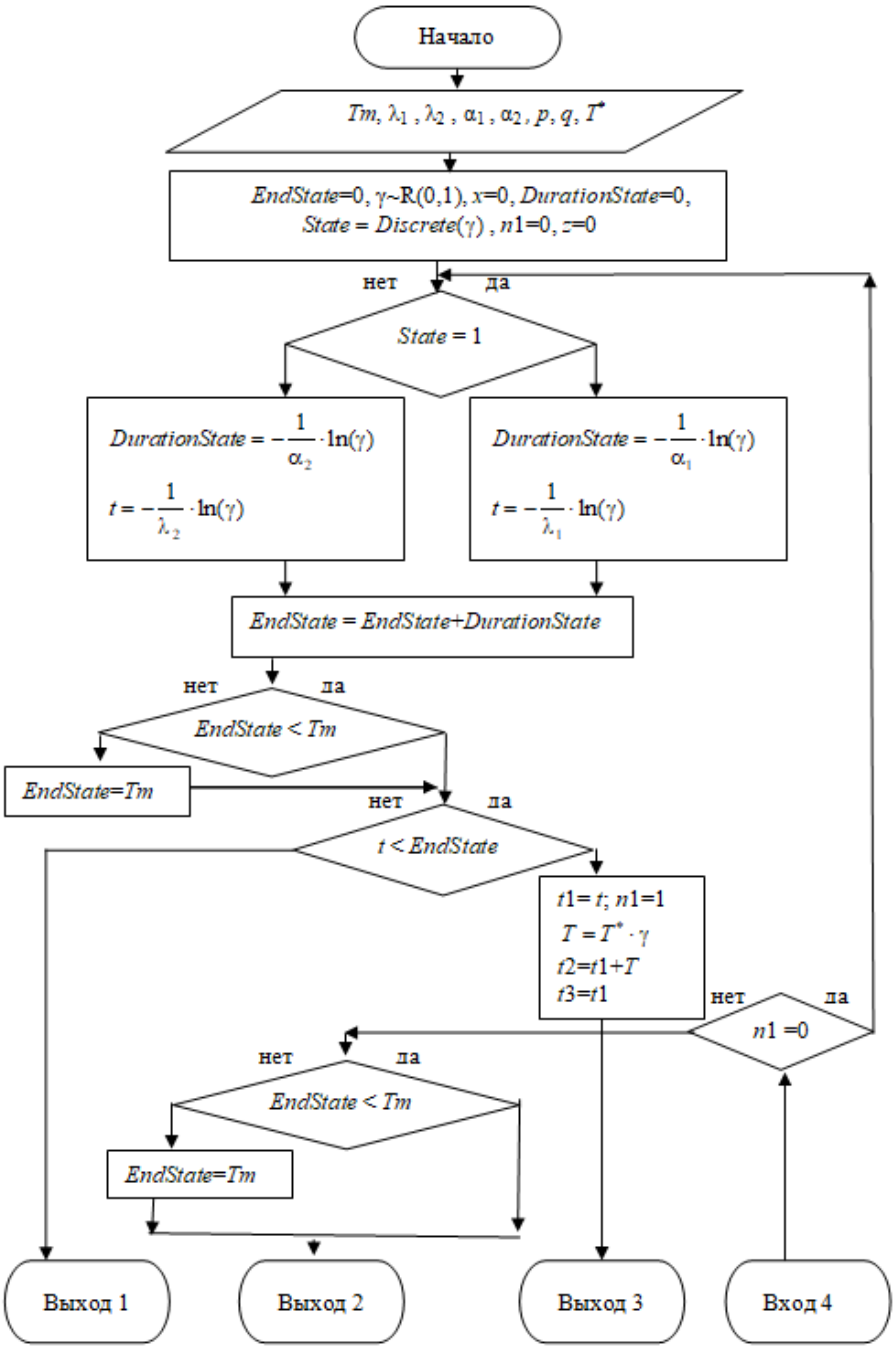


Рис. 2. Блок-схема алгоритма имитационного моделирования обобщённого асинхронного потока событий в условиях непродлевающегося случайного мёртвого времени (часть 1)

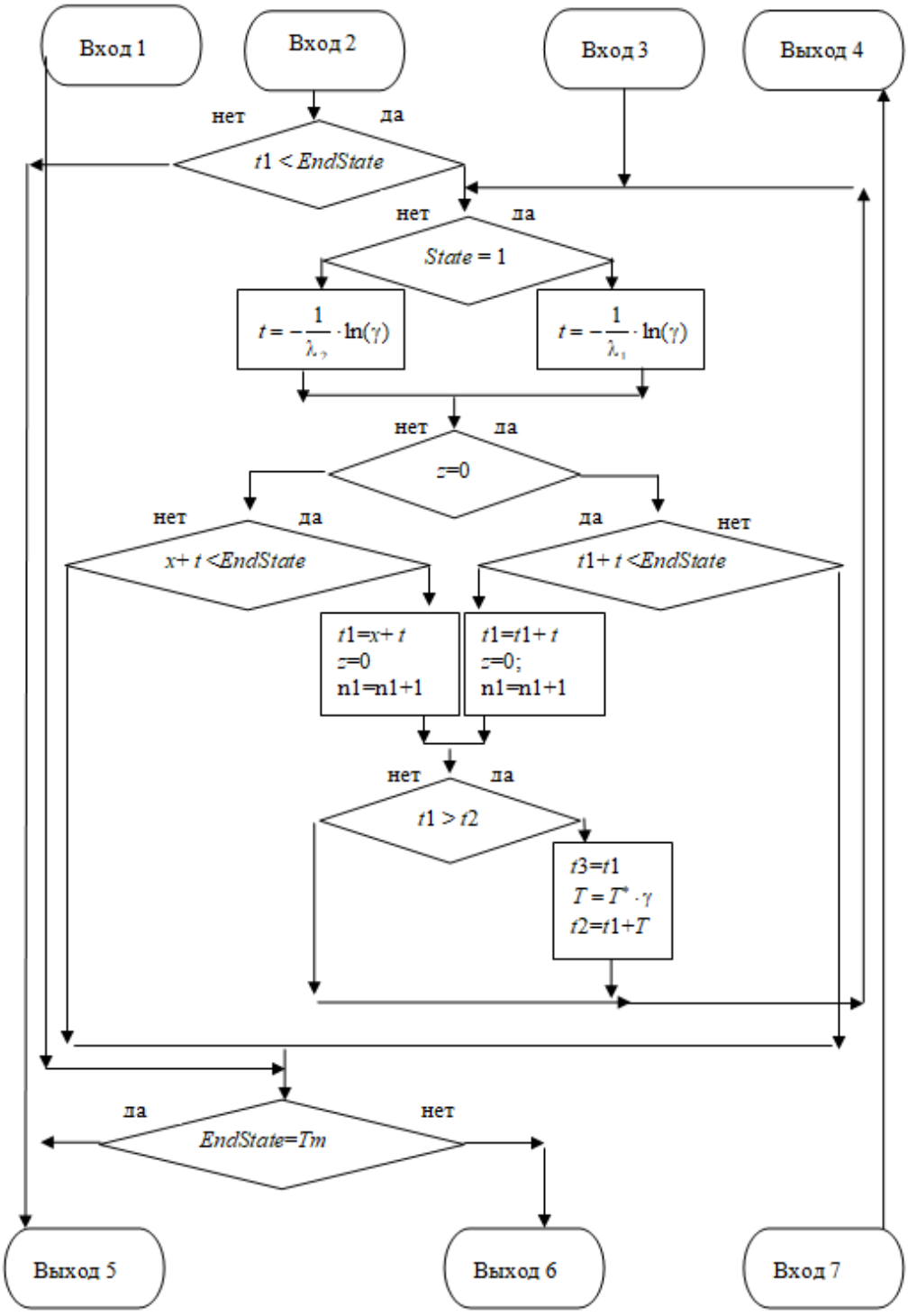


Рис. 3. Блок-схема алгоритма имитационного моделирования обобщённого асинхронного потока событий в условиях непродлевающегося случайного мёртвого времени (часть 2)

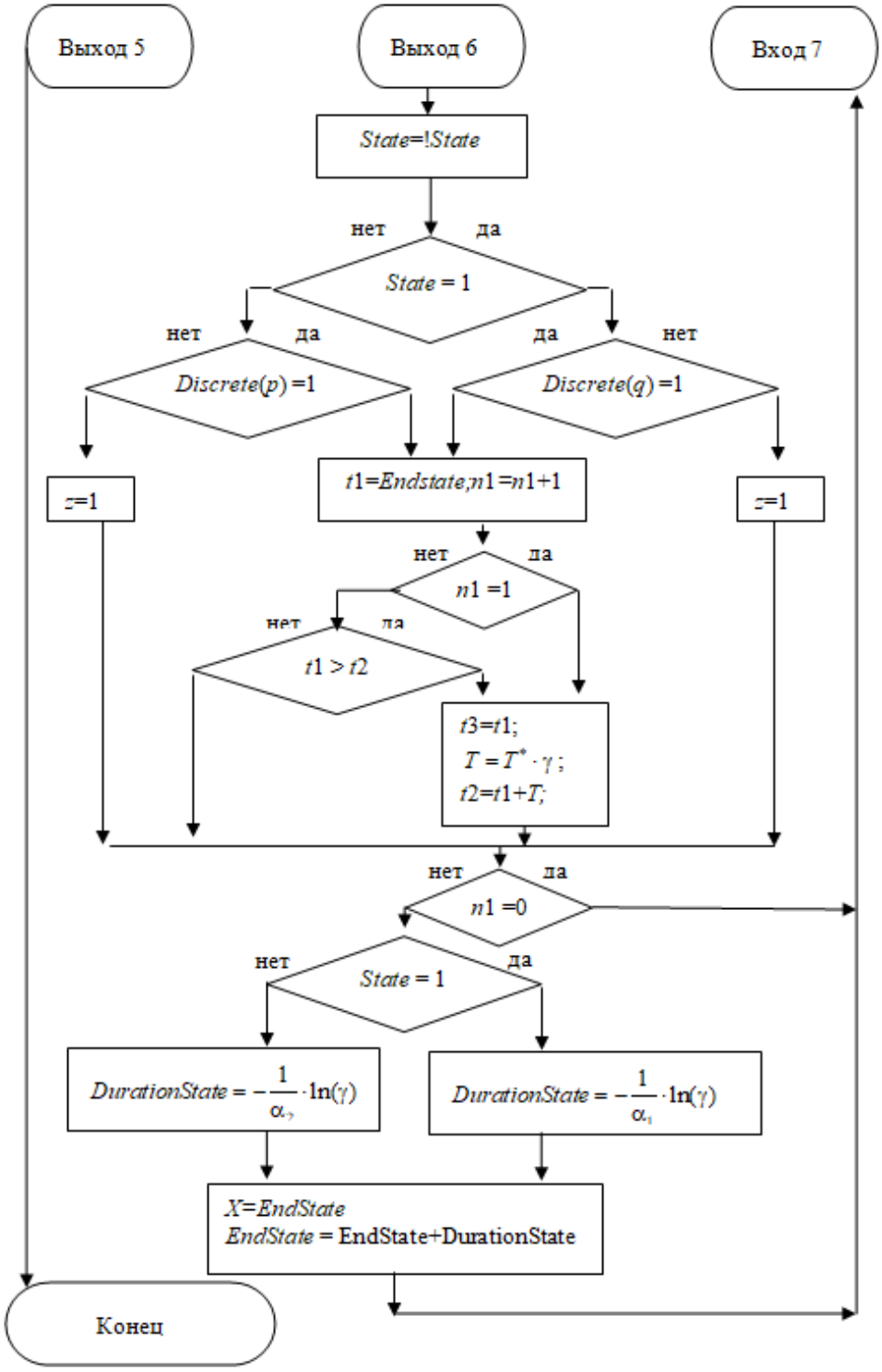


Рис. 4. Блок-схема алгоритма имитационного моделирования обобщённого асинхронного потока событий в условиях непродлевающегося случайного мёртвого времени (часть 3)

3. Статистические эксперименты и численные результаты

С целью получения численных значений периода ненаблюдаемости (мёртвого времени) поставлены статистические эксперименты.

Первый статистический эксперимент: 1) при заданных значениях параметров λ_1, λ_2, α_1, α_2, p, q, T^* и заданном времени моделирования Tm осуществляется имитационное моделирование наблюдаемого потока; выходом имитационной модели являются последовательность значений T_j, $j=\overline{1,N}$, – значение периода ненаблюдаемости, порождённого наблюдаемым событием исходного потока; значение N зависит от заданных значений λ_1, λ_2, α_1, α_2, p, q, T^*, Tm и конкретной реализации наблюдаемого потока; 2) в рамках конкретной реализации наблюдаемого потока вычисляются следующие статистические характеристики [19]:

$\hat{M}(\boldsymbol{T})=\frac{1}{N}\sum_{j=1}^{N}T_j$ – оценка математического ожидания периода ненаблюдаемости;

$\hat{V}(\boldsymbol{T})=\frac{1}{N}\sum_{j=1}^{N}\left[T_j-\frac{T^*}{2}\right]^2$ – оценка вариации периода ненаблюдаемости от истинного значения математического ожидания периода ненаблюдаемости (T^* – известный параметр имитационной модели).

3) после чего время моделирования Tm увеличивается и шаги 1), 2) повторяются.

В табл. 1–6 приведены результаты первого статистического эксперимента.

Таблица 1

Первый статистический эксперимент при $\lambda_1 = 3$, $\lambda_2 = 2$, $\alpha_1 = 0.5$, $\alpha_2 = 0.5$, $p = 0.5$, $q = 0.5$, $T^* = 1$

Tm	**10**	**20**	**30**	**40**	**50**
$\hat{M}(\boldsymbol{T})$	0,568154	0,465002	0,449866	0,505249	0,489746
$\hat{V}(\boldsymbol{T})$	0,046367	0,108388	0,091552	0,053454	0,102945

Таблица 2

Первый статистический эксперимент при $\lambda_1 = 3$, $\lambda_2 = 2$, $\alpha_1 = 0.5$, $\alpha_2 = 0.5$, $p = 0.5$, $q = 0.5$, $T^* = 1$

Tm	**60**	**70**	**80**	**90**	**100**
$\hat{M}(\boldsymbol{T})$	0,457656	0,475564	0,485832	0,534109	0,487774
$\hat{V}(\boldsymbol{T})$	0,083847	0,083804	0,080197	0,086793	0,090204

Таблица 3

Первый статистический эксперимент при $\lambda_1 = 3$, $\lambda_2 = 2$, $\alpha_1 = 0.5$, $\alpha_2 = 0.5$, $p = 0.5$, $q = 0.5$, $T^* = 1$

Tm	**200**	**300**	**400**	**500**	**600**
$\hat{M}(\boldsymbol{T})$	0,48061	0,477752	0,500716	0,504716	0,491727
$\hat{V}(\boldsymbol{T})$	0,091396	0,085622	0,082044	0,080272	0,085313

Таблица 4

Первый статистический эксперимент при $\lambda_1 = 3$, $\lambda_2 = 2$, $\alpha_1 = 0.5$, $\alpha_2 = 0.5$, $p = 0.5$, $q = 0.5$, $T^* = 1$

Tm	**700**	**800**	**900**	**1 000**	**2 000**
$\hat{M}(\boldsymbol{T})$	0,496649	0,483596	0,510026	0,496606	0,498506
$\hat{V}(\boldsymbol{T})$	0,081164	0,082235	0,08134	0,085757	0,083887

Таблица 5

Первый статистический эксперимент при $\lambda_1 = 3$, $\lambda_2 = 2$, $\alpha_1 = 0.5$, $\alpha_2 = 0.5$, $p = 0.5$, $q = 0.5$, $T^* = 1$

Tm	**3 000**	**4 000**	**5 000**	**6 000**	**7 000**
$\hat{M}(\boldsymbol{T})$	0,505192	0,501617	0,499068	0,49942	0,497584
$\hat{V}(\boldsymbol{T})$	0,082071	0,082135	0,083589	0,082173	0,083411

Таблица 6

Первый статистический эксперимент при $\lambda_1 = 3$, $\lambda_2 = 2$, $\alpha_1 = 0.5$, $\alpha_2 = 0.5$, $p = 0.5$, $q = 0.5$, $T^* = 1$

Tm	**8 000**	**9 000**	**10 000**
$\hat{M}(\boldsymbol{T})$	0,498456	0,502564	0,503289
$\hat{V}(\boldsymbol{T})$	0,083767	0,084204	0,083922

Из анализа табл. 1–6 следует, что для оценки математического ожидания $\hat{M}(\boldsymbol{T})$ стационарный режим устанавливается ($\hat{M}(\boldsymbol{T})$ стремится к $T^*/2 = 0.5$ с увеличением Tm) для $Tm > 1000$.

Второй статистический эксперимент: 1) при заданных значениях параметров λ_1, λ_2, α_1, α_2, p, q, T^* и заданном времени моделирования Tm осуществляется имитационное моделирование наблюдаемого потока: выходом имитационной модели является последовательность значений τ_j, $j = \overline{1, N}$, – значение интервала времени $\boldsymbol{\tau}$ между соседними событиями наблюдаемого потока; значение N зависит от заданных значений λ_1, λ_2, α_1, α_2, p, q, T^*, Tm и конкретной реализации наблюдаемого потока; 2) в рамках конкретной реализации наблюдаемого потока вычисляются следующие статистические характеристики [19]:

$\hat{M}(\boldsymbol{\tau}) = \dfrac{1}{N}\sum_{j=1}^{N}\tau_j$ – оценка математического ожидания длительности интервала времени между соседними событиями наблюдаемого потока;

$\hat{D}(\boldsymbol{\tau}) = \dfrac{1}{N}\sum_{j=1}^{N}\left[\tau_j - \hat{M}(\boldsymbol{\tau})\right]^2$ – оценка дисперсии длительности интервала времени между соседними событиями наблюдаемого потока от выборочного математического ожидания $\hat{M}(\boldsymbol{\tau})$;

3) после чего время моделирования Tm увеличивается и шаги 1), 2) повторяются.

В табл. 6–12 приведены результаты второго статистического эксперимента.

Таблица 7

Второй статистический эксперимент при $\lambda_1 = 3$, $\lambda_2 = 2$, $\alpha_1 = 0.5$, $\alpha_2 = 0.5$, $p = 0.5$, $q = 0.5$, $T^* = 1$

Tm	**10**	**20**	**30**	**40**	**50**
$\hat{M}(\boldsymbol{\tau})$	1,012281	0,826491	0,820386	0,987154	0,885746
$\hat{D}(\boldsymbol{\tau})$	0,20338	0,165954	0,22842	0,154153	0,289973

Таблица 8

Второй статистический эксперимент при $\lambda_1 = 3$, $\lambda_2 = 2$, $\alpha_1 = 0.5$, $\alpha_2 = 0.5$, $p = 0.5$, $q = 0.5$, $T^* = 1$

Tm	**60**	**70**	**80**	**90**	**100**
$\hat{M}(\boldsymbol{\tau})$	0,849386	0,783556	0,828049	0,984978	0,92225
$\hat{D}(\boldsymbol{\tau})$	0,206693	0,188462	0,203336	0,297066	0,264798

Таблица 9

Второй статистический эксперимент при $\lambda_1 = 3$, $\lambda_2 = 2$, $\alpha_1 = 0.5$, $\alpha_2 = 0.5$, $p = 0.5$, $q = 0.5$, $T^* = 1$

Tm	**200**	**300**	**400**	**500**	**600**
$\hat{M}(\boldsymbol{\tau})$	0,860066	0,843498	0,872284	0,846505	0,864165
$\hat{D}(\boldsymbol{\tau})$	0,202118	0,210976	0,208881	0,193503	0,240862

Таблица 10

Второй статистический эксперимент при $\lambda_1 = 3$, $\lambda_2 = 2$, $\alpha_1 = 0.5$, $\alpha_2 = 0.5$, $p = 0.5$, $q = 0.5$, $T^* = 1$

Tm	**700**	**800**	**900**	**1 000**	**2 000**
$\hat{M}(\tau)$	0,854676	0,859214	0,901616	0,862568	0,866418
$\hat{D}(\tau)$	0,213439	0,234852	0,244523	0,227471	0,227866

Таблица 11

Второй статистический эксперимент при $\lambda_1 = 3$, $\lambda_2 = 2$, $\alpha_1 = 0.5$, $\alpha_2 = 0.5$, $p = 0.5$, $q = 0.5$, $T^* = 1$

Tm	**3 000**	**4 000**	**5 000**	**6 000**	**7 000**
$\hat{M}(\tau)$	0,878448	0,864936	0,868479	0,867951	0,872117
$\hat{D}(\tau)$	0,229027	0,222929	0,228942	0,228718	0,236436

Таблица 12

Второй статистический эксперимент при $\lambda_1 = 3$, $\lambda_2 = 2$, $\alpha_1 = 0.5$, $\alpha_2 = 0.5$, $p = 0.5$, $q = 0.5$, $T^* = 1$

Tm	**8 000**	**9 000**	**10 000**
$\hat{M}(\tau)$	0,860144	0,874084	0,873261
$\hat{D}(\tau)$	0,2205	0,231285	0,225704

Анализируя численные результаты, представленные в табл. 7–12, можно сделать следующий вывод: стационарный режим для $\hat{M}(\tau)$ устанавливается ($\hat{M}(\tau)$ стремится к константе с увеличением Tm) для $Tm > 1000$.

Заключение

В настоящей статье разработана имитационная модель обобщённого асинхронного потока событий, функционирующего в условиях непродлевающегося случайного мёртвого времени, которая является инструментом для задач оценивания параметров исходного потока и параметра T^* равномерного распределения длительности мёртвого времени. Проведённые статистические эксперименты, а также другие многочисленные испытания имитационной модели подтверждают её работоспособность.

ЛИТЕРАТУРА

1. *Basharin G.P., Gaidamaka Y.V., Samouylov K.E.* Mathematical theory of teletraffic and its application to the analysis of multiservice communication of next generation networks // Automatic Control and Computer Sciences. – 2013. – 47(2). – P. 62–69.

2. *Дудин А.Н., Клименок В.Н.* Системы массового обслуживания с коррелированными потоками. – Минск: Изд-во БГУ, 2000. – 175 с.

3. *Апанасович В.В., Коляда А.А., Чернявский А.Ф.* Статистический анализ случайных потоков в физическом эксперименте.– Минск: Университетское, 1988. – 256 с.

4. *Горцев А.М., Нежельская Л.А.* Оценивание длительности мёртвого времени и параметров синхронного альтернирующего потока событий // Вестник Томского государственного университета. – 2003. – № S6. – С. 232–239.

5. *Горцев А.М., Ниссенбаум О.В.* Оценивание длительности мёртвого времени и параметров асинхронного альтернирующего потока событий с инициированием лишнего события // Вестник Томского государственного университета. – 2004. – № 284. – С. 137–145.

6. *Василевская Т.П., Горцев А.М., Нежельская Л.А.* Оценивание длительности мёртвого времени и параметров синхронного альтернирующего потока с проявлением либо непроявлением событий // Вестник Томского государственного университета. – 2004. – №S 9-2. – С. 129–138.

7. *Горцев А.М., Ниссенбаум О.В.* Оценивание длительности мёртвого времени и параметров асинхронного альтернирующего потока событий при непродлевающемся мёртвом времени // Известия высших учебных заведений. Физика. – 2005. – № 10. – С. 35–49.

8. *Горцев А.М., Леонова М.А., Нежельская Л.А.* Сравнение МП- и ММ-оценок длительности мёртвого времени в обобщенном асинхронном потоке событий // Вестник Томского государственного университета.. Управление, вычислительная техника и информатика.– 2013. – № 4 (25). – С. 32–42.

9. *Горцев А.М., Леонова М.А., Нежельская Л.А.* Совместная плотность вероятностей длительности интервалов обобщенного асинхронного потока событий при непродлевающемся мёртвом времени // Вестник Томского

государственного университета. Управление, вычислительная техника и информатика. – 2012. – № 4 (21). – С. 14–25.

10. *Горцев А.М., Соловьёв А.А.* Совместная плотность вероятностей длительности интервалов потока физических событий при непродлевающемся мёртвом времени // Известия высших учебных заведений. Физика. – 2014. – Т. 57. – № 7. – С. 103–111.

11. *Горцев А.М., Калягин А.А., Нежельская Л.А.* Оценка максимального правдоподобия длительности мёртвого времени в обобщённом полусинхронном потоке // Вестник Томского государственного университета. Управление, вычислительная техника и информатика. – 2015. – № 1. – С. 27–37.

12. *Горцев А.М., Климов И.С.* Оценивание периода ненаблюдаемости и интенсивности пуассоновского потока событий // Радиотехника. – 1996. – № 2. – С. 8–11.

13. *Горцев А.М., Паршина М.Е.* Оценивание параметров альтернирующего потока событий в условиях «мёртвого времени» // Известия высших учебных заведений. Физика. – 1999. – № 4. – С. 8–13.

14. *Горцев А.М., Нежельская Л.А.* Полусинхронный дважды стохастический поток событий при продлевающемся мёртвом времени // Вычислительные технологии. – 2008. – Т. 13. – № 1. – С. 31–41.

15. *Бушланов И.В., Горцев А.М., Нежельская Л.А.* Оценка параметров синхронного дважды стохастического потока событий // Автоматика и телемеханика. – 2008. – № 9. – С. 76–93.

16. *Горцев А.М., Завгородняя М.Е.* Оценивание параметра продлевающегося мёртвого времени случайной длительности в пуассоновском потоке событий // Вестник Томского государственного университета. Управление, вычислительная техника и информатика – 2017. – № 40. – С. 32–40.

17. *Глухова Е.В., Терпугов А.Ф.* Оценка интенсивности пуассоновского потока событий при наличии продлевающегося мёртвого времени // Известия высших учебных заведений. Физика. – 1995. – Т. 38. – № 3. – С. 22–31.

18. *Лифшиц А.Л., Мальц. Э.А.* Статистическое моделирование систем массового обслуживания. – М.: Сов. Радио, 1978. – 248 с.

19. *Шуленин В.П.* Математическая статистика. Ч.1. Параметрическая статистика. – Томск: Изд-во НТЛ, 2012.– 540 с.

СТАТИСТИЧЕСКИЕ ЭКСПЕРЕМЕНТЫ НА ИМИТАЦИОННОЙ МОДЕЛИ АСИНХРОННОГО ПОТОКА СОБЫТИЙ ПРИ НЕПРОДЛЕВАЮЩЕМСЯ СЛУЧАЙНОМ МЕРТВОМ ВРЕМЕНИ

А.М. Горцев, А.В. Шманкеева
Томский государственный университет

Введение

Распространенными математическими моделями физических явлений и процессов являются потоки событий. В частности, такие модели применяются при исследовании информационных потоков сообщений в телекоммуникационных системах, в спутниковых сетях связи и т.п. [1]. В связи с интенсивным развитием компьютерных сетей модель простейшего потока событий перестала быть адекватной реальным информационным потокам. Требования практики послужили стимулом к рассмотрению дважды стохастических потоков [2,3] в качестве математической модели реальных потоков событий в компьютерных сетях. В большинстве публикаций авторы рассматривают математические модели потоков событий, когда события потока доступны наблюдению. Однако на практике возникают ситуации, когда наступившее событие влечет за собой ненаблюдаемость последующих событий. Причиной ненаблюдаемости, как правило, выступает мертвое время регистрирующих приборов [4], в течение которого зарегистрированное событие обрабатывается, другие же события, поступившие в этот период, теряются. Регистрирующие приборы при этом делятся на два вида: с непродлевающемся мертвым временем и продлевающемся. При этом длительность мертвого времени может быть, как детерминированной величиной, так и случайной. Задачи по оценке параметров и состояний потока событий в условиях мёртвого времени фиксированной длительности рассматривались в работах [5–12]. При этом в [5–10] получены результаты для непродлевающегося мёртвого времени, в [11,12] – для продевающегося.

Однако достаточно открытым остается вопрос изучения потоков событий, когда мертвое время является случайной величиной с тем или иным законом распределения. Здесь отметим работу [13], в которой решается задача оценки параметра распределения непродлевающегося случайного мёртвого времени в пуассоновском потоке и работу [14], в которой находятся формулы для начальных моментов общего периода нена-

блюдательности в пуассоновском потоке событий при продлевающемся случайном мёртвом времени. В настоящей статье рассматривается асинхронный поток событий (MMPP-поток), функционирующий в условиях непродлевающегося случайного мертвого времени, приводятся результаты статистических экспериментов, реализованных на имитационной модели изучаемого потока.

1. Математическая модель наблюдаемого потока

Рассматривается асинхронный дважды стохастический поток событий (далее – поток), сопровождающий процесс (интенсивность) которого есть кусочно-постоянный стационарный случайный процесс $\lambda(t)$ с двумя состояниями S_1 и S_2. Будем говорить, что имеет место первое состояние процесса (потока) S_1, если $\lambda(t)=\lambda_1$, и, наоборот, имеет место второе состояние процесса (потока) S_2, если $\lambda(t)=\lambda_2$ ($\lambda_1>\lambda_2\geq 0$). Если имеет место S_i-е, $i=1,2$, состояние процесса $\lambda(t)$, то в течение временного интервала, когда $\lambda(t)=\lambda_i$, наступает пуассоновский поток событий с интенсивностью λ_i. Переход из состояния S_1 в состояние S_2 и, наоборот, может осуществляться в произвольный момент времени, не связанный с моментом наступления события. При этом длительность пребывания процесса $\lambda(t)$ в состоянии S_i, $i=1,2$, есть случайная величина, распределенная по экспоненциальному закону $F_i(t)=1-e^{-\alpha_i t}$, $t\geq 0$, где α_1 – интенсивность смены состояния S_1 на S_2, α_2 – интенсивность смены состояния S_2 на S_1. Т.к. переходы из состояния в состояния происходят в произвольный момент времени, т.е. не привязаны к моментам наступления событий, то поток называется асинхронным дважды стохастическим потоком событий. Матрицы инфинитезимальных коэффициентов процесса $\lambda(t)$ при этом примут вид

$$D_0=\begin{Vmatrix}-(\lambda_1+\alpha_1) & \alpha_1\\ \alpha_2 & -(\lambda_1+\alpha_2)\end{Vmatrix},\quad D_1=\begin{Vmatrix}\lambda_1 & 0\\ 0 & \lambda_2\end{Vmatrix}.$$

Элементами матрицы D_1 являются интенсивности переходов процесса $\lambda(t)$ из состояния в состояние с наступлением события. Недиагональные элементы матрицы D_0 – интенсивности переходов из состояния в состояние без наступления события. Диагональные элементы матрицы D_0 –интенсивности выхода процесса из своих состояний, взятые с противоположным знаком. В сделанных предположениях $\lambda(t)$ – скрытый марковский процесс ($\lambda(t)$ – принципиально ненаблюдаемый процесс).

После каждого зарегистрированного в момент t_k события наступает период мёртвого времени случайной длительности, который порождается этим событием, так что другие события исходного потока, наступившие в течение этого периода мёртвого (периода ненаблюдаемости) недоступны наблюдению (теряются) и не вызывают его продления (непродлевающееся мёртвое время). Принимается, что случайная длительность мёртвого времени распределена по равномерному закону с плотностью вероятности $p(T)=1/T^*$, где T – значение длительности мертвого времени, $0\leq T\leq T^*$.

На рис. 1 приведен вариант возникающей ситуации, где S_1, S_2 – состояния случайного процесса $\lambda(t)$; временная ось $(0,t)$ – ось наступления наблюдаемых событий в моменты времени t_1, t_2, …; временная ось $\left(0,t^{(1)}\right)$ – ось наступления наблюдаемых событий в моменты времени $t_1^{(1)}$, $t_2^{(1)}$, … в первом (S_1) состоянии процесса $\lambda(t)$, на которой также указаны значения длительностей $T_1^{(1)}$, $T_2^{(1)}$, … мертвых времён, порожденных наблюдаемыми событиями потока; аналогично для временной оси $\left(0,t^{(2)}\right)$; белыми кружками обозначены наблюдаемые события, черными – ненаблюдаемые, штриховкой

– значения периодов мертвого времени; траектория процесса $\lambda(t)$ привязана к осям $\left(0,t^{(1)}\right)$ и $\left(0,t^{(2)}\right)$.

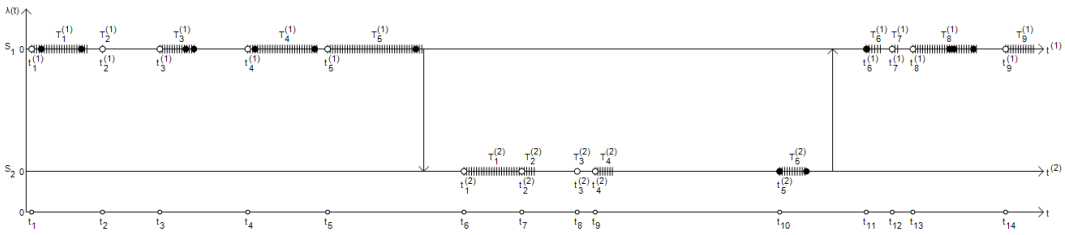


Рис.1. Асинхронный поток событий в условиях непродлевающегося случайного мертвого времени

Настоящая работа посвящена вопросу построения имитационной модели асинхронного потока событий, функционирующего в условиях непродлевающегося случайного мертвого времени, с целью проведения статистических экспериментов на этой модели для изучения содержательной информации о статистических характеристиках периода ненаблюдаемости и установления её работоспособности.

2. Имитационная модель асинхронного потока событий при непродлевающемся случайном мертвом времени

Метод имитационного моделирования широко применяется на этапе проектирования сложных систем. Основным средством реализации имитационного моделирования служит компьютер, позволяющий осуществлять цифровое моделирование систем и процессов. При имитационном моделировании используемая математическая модель воспроизводит алгоритм функционирования исследуемой системы во времени при различных сочетаниях значений параметров системы.

Основные преимущества имитационного моделирования по сравнению с аналитическими методами [15] заключаются в следующем: 1) бóльшая адекватность между сущностью физического явления (объекта) и математической моделью физического явления (объекта); 2) более широкий охват физических процессов и систем по сравнению с аналитическими методами; 3) имитационное моделирование функционирования процессов и систем при разнообразных законах распределения случайных величин как в установившихся, так и в переходных режимах; 4) получение в результате имитационного моделирования самых разнообразных статистических характеристик исследуемых процессов и систем.

На рис. 2 представлена блок-схема алгоритма имитационного моделирования асинхронного потока событий в условиях непродлевающегося случайного мертвого времени.

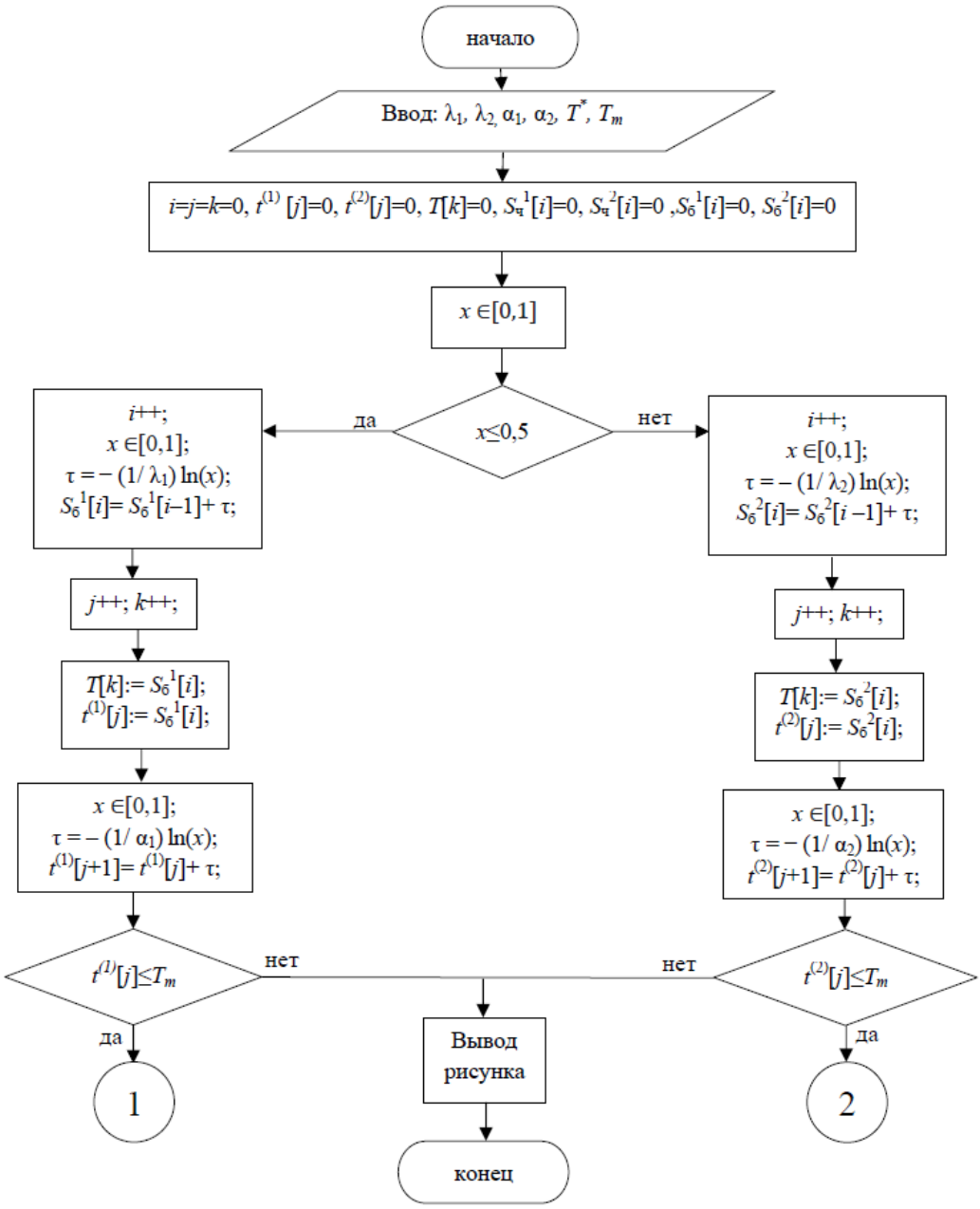


Рис. 2. Блок- схема алгоритма имитационного моделирования асинхронного потока событий в условиях непродлевающегося случайного мертвого времени (часть 1)

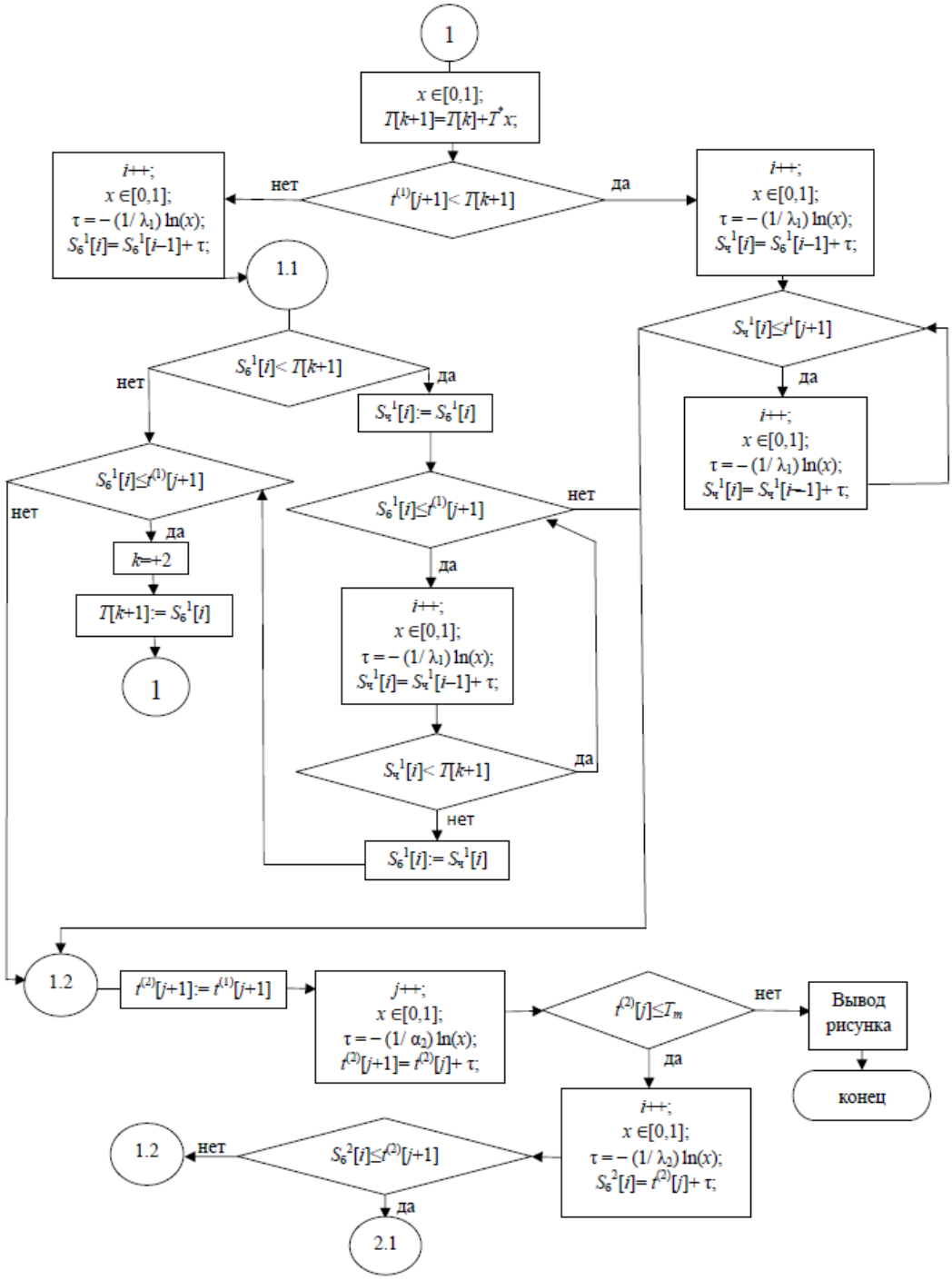


Рис. 2. Блок- схема алгоритма имитационного моделирования асинхронного потока событий в условиях непродлевающегося случайного мертвого времени (часть 2)

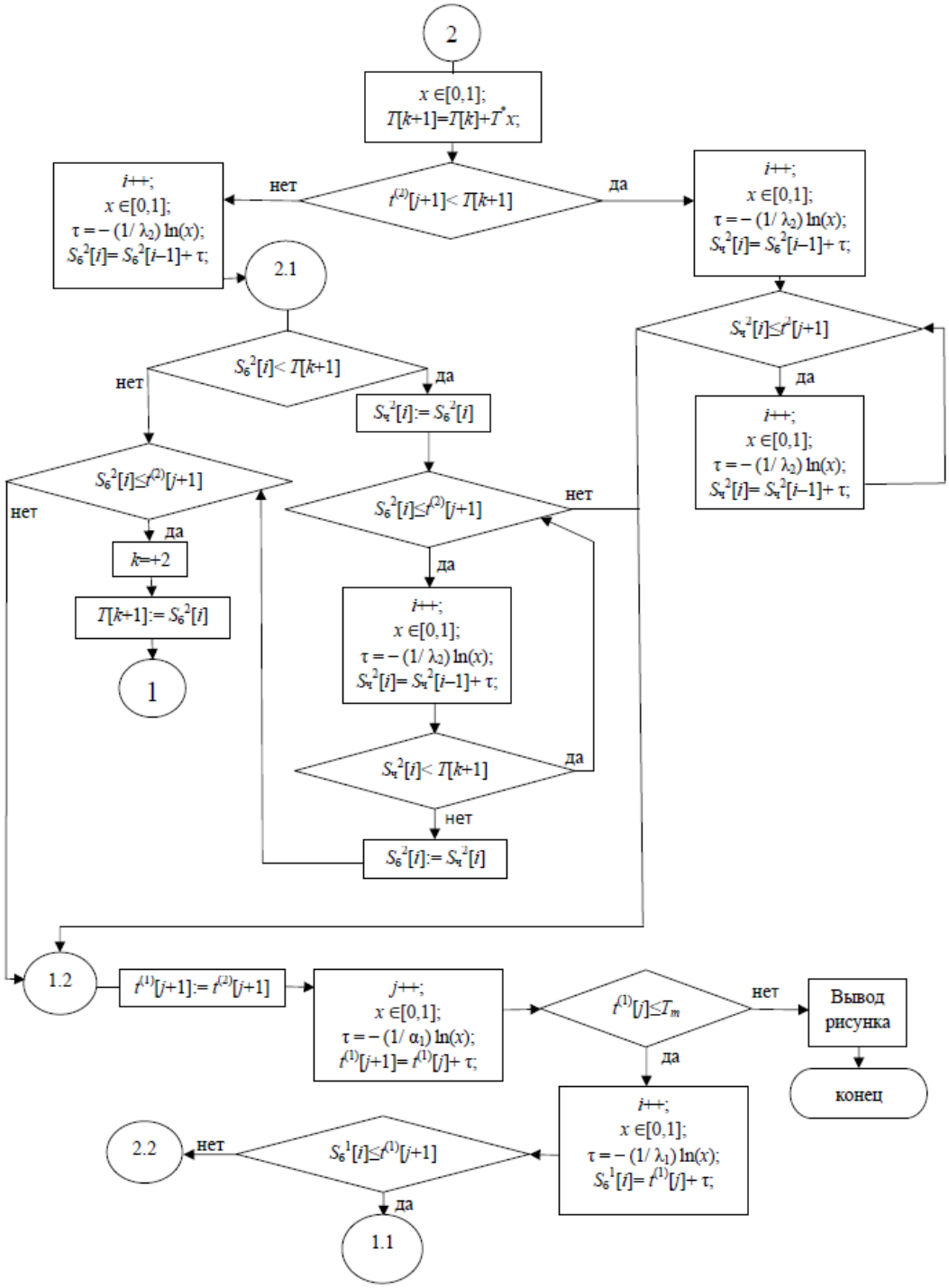


Рис. 2. Блок- схема алгоритма имитационного моделирования асинхронного потока событий в условиях непродлевающегося случайного мертвого времени (часть 3)

На рис. 2 введены следующие обозначения: T_m – время моделирования; T – значение длительности мертвого времени; x – равномерно распределённая на отрезке [0,1] случайная величина; τ – либо интервал времени между соседними событиями потока, либо интервал пребывания процесса $\lambda(t)$ в состоянии S_1 или в состоянии S_2; $t^{(1)}$ – теку-

щее время в состоянии S_1; $t^{(2)}$ – текущее время в состоянии S_2; $S_б^1$ и $S_б^2$ – наблюдаемые события; $S_ч^1$ и $S_ч^2$ – ненаблюдаемые события.

Для получения численных результатов разработана (в соответствии с приведенной на рис. 2 блок- схемой) программа, реализованная на языке программирования C# в среде Microsoft Visual Studio C#. Первый этап расчета предполагает имитационное моделирование асинхронного потока событий, схемы создания непродлевающегося случайного мертвого вмени и наблюдаемого потока событий.

Результатом работы имитационной модели является последовательность значений длительности периодов ненаблюдаемости событий асинхронного потока, а также последовательность значений длительности интервалов между соседними событиями наблюдаемого потока.

3. Статистические эксперименты и численные результаты

С целью установления работоспособности имитационной модели асинхронного потока, функционирующего в условиях непродлевающегося случайного мёртвого времени, поставлены следующие статистические эксперименты.

Первый статистический эксперимент: 1) При заданных значениях параметров λ_1, λ_2, α_1, α_2, T^* и заданном времени моделирования T_m осуществляется имитационное моделирование наблюдаемого потока; выходом имитационной модели является последовательность значений T_1, T_2, …, T_N периодов мёртвого времени, где T_i – значение периода ненаблюдаемости, порожденного i-ым наблюдаемым событием; значение N – зависит от заданных значений λ_1, λ_2, α_1, α_2 , T_m, T^* и конкретной реализации наблюдаемого потока; 2) в рамках конкретной реализации наблюдаемого потока вычисляется статистическая характеристика $\hat{M}(\boldsymbol{T}) = (1/N)\sum_{i=1}^{N} T_i$ – оценка математического ожидания периода ненаблюдаемости $M(\boldsymbol{T})$.

В табл. 1–3 приведены результаты первого статистического эксперимента. В первой строке табл. 1–3 указано время моделирования T_m (время наблюдения за потоком) (T_m = 50, 100, 150, …, 1000 ед. времени), во второй строке указана оценка математического ожидания $\hat{M}(\boldsymbol{T})$ случайной величины $\boldsymbol{T}$ – длительности мёртвого времени. Значения параметров λ_1, λ_2, α_1, α_2, T^*, при которых осуществляется имитационное моделирование, приведены в заголовках таблиц.

Таблица 1

Оценка математического ожидания периода ненаблюдаемости при $\lambda_1 = 2$, $\lambda_2 = 1$, $\alpha_1 = 0.1$, $\alpha_2 = 0.2$, $T^* = 1$

T_m	50	100	150	200	250	300	350	400	450	500
$\hat{M}(\boldsymbol{T})$	0.601	0.505	0.544	0.525	0.506	0.509	0.490	0.510	0.481	0.810
T_m	550	600	650	700	750	800	850	900	950	1000
$\hat{M}(\boldsymbol{T})$	0.497	0.480	0.500	0.516	0.512	0.488	0.501	0.503	0.499	0.494

Таблица 2

Оценка математического ожидания периода ненаблюдаемости при $\lambda_1 = 2$, $\lambda_2 = 1$, $\alpha_1 = 0.1$, $\alpha_2 = 0.2$, $T^* = 2$

T_m	50	100	150	200	250	300	350	400	450	500
$\hat{M}(\boldsymbol{T})$	1.070	1.050	1.010	1.110	0.950	0.990	0.930	1.030	1.014	1.016
T_m	550	600	650	700	750	800	850	900	950	1000
$\hat{M}(\boldsymbol{T})$	0.988	0.960	0.988	1.020	1.000	1.030	0.980	0.970	0.980	0.990

Таблица 3

Оценка математического ожидания периода ненаблюдаемости при $\lambda_1 = 2$, $\lambda_2 = 1$, $\alpha_1 = 0.1$, $\alpha_2 = 0.2$, $T^* = 3$

T_m	50	100	150	200	250	300	350	400	450	500
$\hat{M}(\boldsymbol{T})$	1.490	1.720	1.580	1.630	1.360	1.490	1.501	1.470	1.570	1.505
T_m	550	600	650	700	750	800	850	900	950	1000
$\hat{M}(\boldsymbol{T})$	1.479	1.550	1.440	1.464	1.508	1.480	1.490	1.470	1.511	1.512

Для наглядности на рис. 3 приведены графики зависимости $\hat{M}(\boldsymbol{T})$ от времени моделирования T_m, построенные по данным табл. 1–3.

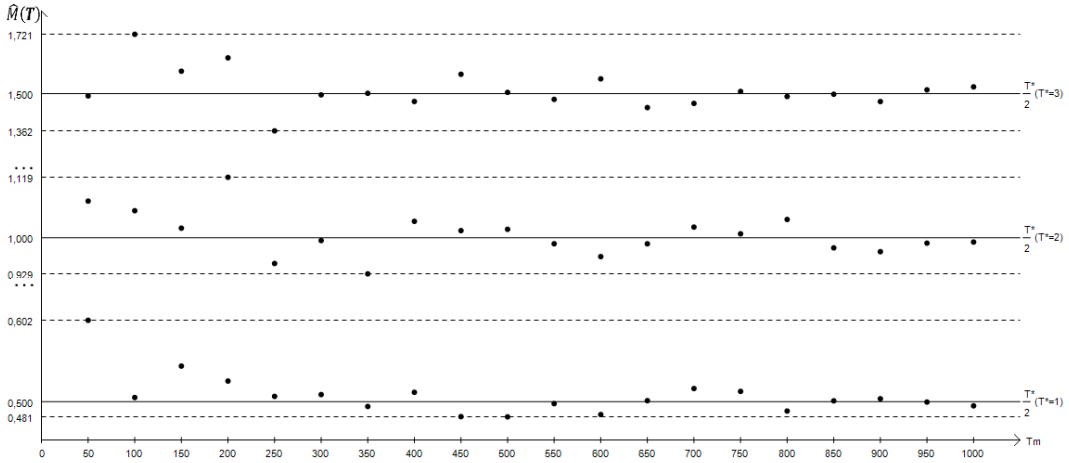


Рис. 3. График зависимости $\hat{M}(\boldsymbol{T})$ от значений T_m

Из анализа результатов табл. 1–3 и графиков, построенных по данным табл. 1–3, следует, что стационарный режим устанавливается при $T_m > 800$, т.е. оценка математического ожидания $\hat{M}(\boldsymbol{T})$ при увеличении T_m стремится к теоретическому значению $M(\boldsymbol{T}) = T^*/2$.

Второй статистический эксперимент: 1) При заданных значениях параметров λ_1, λ_2, α_1, α_2, T^* и заданном времени моделирования T_m осуществляется имитационное моделирование наблюдаемого потока; выходом имитационной модели является последовательность значений τ_i, $i = \overline{1, N}$; значение N зависит от заданных значений λ_1, λ_2, α_1, α_2, T^*, T_m и конкретной реализации наблюдаемого потока; 2) в рамках конкретной реализации наблюдаемого потока вычисляется статистическая характеристика $\hat{M}(\boldsymbol{\tau}) = (1/N)\sum_{i=1}^{N}\tau_i$ – оценка математического ожидания длительности интервала между соседними наблюдаемым событиями наблюдаемого потока, где $\tau_i = t_{i+1} - t_i$.

В табл. 4–6 приведены результаты второго статистического эксперимента. Структура табл. 4–6 аналогична структуре табл. 1–3.

Таблица 4

Оценка математического ожидания периода между наблюдаемыми событиями при $\lambda_1 = 2$, $\lambda_2 = 1$, $\alpha_1 = 0.1$, $\alpha_2 = 0.2$, $T^* = 1$

T_m	50	100	150	200	250	300	350	400	450	500
$\hat{M}(\tau)$	0.930	1.090	1.160	1.130	1.070	1.110	1.080	1.140	1.080	1.124
T_m	550	600	650	700	750	800	850	900	950	1000
$\hat{M}(\tau)$	1.510	1.154	1.164	1.134	1.202	1.102	1.103	1.110	1.101	1.070

Таблица 5

Оценка математического ожидания периода между наблюдаемыми событиями при $\lambda_1 = 2$, $\lambda_2 = 1$, $\alpha_1 = 0.1$, $\alpha_2 = 0.2$, $T^* = 2$

T_m	50	100	150	200	250	300	350	400	450	500
$\hat{M}(\tau)$	1.340	1.659	1.650	1.540	1.593	1.599	1.657	1.643	1.663	1.639
T_m	550	600	650	700	750	800	850	900	950	1000
$\hat{M}(\tau)$	1.594	1.600	1.560	0.516	1.600	1.580	1.560	1.580	1.640	1.610

Таблица 6

Оценка математического ожидания периода между наблюдаемыми событиями при $\lambda_1 = 2$, $\lambda_2 = 1$, $\alpha_1 = 0.1$, $\alpha_2 = 0.2$, $T^* = 3$

T_m	50	100	150	200	250	300	350	400	450	500
$\hat{M}(\tau)$	2.070	2.100	2.070	2.120	1.950	2.110	2.160	2.017	2.096	1.910
T_m	550	600	650	700	750	800	850	900	950	1000
$\hat{M}(\tau)$	2.110	2.186	2.121	2.128	2.130	2.210	2.202	2.090	2.070	2.130

Для наглядности на рис. 4 приведены графики зависимости $\hat{M}(\tau)$ от времени моделирования T_m, построенные по данным табл. 4–6.

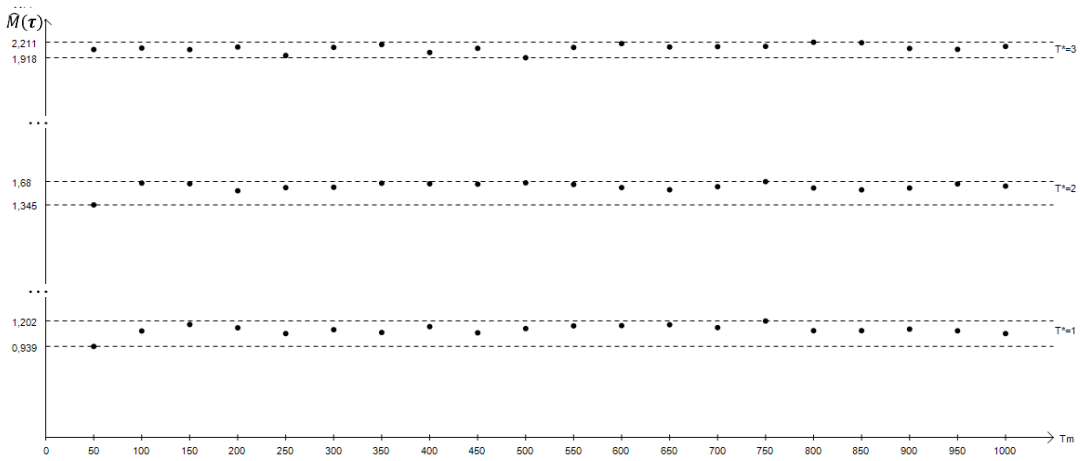


Рис. 4. График зависимости $\hat{M}(\tau)$ от значений T_m

Из анализа результатов табл. 4–6 и графиков, построенных по данным табл. 4–6, следует, что стационарный режим устанавливается при $T_m > 800$, т.е. оценка математического ожидания $\hat{M}(\tau)$ стремится к некоторому постоянному значению.

Заключение

В настоящей статье разработана имитационная модель асинхронного дважды стохастического потока событий, функционирующего в условиях непродлевающегося случайного мертвого времени, которая является инструментом для задач оценивания

параметров исходного асинхронного потока λ_1, λ_2, α_1, α_2 и параметра равномерного распределения длительности мертвого времени T^*. Проведенные статистические эксперименты, а также другие многочисленные испытания имитационной модели подтверждают ее работоспособность.

ЛИТЕРАТУРА

1. *Дудин А.Н., Клименок В.И.* Системы массового обслуживания с коррелированными потоками. – Минск: Изд-во БГУ, 2000. – 175 с.

2. *Башарин Г.П., Кокотушкин В.А., Наумов В.А.* О методе эквивалентных замен расчёта фрагментов сетей связи. Ч. 1 // Изв. АН СССР. Техн. кибернетика. – 1979. – № 6. – С. 92–99.

3. *Neuts M.F.* A versatile Markovian point process // Journal of Applied Probability. – 1979. – 16. – P. 764–779.

4. *Апанасович В.В., Коляда А.А., Чернявский А.Ф.* Статистический анализ случайных потоков в физическом эксперементе. – Минск: Университетское, 1988. – 256 с.

5. *Горцев А.М., Калягин А.А., Нежельская Л.А.* Оценка максимального правдоподобия длительности мертвого времени в обобщенном полусинхронном потоке // Вестник Томского государственного университета. Управление, вычислительная техника и информатика. – 2015. – № 1(30). – С. 27–37.

6. *Горцев А.М., Леонова М.А., Нежельская Л.А.* Совместная плотность вероятностей длительности интервалов обобщенного асинхронного потока событий при непродлевающемся мёртвом времени // Вестник Томского государственного университета. Управление, вычислительная техника и информатика. – 2012. – № 4 (21). – С. 14–25.

7. *Горцев А.М., Соловьёв А.А.* Совместная плотность вероятностей длительности интервалов потока физических событий при непродлевающемся мёртвом времени // Известия высших учебных заведений. Физика. – 2014. – Т. 57. – № 7. – С. 103–111.

9. *Горцев А.М., Василевская Т.П., Нежельская Л.А.* Оценивание длительности мёртвого времени и параметров синхронного альтернирующего потока с проявлением либо непроявлением событий // Вестник Томского государственного университета. – 2004. – № S9-2. – С. 129–138.

10. *Nezhel'skaya L.A.* Optimal state estimation in modulated MAP event flows with unextendable dead time // Communications in Computer and Information Sciences. – 2014. – 487. – P. 342–350.

11. *Горцев А.М., Нежельская Л.А.* Полусинхронный дважды стохастический поток событий при продлевающемся мёртвом времени // Вычислительные технологии. – 2008. – Т. 13. – № 1. – С. 31–41.

12. *Горцев А.М., Климов И.С.* Оценивание периода ненаблюдаемости и интенсивности пуассоновского потока событий // Радиотехника. – 1996. – № 2. – С. 8–11.

13. *Горцев А.М., Завгородняя М.Е.* Оценивание параметра непродлевающегося мертвого времени случайной длительности в пуассоновском потоке событий // Вестник Томского государственного университета. Управление, вычислительная техника и информатика. – 2017. – № 40. – С. 32–40.

14. *Глухова Е.В., Терпугов А.Ф.* Оценка интенсивности пуассоновского потока событий при наличии продлевающегося мертвого времени // Известия высших учебных заведений. Физика. – 1995. – Т. 38. – № 3. – С. 22–31.

15. *Лифшиц А.Л., Мальц Э.А.* Статистическое моделирование систем массового обслуживания. – М.: Советское радио, 1978. – 248 с.

ОЦЕНИВАНИЕ ПАРАМЕТРОВ ПРОСТЕЙШЕГО ПОТОКА СОБЫТИЙ И МЕРТВОГО ВРЕМЕНИ, РАСПРЕДЕЛЕННОГО ПО ЗАКОНУ РЭЛЕЯ, МЕТОДОМ МОМЕНТОВ

М.Е. Завгородняя, Ю.Е. Табольжина

Томский государственный университет

Введение

Потоки однородных событий являются распространенными математическими моделями многих реальных физических процессов. Часто на практике возникают ситуации, когда наступившее событие может повлечь за собой ненаблюдаемость последующих событий. Одним из искажающих факторов при оценке параметров потока событий выступает мертвое время регистрирующих приборов, которое порождается зарегистрированным событием. Другие же события, наступившие в течение периода мертвого времени, недоступны наблюдению. Все устройства регистрации делятся на две группы. Первую группу составляют устройства с непродлевающимся мертвым временем, вторую – устройства с продлевающимся мертвым временем. Отметим, что подобные задачи по оценке параметров случайных потоков событий, функционирующих в условиях непродлевающегося мертвого времени, рассматриваются в работах [1–10].

В настоящей статье исследуется простейший поток событий, функционирующий в условиях непродлевающегося мертвого времени, имеющего распределение Рэлея.

1. Постановка задачи

Рассматривается простейший поток событий интенсивности λ. После каждого события, зарегистрированного в момент t_k, наступает период мертвого времени, в течении которого остальные события недоступны наблюдению. По окончании периода мертвого времени первое наступившее событие снова создает период мертвого времени случайной длительности и т.д. Длительность мертвого времени распределена по закону распределения Рэлея с параметром σ, т.е. плотность распределения вероятностей имеет вид $p(t) = \frac{t}{\sigma^2} e^{-\frac{t^2}{2\sigma^2}}$. Вариант возникновения возможной ситуации приведен на рис. 1, где прямоугольниками обозначены периоды мертвого времени длительности T_1, T_2; ,...; t_1, t_2,... – моменты наступления событий в наблюдаемом потоке.

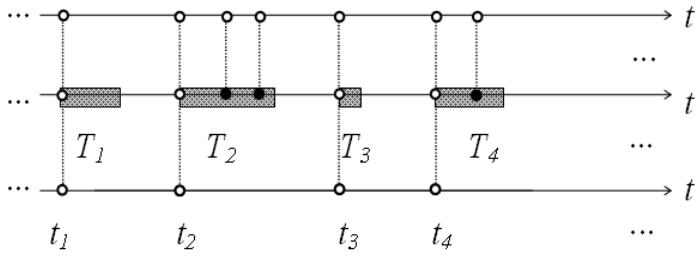


Рис. 1.Формирование наблюдаемого потока событий

Рассматривается установившийся (стационарный) режим функционирования наблюдаемого потока событий, поэтому переходными процессами на полуинтервале наблюдения $(t_0, t]$, где t_0 – начало наблюдений, t – окончание наблюдений, пренебрегаем. Тогда без потери общности можно положить $t_0 = 0$.

Необходимо в момент окончания наблюдений (в момент времени t) на основании выборки $t_1,..., t_k$ наблюденных моментов наступления событий потока осуществить методом моментов построение оценок $\hat{\lambda}$ и $\hat{\sigma}$.

2. ММ – оценки параметров λ и σ

Обозначим через $\tau_k = t_{k+1} - t_k$, $k = 1,2,...$ значение длительности k-го интервала между соседними событиями наблюдаемого потока ($\tau_k \geq 0$). Т.к. рассматривается стационарный режим, то плотность вероятности значений длительности k-го интервала есть $\rho_{\tau_k}(x) = \rho_\tau(x)$, $x \geq 0$, для любого k в наблюдаемом потоке событий. В силу этого момент времени t_k без потери общности можно положить равным нулю, т.е. момент наступления события есть $\tau = 0$. В [1] получено выражение для плотности вероятности $\rho(x \mid t)$, когда длительность мертвого времени является детерминированной величиной:

$$\rho(x \mid t) = 0,\ x < t;\ \rho(x \mid t) = \lambda e^{-\lambda(x-t)},\ x \geq t. \qquad (1)$$

Тогда плотность вероятности $\rho_\tau(x)$ для наблюдаемого потока выпишется в виде

$$p_\tau(x) = \int_0^x p(t)\rho(x \mid t)dt. \qquad (2)$$

Подставляя в (2) выражение (1) и учитывая, что $p(t)=\frac{t}{\sigma^2}e^{-\frac{t^2}{2\sigma^2}}$, запишем (2) в следующем виде:

$$p_\tau(x)=\int_0^x \frac{t}{\sigma^2}e^{-\frac{t^2}{2\sigma^2}}\lambda e^{-\lambda(x-t)}dt\,. \tag{3}$$

Методом моментов получим оценки параметров λ и σ. Суть метода моментов заключается в приравнивании теоретических моментов распределения к соответственным эмпирическим моментам, найденных по выборке:

$$\hat{M}\left\{\tau^k\right\}=\frac{1}{n}\sum_{i=1}^{n}\tau_i^k\,. \tag{4}$$

Т.к. необходимо найти оценки двух параметров, соотношение (4) рассмотрим для случая $k=1$, $k=2$.

Для случая, когда $k=1$ формула (4) принимает вид

$$\hat{M}\{\tau\}=\frac{1}{n}\sum_{i=1}^{n}\tau_i\,. \tag{5}$$

Теоретическое среднее значение длительности интервала между соседними моментами наступления событий наблюдаемого потока (начальный момент первого порядка) определяется по формуле:

$$M\{\tau\}=\int_0^\infty xp_\tau(x)dx\,. \tag{6}$$

Подставляя в (6) выражение (3), получим

$$M\{\tau\}=\int_0^\infty\int_0^x x\frac{t}{\sigma^2}e^{-\frac{t^2}{2\sigma^2}}\lambda e^{-\lambda(x-t)}dtdx\,. \tag{7}$$

Поменяем порядок интегрирования в (7) и перепишем его в виде

$$M\{\tau\}=\int_0^\infty \frac{t}{\sigma^2}e^{-\frac{t^2}{2\sigma^2}}e^{\lambda t}dt\int_t^\infty x\lambda e^{-\lambda x}dx\,. \tag{8}$$

Отдельно рассмотрим интеграл $\int_t^\infty x\lambda e^{-\lambda x}dx$:

$$\int_t^\infty x\lambda e^{-\lambda x}dx=te^{-\lambda t}+\frac{1}{\lambda}e^{-\lambda t}\,. \tag{9}$$

С учетом (9) перепишем выражение (8) в виде

$$M\{\tau\}=\int_0^\infty \frac{t}{\sigma^2}e^{-\frac{t^2}{2\sigma^2}}e^{\lambda t}e^{-\lambda t}\left(t+\frac{1}{\lambda}\right)dt\,. \tag{10}$$

В выражении (10) сделаем замену $\frac{t}{\sigma}=z$ и запишем его в виде

$$M\{\tau\}=\int_0^\infty ze^{-\frac{z^2}{2}}\left(z\sigma+\frac{1}{\lambda}\right)dz=I_1+I_2\,. \tag{11}$$

Учитывая, что $I_1=\sigma\int_0^\infty z^2e^{-\frac{z^2}{2}}dz=\sigma\frac{\sqrt{2\pi}}{2}$, $I_2=\frac{1}{\lambda}\int_0^\infty ze^{-\frac{z^2}{2}}dz=\frac{1}{\lambda}$, выражение (11) примет вид $M\{\tau\}=\sigma\frac{\sqrt{2\pi}}{2}+\frac{1}{\lambda}$.

Перепишем формулу (5) для уравнения моментов первого порядка следующим образом:

$$\sigma\frac{\sqrt{2\pi}}{2}+\frac{1}{\lambda}=\tau_{\text{ср}},\ \tau_{\text{ср}}=\frac{1}{n}\sum_{i=1}^{n}\tau_i\ . \tag{12}$$

Для случая, когда $k=2$, формула (4) принимает вид

$$\hat{M}\left\{\tau^2\right\}=\frac{1}{n}\sum_{i=1}^{n}\tau_i^2\ . \tag{13}$$

Начальный момент второго порядка определяется по формуле

$$M\left\{\tau^2\right\}=\int_0^{\infty}x^2p_{\tau}(x)dx\ . \tag{14}$$

Подставляя в (14) выражение (3), получим $M\{\tau^2\}=\int_0^{\infty}\int_0^{x}x^2\frac{t}{\sigma^2}e^{-\frac{t^2}{2\sigma^2}}\lambda e^{-\lambda(x-t)}dtdx$.

Поменяем порядок интегрирования в (8) и перепишем его в виде

$$M\{\tau^2\}=\int_0^{\infty}\frac{t}{\sigma^2}e^{-\frac{t^2}{2\sigma^2}}e^{\lambda t}dt\int_t^{\infty}x^2\lambda e^{-\lambda x}dx\ . \tag{15}$$

Отдельно рассмотрим интеграл $\int_t^{\infty}x^2\lambda e^{-\lambda x}dx$:

$$\int_t^{\infty}x\lambda e^{-\lambda x}dx=te^{-\lambda t}+\frac{2}{\lambda}e^{-\lambda t}\left(t+\frac{1}{\lambda}\right). \tag{16}$$

С учетом (16) перепишем выражение (15) в виде

$$M\left\{\tau^2\right\}=\int_0^{\infty}\frac{t}{\sigma^2}e^{-\frac{t^2}{2\sigma^2}}e^{\lambda t}e^{-\lambda t}\left(t^2+\frac{2}{\lambda}\left(t+\frac{1}{\lambda}\right)\right)dt\ . \tag{17}$$

В выражении (17) сделаем замену $\frac{t}{\sigma}=z$ и запишем его в виде

$$M\left\{\tau^2\right\}=\int_0^{\infty}ze^{-\frac{z^2}{2}}\left(z^2\sigma^2+\frac{2}{\lambda}z\sigma+\frac{2}{\lambda^2}\right)dz=I_1+I_2+I_3\ . \tag{18}$$

С учетом того, что $I_1=\sigma^2\int_0^{\infty}z^3e^{-\frac{z^2}{2}}dz=2\sigma^2$, $I_2=\frac{2\sigma}{\lambda}\int_0^{\infty}ze^{-\frac{z^2}{2}}dz=\frac{\sigma\sqrt{2\pi}}{\lambda}$, $I_3=\frac{2}{\lambda^2}\int_0^{\infty}ze^{-\frac{z^2}{2}}dz=\frac{2}{\lambda^2}$, выражение (18) примет вид $M\left\{\tau^2\right\}=2\sigma^2+\sigma\frac{\sqrt{2\pi}}{\lambda}+\frac{2}{\lambda^2}$.

Перепишем формулу (13) для уравнения моментов второго порядка следующим образом:

$$2\sigma^2+\sigma\frac{\sqrt{2\pi}}{\lambda}+\frac{2}{\lambda^2}=\tau_{\text{ср}}{}^2,\ \tau_{\text{ср}}^2=\frac{1}{n}\sum_{i=1}^{n}\tau_i{}^2\ . \tag{19}$$

Составим систему из уравнений (12) и (19):

$$\begin{cases}\sigma\frac{\sqrt{2\pi}}{2}+\frac{1}{\lambda}=\tau_{\text{ср}}=C_1,\\ 2\sigma^2+\sigma\frac{\sqrt{2\pi}}{\lambda}+\frac{2}{\lambda^2}=\tau_{\text{ср}}^2=C_2.\end{cases} \tag{20}$$

Решая систему уравнений (20), получим две пары значений $\left(\hat{\lambda}_1,\hat{\sigma}_1\right)$, $\left(\hat{\lambda}_2,\hat{\sigma}_2\right)$:

$$\hat{\lambda}_1 = \frac{2}{2C_1 - \sigma_1\sqrt{2\pi}},$$
$$\hat{\sigma}_1 = \frac{C_1\sqrt{2\pi} + \sqrt{2C_1^2\pi - 8\left(2C_1^2 - C_2\right)}}{4},$$
$$\hat{\lambda}_2 = \frac{2}{2C_1 - \sigma_2\sqrt{2\pi}}, \qquad (21)$$
$$\hat{\sigma}_2 = \frac{C_1\sqrt{2\pi} - \sqrt{2C_1^2\pi - 8\left(2C_1^2 - C_2\right)}}{4}.$$

3. Результаты численных расчетов

Для получения численных результатов разработан алгоритм оценивания параметров простейшего потока событий, функционирующего в условиях мертвого времени, имеющего распределение Рэлея. Программа реализована в среде Microsoft Visual Studio 2015 на языке программирования C#. Первый этап расчета заключается в нахождении численных значений ММ-оценок рассматриваемого потока. Второй этап расчета предполагает имитационное моделирование исходного простейшего потока событий, схемы создания мертвого времени и наблюдаемого потока событий. Результатом работы имитационной модели является получение оценок параметров простейшего потока событий и мертвого времени, распределенного по закону Рэлея.

С целью установления свойств получаемых оценок проведены статистические эксперименты. Результаты приведены в табл. 1–3. В первой строке таблиц указано время моделирования T и соответствующие величины $\tau_{ср}$, $\tau_{ср}^2$ (начальные моменты первого и второго порядка, соответственно). Во второй строке табл. 1–3 для соответствующих T и $\tau_{ср}$, $\tau_{ср}^2$ приведены численные значения ММ-оценок.

Таблица 1

$T = 1000$, $\tau_{ср} = 4.19$, $\tau_{ср}^2 = 21.59$	$T = 5000$, $\tau_{ср} = 4.15$, $\tau_{ср}^2 = 21.12$	$T = 10000$, $\tau_{ср} = 4.27$, $\tau_{ср}^2 = 22.37$	$T = 15000$, $\tau_{ср} = 4.15$, $\tau_{ср}^2 = 21.12$
$\hat{\lambda}_1 = 0.73$	$\hat{\lambda}_1 = 0.81$	$\hat{\lambda}_1 = 0.73$	$\hat{\lambda}_1 = 0.74$
$\hat{\lambda}_2 = 2.38$	$\hat{\lambda}_2 = 1.83$	$\hat{\lambda}_2 = 2.18$	$\hat{\lambda}_2 = 2.06$
$\hat{\sigma}_1 = 2.24$	$\hat{\sigma}_1 = 2.33$	$\hat{\sigma}_1 = 2.31$	$\hat{\sigma}_1 = 2.34$
$\hat{\sigma}_2 = 3.01$	$\hat{\sigma}_2 = 2.88$	$\hat{\sigma}_2 = 3.04$	$\hat{\sigma}_2 = 3.04$

В табл. 1 приведены значения оценок $\hat{\lambda}$, $\hat{\sigma}$ при заданных параметрах $\lambda = 2$, $\sigma = 3$.

Таблица 2

$T = 1000$, $\tau_{ср} = 5.74$, $\tau_{ср}^2 = 40.36$	$T = 5000$, $\tau_{ср} = 5.86$, $\tau_{ср}^2 = 42.19$	$T = 10000$, $\tau_{ср} = 6.15$, $\tau_{ср}^2 = 45.99$	$T = 15000$, $\tau_{ср} = 5.98$, $\tau_{ср}^2 = 43.87$
$\hat{\lambda}_1 = 0.59$	$\hat{\lambda}_1 = 0.53$	$\hat{\lambda}_1 = 0.65$	$\hat{\lambda}_1 = 0.54$
$\hat{\lambda}_2 = 1.29$	$\hat{\lambda}_2 = 1.58$	$\hat{\lambda}_2 = 0.91$	$\hat{\lambda}_2 = 1.42$
$\hat{\sigma}_1 = 3.23$	$\hat{\sigma}_1 = 3.17$	$\hat{\sigma}_1 = 3.67$	$\hat{\sigma}_1 = 3.29$
$\hat{\sigma}_2 = 3.97$	$\hat{\sigma}_2 = 4.17$	$\hat{\sigma}_2 = 4.03$	$\hat{\sigma}_2 = 4.21$

В табл. 2 приведены значения оценок $\hat{\lambda}$, $\hat{\sigma}$ при заданных параметрах $\lambda = 1$, $\sigma = 4$.

Таблица 3

$T=1000$, $\tau_{ср}=3.01$, $\tau_{ср}^2=11.19$	$T=5000$, $\tau_{ср}=3.09$, $\tau_{ср}^2=11.67$	$T=10000$, $\tau_{ср}=4.27$, $\tau_{ср}^2=22.37$	$T=15000$, $\tau_{ср}=3.02$, $\tau_{ср}^2=11.10$
$\hat{\lambda}_1=0.97$ $\hat{\lambda}_2=3.80$ $\hat{\sigma}_1=1.58$ $\hat{\sigma}_2=2.19$	$\hat{\lambda}_1=1.04$ $\hat{\lambda}_2=2.75$ $\hat{\sigma}_1=1.70$ $\hat{\sigma}_2=2.17$	$\hat{\lambda}_1=0.73$ $\hat{\lambda}_2=2.18$ $\hat{\sigma}_1=2.31$ $\hat{\sigma}_2=3.04$	$\hat{\lambda}_1=1.30$ $\hat{\lambda}_2=1.91$ $\hat{\sigma}_1=1.79$ $\hat{\sigma}_2=1.99$

В табл. 3 приведены значения оценок $\hat{\lambda}$, $\hat{\sigma}$ при заданных параметрах $\lambda=2$, $\sigma=2$.

По результатам статистического эксперимента можно сделать вывод о том, что при увеличении времени моделирования T оценки $\hat{\lambda}$, $\hat{\sigma}$ улучшаются, становясь все ближе к истинным значениям.

Отметим, что пара значений $\left(\hat{\lambda}_2, \hat{\sigma}_2\right)$ при достаточно большом времени моделирования, оказывается ближе к истинным значениям параметров λ, σ. Но для того, чтобы отсечь одну из пар решений (21) системы (20), необходимо использовать дополнительную информацию.

Заключение

В настоящей статье был исследован простейший поток событий, функционирующий в условиях непродлевающегося мертвого времени, имеющего распределение Рэлея. Разработан и реализован алгоритм оценивания по методу моментов параметров рассматриваемого потока и мертвого времени. Программа реализована в среде Microsoft Visual Studio 2015 на языке программирования C#. Проведены статистические эксперименты с целью установления свойств оценок. Полученные результаты показывают возможность оценивания параметров λ и σ по результатам текущих наблюдений за потоком событий методом моментов.

ЛИТЕРАТУРА

1. *Горцев А.М., Климов И.С.* Оценка интенсивности пуассоновского потока событий в условиях частичной его ненаблюдаемости // Радиотехника. – 1991. – № 12. – С. 3–7.

2. *Горцев А.М., Завгородняя М.Е.* Оценка параметров альтернирующего потока событий при условии его частичной наблюдаемости // Оптика атмосферы и океана. – 1997. – Т. 10. – № 3. – С. 273–280.

3. *Горцев А.М., Куснатдинов Р.Т.* Оценивание состояний МС-потока событий при его частичной наблюдаемости // Известия высших учебных заведений. Физика. – 1998. – № 4. – С. 22–30.

4. *Горцев А.М., Паршина М.Е.* Оценивание параметров альтернирующего потока событий в условиях «мертвого времени» // Известия высших учебных заведений. Физика. – 1999. – № 4. – С. 8–13.

5. *Vasileva L.A., Gortsev A.M.* Estimation of parameters of a double-stochastic flow of events under conditions of its incomplete observability // Automation and Remote Control. – 2002. – V. 63. – No. 3. – P. 511–515.

6. *Vasileva L.A., Gortsev A.M.* Estimation of the dead time of an asynchronous double stochastic flow of events under incomplete observability // Automation and Remote Control. – 2003. – V. 63. – No. 12. – P. 1890–1898.

7. *Горцев А.М., Нежельская Л.А.* Оценивание длительности мертвого времени и параметров синхронного альтернирующего потока событий // Вестник Томского государственного университета. – 2003. – № 6. – С. 232–239.

8. *Горцев А.М., Нежельская Л.А.* Оценивание длительности «мертвого времени» и интенсивностей синхронного дважды стохастического потока событий // Радиотехника. – 2004. – № 10. – С. 8–16.

9. *Горцев А.М., Ниссенбаум О.В.* Оценивание длительности мертвого времени и параметров асинхронного альтернирующего потока событий при непродлевающемся мертвом времени // Известия высших учебных заведений. Физика. – 2005. – Т. 10. – № 3. С. 35–49.

10. *Горцев А.М., Завгородняя М.Е.* Оценивание параметра непродлевающегося мертвого времени случайной длительности в пуассоновском потоке событий // Вестник Томского государственного университета. – 2017. – № 40. – С. 1–8.

МОДЕЛИРОВАНИЕ СЛУЧАЙНЫХ ВЕЛИЧИН МЕТОДОМ ОБРАТНЫХ ФУНКЦИЙ

А.В. Холкина
Томский государственный университет

Введение

Управление в современном мире становится всё более трудным делом. Это объясняется усложнением самих объектов и, следовательно, усложнением их моделей. Система, объединяющая все объекты и связи между ними, является сложной системой. Изменение одной из характеристик такой системы может привести к изменениям или создать потребность в изменениях в других частях системы. В связи с этим появилось имитационное моделирование.

Имитационное моделирование есть процесс конструирования модели реальной системы и постановка экспериментов на этой модели с целью либо понять поведение системы, либо оценить различные стратегии, обеспечивающие функционирование данной системы. Имитационное моделирование – наиболее мощный и универсальный метод исследования и оценки эффективности систем, поведение которых зависит от воздействия случайных факторов. В основе имитационного моделирования лежит статистический эксперимент, реализация которого невозможна без применения ЭВМ.

Имитационное моделирование имеет следующие цели:

1. Описание поведения систем;
2. Построение теорий и гипотез, которые могут объяснить наблюдаемое поведение;
3. Прогнозирование поведения систем, т.е. тех воздействий, которые могут быть вызваны изменениями в самой системе или ее способов функционирования.

1. Постановка задачи

Пусть x имеет заданную функцию распределения $F_\xi(x)$, где

1) $F_\xi(x)=\dfrac{x-a}{b-a}$, $a \le x \le b$,

2) $F_\xi(x)=1-e^{-\lambda x}$, $x \ge 0$, $\lambda > 0$,

3) $F_\xi(x)=1-e^{-\frac{x^2}{2\sigma^2}}$, $x \ge 0$, $\sigma > 0$,

4) $F_\xi(x)=1-e^{-\left(\frac{x}{\lambda}\right)^k}$, $x \ge 0$, $k > 0$, $\lambda > 0$,

5) $F_\xi(x)=\dfrac{1}{\sigma\sqrt{2\pi}}\int\limits_{-\infty}^{\infty} e^{-\frac{(x-\mu)^2}{2\sigma^2}}\,dx$, $-\infty < x < \infty$, $\sigma > 0$.

Требуется построить формулу моделирования случайной величины ξ на ЭВМ методом обратных функций и оценить адекватность моделирования с помощью критерия χ^2.

Как правило, в качестве базовой случайно величины (БСВ) выбирают случайную величину γ, которая распределена по равномерному закону на интервале (0,1). Эта случайная величина имеет функцию распределения $F_\gamma(x)=x$, $0<x<1$, и плотность распределения, $p(x)=1$, её математическое ожидание $M_\gamma = 1/2$, а дисперсия $D_\gamma = 1/12$. Величину γ будем далее называть случайным числом.

2. Метод обратных функций

Рассмотрим преобразование, которое позволит нам с помощью случайных чисел γ получить значение любой случайно величины ξ. Такие вычисления называют моделированием случайной величины ξ.

Основным приемом моделирования случайных величин является метод обратных функций, который заключается в том, что на основании условия существования обратной функции и приведённой ниже теоремы, мы получаем конкретные формулы моделирования случайно величины ξ.

Теорема. Случайная величина $\xi \sim F^{-1}(\gamma)$ имеет функцию распределения $F(x)$, где γ – БСВ, $F^{-1}(x)$ – функция, обратная к $F(x)$.

Т.к. случайная величина имеет то же распределение, что и γ, то иногда (когда это удобно) можно вместо случайной величины γ можно использовать величину $1-\gamma$.

Алгоритм построения формулы моделирования случайной величины x на основе теоремы выглядит следующим образом: 1. Генерируем γ – БСВ. 2. Составляем уравнение $F(\gamma)=\xi$. 3. Находим $\xi = F^{-1}(\gamma)$.

3. Метод обратных функций для конкретных распределений

Подробно пример построения формулы моделирования случайной величины ξ рассмотрим на примере экспоненциального распределения. Случайная величина имеет экспоненциальное распределение с параметром $\lambda > 0$, если её плотность распределения имеет вид $p(x) = \lambda e^{-\lambda x}$, $x \geq 0$.

Шаг 1. Генерируем γ – БСВ: $F(\xi) = \gamma$, $F(\xi) = \int_0^{\xi} p(x)\,dx$.

Шаг 2. Составляем уравнение: $\int_0^{\xi} \lambda e^{-\lambda x} dx = \gamma$, $-e^{-\lambda x}\Big|_0^{\xi} = \gamma$.

Шаг 3. Разрешаем уравнение относительно ξ: $-e^{-\lambda\xi} + 1 = \gamma$, $-\lambda\xi = \ln(1-\gamma)$, $\xi = -\frac{1}{\lambda}\ln(1-\gamma)$.

Аналогично можно получить формулы и для других распределений: равномерное ($\xi = a + (b-a)\gamma$), Рэлея ($\xi = \sigma\sqrt{-2\ln(1-\gamma)}$), Вейбулла ($\xi = \sqrt[k]{-\lambda^k \ln\left(1-\frac{\gamma}{k\lambda^k}\right)}$), нормальное ($\xi = \mu + \sigma\sqrt{-2\ln(1-\gamma_1)}\cos(2\pi\gamma_2)$).

4. Проверка адекватности моделирования

Рассмотрим гипотезу H_0, которая заключается в том, что смоделированная случайная величина ξ имеет искомую функцию распределения. Рассмотрим статистику χ^2, которая имеет вид

$$\chi^2 = \sum_{j=1}^{r} \frac{(\nu_j - Np_j)^2}{Np_j},$$

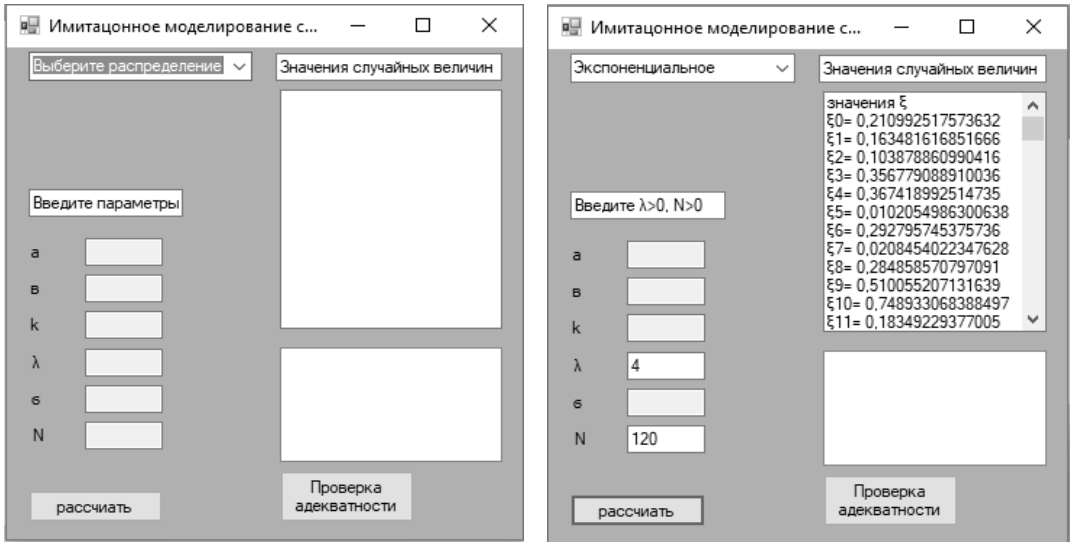
где r - число попарно не пересекающихся отрезков разбиения множества значений x, n_j – количество значений, принадлежащих j-му отрезку, p_j – вероятность попадания в j-й отрезок, N – объём выборки.

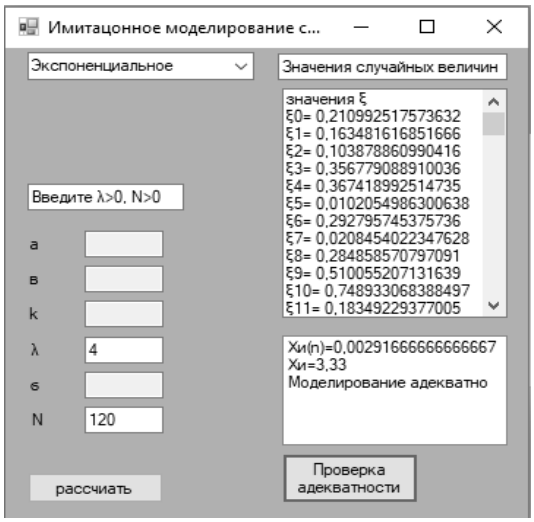
Зададим уровень значимости $\alpha = 0.05$ и число степеней свободы $n = 9$. В таблице распределения χ^2 находим критическое значение $t = 3.33$.

Если вычисленное значение статистики χ^2 превышает ее критическое значение t, то гипотеза H_0 отвергается на уровне значимости α, если $\chi^2 < t$, то гипотеза H_0 принимается, и в этом случае мы делаем вывод, что смоделированная случайная величина ξ имеет заданную функцию распределения.

5. Имитационное моделирование метода обратных функций

Имитационное моделирование заключалось в построении выборки случайной величины по формулам моделирования, полученным по методу обратных функций и оценки адекватности с помощью критерия χ^2. Рассмотрим реализацию имитационного моделирования на примере экспоненциального распределения.





Делаем вывод, что смоделированная случайная величина ξ имеет экспоненциальный закон распределения.

Заключение

В данной работе рассмотрена задача моделирования случайной величины ξ с заданной функцией распределения $F_\xi(x)$ по методу обратной функции, где $F_\xi(x)$ может иметь различный вид. С помощью имитационного моделирования была получена выборка для каждой конкретной функции распределения и оценена адекватность моделирования на основе критерия согласия χ^2. Имитационное моделирование было осуществлено на языке C#.

ИМИТАЦИОННОЕ МОДЕЛИРОВАНИЕ СЛУЧАЙНЫХ ВЕЛИЧИН МЕТОДОМ СУПЕРПОЗИЦИИ

Д.П. Карлина
Томский государственный университет

Введение

В современном мире, с развитием теории вероятностей, для описания и моделирования процессов и явлений необходимы случайные величины, удовлетворяющие конкретным распределениям, в зависимости от условий поставленной задачи [1]. Существует несколько инструментов для моделирования случайной величины, такие, как таблицы случайных цифр, датчики и методы псевдослучайных чисел, один из которых и описан в данной работе, а именно метод суперпозици. В реальном мире редко встречаются процессы или события, которые можно описать только одним законом распределения, или же исходная функция распределения слишком сложна для расчета случайной величины методом обратных функций. И в том, и в другом случае чаще всего используется метод суперпозиции, который и будет рассматриваться в данном работе.

1. Постановка задачи

Моделирование случайной величины заключается в определении («розыгрыше») в нужный по ходу имитации момент времени конкретного значения случайной величины в соответствии с требуемым (заданным) законом распределения [2].

Пусть случайная величина ξ имеет функцию распределения вида $F(x)=\sum_{k=1}^{m} c_k F_k(x)$, где $F_k(x)$ может быть 1) равномерным распределением, 2) экспоненциальным распределением, 3) стандартным нормальным распределением. Требуется построить формулу моделирования случайной величины ξ с помощью метода суперпозиции и оценить адекватность моделирования с помощью критерия χ^2.

2. Метод суперпозиции

Пусть функция распределения $F(x)$ интересующей нас случайной величины ξ представима в виде $F(x)=\sum_{k=1}^{m} c_k F_k(x)$, где $F_k(x)$ – функции распределений, $c_k > 0$, $k=\overline{1,m}$. При $x \to +\infty$ следует, что $\sum_{i=1}^{m} c_i = 1$. Следовательно, можно ввести фиктивную дискретную случайную величину η с распределением

$$\begin{pmatrix} 1 & 2 & \dots & m \\ c_1 & c_2 & \dots & c_m \end{pmatrix}$$

так, что $P\{\eta = k\} = c_k$.Тогда если по числу γ_1 разыграть значение $\eta = k$ случайной величины η, а затем из уравнения $F_k(x) = \gamma_2$ определить ξ, то функция распределения ξ равна $F(x)$, где γ_1, γ_2 имеют равномерное распределение на отрезке [0;1].

Приведем формулы для некоторых случайных величин с известными распределениями, полученные методом обратных функций.

1. Равномерная случайная величина: $\xi = a + (b-a)\gamma$.

2. Экспоненциальная случайная величина: $\xi = -\frac{1}{\lambda}\ln(1-\gamma)$.

3. Гауссовская стандартная случайная величина: $\xi = \sqrt{2\ln(1-\gamma_1)}\cos 2\pi\gamma_2$.

3. Примеры использования метода суперпозиции

Пример 1. Пусть плотность распределения вероятностей случайной величины ξ имеет вид

$$p(x) = \frac{1}{5}\cdot 2e^{-2x} + \frac{4}{5}\cdot 3e^{-3x} .$$

Введем фиктивную дискретную случайную величину η, имеющую следующее распределение (согласно методу суперпозиции):

$$\eta : \begin{pmatrix} 1 & 2 \\ 1/5 & 4/5 \end{pmatrix}.$$

Тогда если $\gamma_1 < 1/5$, то $\eta = 1$, иначе $\eta = 2$.

При $\eta = 1$: $F_1(x) = \gamma_2$, $\int_0^{\xi} 2e^{-2x}dx = \gamma_2$, $\xi = -\frac{1}{3}\ln(1-\gamma_2)$.

При $\eta = 2$: $F_2(x) = \gamma_2$, $\int_0^{\xi} 3e^{-3x}dx = \gamma_2$, $\xi = -\frac{1}{3}\ln(1-\gamma_2)$.

Таким образом, согласно методу суперпозиции, получаем формулу моделирования:

$$\xi = \begin{cases} -\frac{1}{2}\ln(1-\gamma_2), & \gamma_1 < \frac{1}{5}, \\ -\frac{1}{3}\ln(1-\gamma_2), & \gamma_1 \geq \frac{1}{5}. \end{cases}$$

Пример 2. Пусть плотность распределения вероятностей случайной величины ξ имеет вид:

$$p(x) = \frac{1}{6}\cdot\frac{5}{2}(x-1)^4 + \frac{5}{6}\cdot\frac{1}{2} .$$

Введем фиктивную дискретную случайную величину η, имеющую следующее распределение согласно методу суперпозиции:

$$\eta : \begin{pmatrix} 1 & 2 \\ 1/6 & 5/6 \end{pmatrix}.$$

В результате применения метода суперпозиции получаем формулу моделирования для случайной величины ξ:

$$\xi = \begin{cases} 1 + \sqrt[5]{2\gamma_2 - 1}, & \gamma_1 < \dfrac{1}{6}, \\ 2\gamma_2, & \gamma_1 \geq \dfrac{1}{6}. \end{cases}$$

Пример 3. Пусть плотность распределения вероятностей случайной величины ξ имеет вид

$$p(x) = \frac{1}{3} \cdot \frac{1}{2\sqrt{2\pi}} \exp\left\{-\frac{(x-3)^2}{8}\right\} + \frac{2}{3} \cdot \frac{1}{3\sqrt{2\pi}} \exp\left\{-\frac{(x-2)^2}{18}\right\}.$$

Тогда значение η в данном случае примет вид

$$\eta: \begin{pmatrix} 1 & 2 \\ 1/3 & 2/3 \end{pmatrix}.$$

В результате применения метода суперпозиции получаем формулу моделирования для случайной величины ξ:

$$\xi = \begin{cases} 3 + 2\sqrt{-2\ln(1-\gamma_2)}\cos 2\pi\gamma_3, & \gamma_1 < \dfrac{1}{3}, \\ 2 + 3\sqrt{-2\ln(1-\gamma_2)}\cos 2\pi\gamma_3, & \gamma_1 \geq \dfrac{1}{3}. \end{cases}$$

4. Оценка адекватности моделирования с помощью критерия χ^2

Пусть требуется проверить гипотезу H_0 о том, что построенная случайная величина ξ имеет исходную функцию распределения с помощью критерия согласия χ^2. Для этого необходимо вычислить статистику χ^2:

$$\chi^2 = \sum_{i=1}^{k} \frac{(n_i - np_i)^2}{np_i},$$

где k – число разбиений отрезка, n – объем выборки, n_i – количество попаданий случайной величины в i-й отрезок, p_i – вероятность попадания случайной величины в i-й отрезок [3].

Далее задаем уровень значимости α. В таблице распределения χ^2 находим критическое значение t при степенях свободы $k-1$. Если вычисленное значение статистики χ^2 превышает ее критическое значение t, то гипотеза H_0 отвергается на уровне значимости α, если $\chi^2 =< t$, то гипотеза H_0 принимается, и в этом случае мы делаем вывод, что смоделированная случайная величина ξ имеет заданную функцию распределения.

5. Имитационное моделирование

Имитационное моделирование заключалось в построении выборки случайной величины по формулам моделирования, полученным по методу суперпозиции и оценки адекватности с помощью критерия χ^2. Результат моделирования представлен на рис. 1.

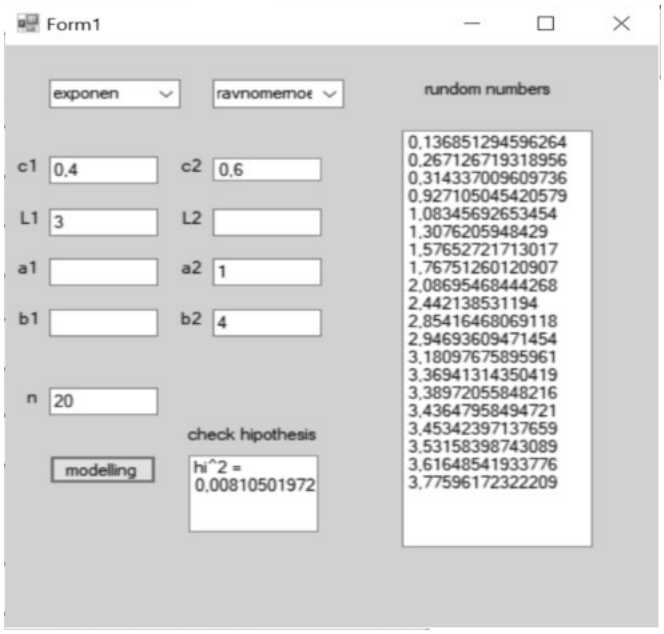


Рис. 1. Пример работы программы моделирования случайной величины методом суперпозиции

Заключение

В данной работе рассмотрена задача моделирования случайной величины ξ, имеющей функцию распределения в виде $F(x)=\sum_{k=1}^{m} c_k F_k(x)$, где $F_k(x)$ может быть равномерным, экспоненциальным или стандартным нормальным распределением. С помощью имитационного моделирования была получена выборка для каждой конкретной функции распределения и оценена адекватность моделирования на основе критерия согласия χ^2. Имитационное моделирование было осуществлено на языке C#.

ЛИТЕРАТУРА

1. *Соболь И.М.* Численные методы Монте-Карло. – М.: Изд-во Наука, 1973. –307 с.

2. *Марголис Н.Ю.* Имитационное моделирование. – Томск: Издательский Дом Томского государственного университета, 2015. – 128 с.

3. *Шуленин В.П.* Математическая статистика. Часть 2. – Томск: Изд-во НТЛ, 2012. – 387 с.

ФИЛЬТРАЦИЯ В ДИСКРЕТНЫХ СИСТЕМАХ СО СЛУЧАЙНЫМИ СКАЧКООБРАЗНЫМИ ПАРАМЕТРАМИ И НЕИЗВЕСТНЫМ ВХОДОМ[1]

К.С. Ким, В.И. Смагин

Томский государственный университет

Введение

Задачи построения оценок для дискретных объектов со случайными скачкообразными параметрами рассматривались в работах [1–4] и др. В настоящей статье рассматривается задача оптимальной фильтрации для дискретной модели со случайными скачкообразными параметрами и неизвестным входом. Предлагаются алгоритмы оптимальной калмановской фильтрации, использующие оценки вектора неизвестного входа [5–10], которые вычисляются с помощью метода наименьших квадратов и его модифика-

[1] Работа выполнена при финансовой поддержке РФФИ, проект № 17-08-00920.

ций, на основе метода скользящего среднего и метода непараметрического сглаживания. Приводится результаты численного моделирования.

1. Постановка задачи

Пусть модель объекта описывается разностным уравнением:

$$x(k+1) = A_\gamma x(k) + B_\gamma f(k) + q_\gamma(k),\ x(0) = x_0,$$

где $x(k) \in \mathrm{R}^m$ – вектор состояния, $\gamma(k)$ – марковская цепь с n состояниями $\gamma_1, \gamma_2, ...\gamma_n$; $f(k) \in \mathrm{R}^r$ – неизвестный вектор, x_0 – случайный вектор, $\bar{x}_0 = \mathrm{M}\{x_0\}$, предполагаются известными матрицы начальных дисперсий $N_{0,i} = \mathrm{M}\left\{(x_0 - \bar{x}_0)(x_0 - \bar{x}_0)^{\mathrm{T}} / \gamma = \gamma_i\right\}$, $i = \overline{1,n}$; A_γ, B_γ – заданные матрицы; $q_\gamma(k)$ – случайные ошибки наблюдений с характеристиками $\mathrm{M}\{q_\gamma(k)\} = 0$, $\mathrm{M}\{q_\gamma(k) q_\gamma^{\mathrm{T}}(j)\} = Q_\gamma \delta_{kj}$ (δ_{kj} – символ Кронекера).

Канал наблюдений имеет вид $y(k) = S_\gamma x(k) + v_\gamma(k)$, где $v_\gamma(k)$ – гауссовская случайная последовательность с характеристиками $\mathrm{M}\{v_\gamma(k)\} = 0$, $\mathrm{M}\{v_\gamma(k) v_\gamma^{\mathrm{T}}(j)\} = V_\gamma \delta_{kj}$. Предполагается, что последовательности $q_\gamma(k)$, $v_\gamma(k)$ независимы между собой, пары матриц S_γ, A_γ детектируемы, матрицы $Q_\gamma \ge 0$, $V_\gamma > 0$.

Вероятность $p_i(k) = \mathrm{P}\{\gamma(k) = i\}$, $i = \overline{1,n}$ удовлетворяет уравнению $p_i(k+1) = \sum_{j=1}^{n} p_{ij} p_j(k)$, $p_i(0) = p_{i,0}$, где $p_{ij} = \mathrm{P}\{\gamma(k+1) = \gamma_j \,|\, \gamma(k) = \gamma_i\}$ – вероятность перехода из состояния i в состояние j за один такт времени, $p_{i,0}$ – начальная вероятность i-го состояния.

По информации, поступившей в момент k, предполагая, что состояние марковской цепи $\gamma(k)$ известно без ошибок, требуется найти оценку вектора состояния $\hat{x}(k)$.

2. Синтез нестационарного фильтра

Вычислим оценку вектора состояния на основе минимизации следующего критерия:

$$J\left[0; T_f\right] = \mathrm{M}\left\{\sum_{k=0}^{T_f} e^{\mathrm{T}}(k) R_\gamma(k) e(k)\right\}, \qquad (1)$$

где $R_\gamma(k) > 0$ – весовая матрица, $e(k) = x(k) - \hat{x}(k)$ – вектор ошибок.

Для решения задачи воспользуемся принципом разделения. Это означает, что сначала мы построим оценки вектора $\hat{x}(k)$ в предположении, что вектор $f(k)$ известен, затем строятся оценки вектора $\hat{f}(k)$ в предположении, что оценка вектора состояния $\hat{x}(k)$ известна.

Для построения оценки воспользуемся рекуррентным алгоритмом фильтрации Калмана

$$\hat{x}(k+1) = A_\gamma \hat{x}(k) + B_\gamma f(k) + K_\gamma(k)\left[y(k+1) - S_\gamma\left(A_\gamma \hat{x}(k) + B_\gamma f(k)\right)\right],\ \hat{x}(0) = \bar{x}_0. \qquad (2)$$

В (2) матрицу коэффициентов передачи $K_\gamma(k)$ определим из следующей теоремы.

Теорема. Пусть существуют положительно определенные матрицы N_i, являющиеся решением задачи Коши:

$$N_i(k+1)=\left(A_i-K_i(k)S_iA_i\right)\left(\sum_{j=1}^{n}p_{i,j}N_j(k)\right)\left(A_i-K_i(k)S_iA_i\right)^{\mathrm{T}}+$$
$$+\left(E-K_i(k)S_i\right)Q_i\left(E-K_i(k)S_i\right)^{\mathrm{T}}+K_i(k)V_iK_i(k)^{\mathrm{T}},\ N_i(0)=N_0. \tag{3}$$

Тогда матрицы оптимальных коэффициентов передачи определятся по формуле:

$$K_i(k)=\left(A_i\left(\sum_{j=1}^{n}p_{i,j}N_j(k)\right)A_i^{\mathrm{T}}S_i^{\mathrm{T}}+EQ_i(k)S_i^{\mathrm{T}}\right)\times$$
$$\times\left[S_iA_i\left(\sum_{j=1}^{r}p_{i,j}N_j(k)\right)A_i^{\mathrm{T}}S_i^{\mathrm{T}}+S_iQ_iS_i^{\mathrm{T}}+V_i(k+1)\right]^{-1}. \tag{4}$$

Здесь введены обозначения для матриц Q_γ, V_γ, R_γ, N_γ, L_γ, A_γ, S_γ, K_γ при $\gamma=\gamma_i$: Q_i, V_i, R_i, N_i, L_i, A_i, S_i, K_i соответственно ($i=\overline{1,n}$).

Доказательство. Представим критерий (1) в виде суммы

$$J\left[0;T_f\right]=\sum_{i=1}^{n}J_i\left[0;T_f\right]. \tag{5}$$

В (5) $J_i\left[0;T_f\right]$ определятся по формуле

$$J_i\left[0;T_f\right]=\sum_{k=0}^{T_f-1}\mathrm{tr}\,N_i(k)R_i(k)+\mathrm{tr}\,N_i\left(T_f\right)L_i\left(T_f\right), \tag{6}$$

где tr – след матрицы, матрица $N_i(k)=\mathrm{M}\left\{e(k)e(k)^{\mathrm{T}}/\gamma=\gamma_i\right\}$ будет определяться из уравнения (3).

Введем функцию Ляпунова следующего вида:

$$W\left(k,N_i(k)\right)=\mathrm{tr}\,N_i(k)L_i(k)+\mathrm{tr}\sum_{t=k}^{T_f}\bar{\Psi}_i(t)L_i(t),$$

где матрицы $L_i(k)>0$ удовлетворяют уравнениям

$$L_i(k)=\left(A_i-K_i(k)S_iA_i\right)^{T}\left(\sum_{j=1}^{n}p_{i,j}L_j(k+1)\right)\left(A_i-K_i(k)S_iA_i\right)+R_i(k),\ L_i\left(T_f\right)=L_{T,i}, \tag{7}$$

матрицы $\bar{\Psi}_i(k)=\left(E-K_i(k)S_i\right)Q_i\left(E-K_i(k)S_i\right)^{\mathrm{T}}+K_i(k)V_iK_i(k)^{\mathrm{T}}+\Psi_i(k)$ (здесь $\Psi_i(k)$ – некоторые положительно определенные матрицы).

Просуммируем по $k=\overline{0,T_f-1}$ конечные разности функции $W\left(k,N_i(k)\right)$, учитывая формулу (7), получим:

$$\sum_{k=0}^{T_f-1}\Delta W\left(k,N_i(k)\right)=\sum_{k=t}^{T_f-1}\left[W\left(k+1,N_i(k+1)\right)-W\left(k,N_i(k)\right)\right]=$$
$$=\sum_{k=0}^{T_f-1}\mathrm{tr}\left[N_i(k+1)L_i(k+1)-N_i(k)L_i(k)-\bar{\Psi}_i(k)L_i(k)\right]. \tag{8}$$

С другой стороны, данное выражение можно представить в следующем виде:

$$\sum_{k=0}^{T_f-1}\Delta W\left(k,N_i(k)\right)=W\left(1,N_i(1)\right)-W\left(0,N_i(0)\right)+\ldots+W\left(T_f,N_i\left(T_f\right)\right)-$$
$$-W\left(T_f-1,N_i\left(T_f-1\right)\right)=\mathrm{tr}\,N_i\left(T_f\right)L_i\left(T_f\right)-\mathrm{tr}\,N_i(0)L_i(0)-\mathrm{tr}\sum_{k=0}^{T_f-1}\bar{\Psi}_i(k)L_i(k). \tag{9}$$

Добавим в формулу (6) разность правых частей (8) и (9). Учитывая, что данная разность равна нулю, то критерий (6) примет вид:

$$J_i\left[0;T_f\right]=\left\{\sum_{k=0}^{T_f-1}\operatorname{tr}N_i(k)R_i(k)-\sum_{k=1}^{T_f-1}\operatorname{tr}N_i(k)L_i(k)+\right.$$

$$+\sum_{k=0}^{T_f-1}\operatorname{tr}\left[\left(A_i-K_i(k)S_iA_i\right)\left(\sum_{j=1}^{n}p_{i,\mathrm{j}}N_j(k)\right)\left(A_i-K_i(k)S_iA_i\right)^{\mathrm{T}}+\right. \tag{10}$$

$$\left.\left.+\left(E-K_i(k)S_i\right)Q_i\left(E-K_i(k)S_i\right)^{\mathrm{T}}+K_i(k)V_iK_i(k)^{\mathrm{T}}\right]L_i(k+1)\right\}.$$

Тогда из условия оптимальности $\partial J/\partial K_i(k)=0$ в силу (5) имеем $\partial J/\partial K_i(k)=\partial J_i/\partial K_i(k)=0$. Применяя правила дифференцирования функции след от матрицы [11] и учитывая (10) получим уравнение:

$$\frac{\partial J_i}{\partial K_i(k)}=\sum_{k=0}^{T_f-1}\left[-L_i(k+1)A_i\left(\sum_{j=1}^{n}p_{i,\mathrm{j}}N_j(k)\right)A_i^{\mathrm{T}}S_i^{\mathrm{T}}-L_i(k+1)A_i\left(\sum_{j=1}^{n}p_{ij}N_j(k)\right)A_i^{\mathrm{T}}S_i^{\mathrm{T}}+\right.$$

$$+L_i(k+1)K_i(k)S_iA_i\left(\sum_{j=1}^{n}p_{i,\mathrm{j}}N_j(k)\right)A_i^{\mathrm{T}}S_i^{\mathrm{T}}+L_i(k+1)K_i(k)S_iA_i\left(\sum_{j=1}^{n}p_{i,\mathrm{j}}N_j(k)\right)A_i^{\mathrm{T}}S_i^{\mathrm{T}}+ \tag{11}$$

$$-L_i(k+1)EQ_iS_i^{\mathrm{T}}-L_i(k+1)EQ_iS_i^{\mathrm{T}}+L_i(k+1)K_i(k)S_iQ_iS_i^{\mathrm{T}}+L_i(k+1)K_i(\xi)S_iQ_iS_i^{\mathrm{T}}+$$

$$\left.+L_i(k+1)K_i(k)V_i+L_i(k+1)K_i(k)V_i\right]=0.$$

Выражая из (11) значение матрицы $K_i(k)$, получим выражение (4).

Вычислим конечную разность функции Ляпунова

$$\Delta W\left(k,N_i(k)\right)=W\left(k+1,N_i(k+1)\right)-W\left(k,N_i(k)\right)=$$

$$=\operatorname{tr}N_i(k+1)R_i(k+1)-\operatorname{tr}N_i(k)R_i(k)+$$

$$+\operatorname{tr}\sum_{t=k+1}^{T_f}\left[\left(E-K_i(k)S_i\right)Q_i\left(E-K_i(k)S_i\right)^{\mathrm{T}}+K_i(t)V_iK_i(t)^{\mathrm{T}}+\Psi_i(t)\right]L_i(t)-$$

$$-\operatorname{tr}\sum_{t=k}^{T_f}\left[\left(E-K_i(k)S_i\right)Q_i\left(E-K_i(k)S_i\right)^{\mathrm{T}}+K_i(t)V_iK_i(t)^{\mathrm{T}}+\Psi_i(t)\right]L_i(t)= \tag{12}$$

$$=\operatorname{tr}N_i(k+1)R_i(k+1)-\operatorname{tr}N_i(k)R_i(k)-$$

$$-\operatorname{tr}\left[\left(E-K_i(k)S_i\right)Q_i\left(E-K_i(k)S_i\right)^{\mathrm{T}}+K_i(k)V_iK_i(k)^{\mathrm{T}}+\Psi_i(k)\right]L_i(k).$$

В силу того, что матрицы N_i, L_i положительно определены, а матрица $\Psi_i(t)$ также положительно определена и задается произвольно, то очевидно, что ее можно задать так, чтобы конечная разность (12) стала отрицательной. Выполнение этого условия гарантирует устойчивость динамики фильтра. Теорема доказана.

3. Синтез стационарного фильтра

В стационарном случае предполагается, что марковская цепь $\gamma_i(k)$ – является эргодической. Тогда матрицы коэффициентов передачи будут постоянными и определятся по формуле:

$$K_i=\left(A_i\left(\sum_{j=1}^{r}p_{i,j}N_j\right)A_i^{\mathrm{T}}S_i^{\mathrm{T}}+EQ_iS_i^{\mathrm{T}}\right)\left[S_iA_i\left(\sum_{j=1}^{r}p_{i,j}N_j\right)A_i^{\mathrm{T}}S_i^{\mathrm{T}}+S_iQ_iS_i^{\mathrm{T}}+V_i\right]^{-1},\ i=\overline{1,n},$$

где матрицы N_i являются решениями матричных уравнений:

$$N_i = \left(A_i - K_i S_i A_i\right)\left(\sum_{j=1}^{n} p_{i,j} N_j\right)\left(A_i - K_i S_i A_i\right)^{\mathrm{T}} + \left(E - K_i S_i\right) Q_i \left(E - K_i S_i\right)^{\mathrm{T}} + \\ + K_i V_i K_i^{\mathrm{T}},\ i = \overline{1,n}. \tag{13}$$

В этом случае, если решение уравнения (13) существует, $N_i > 0$ и справедливо матричное неравенство $\left(E - K_i S_i\right) Q_i \left(E - K_i S_i\right)^{\mathrm{T}} + K_i V_i K_i^{\mathrm{T}} > 0$, следует справедливость следующего условия:

$$\left(A_i - K_i S_i A_i\right)\left(\sum_{j=1}^{n} p_{i,j} N_j\right)\left(A_i - K_i S_i A_i\right)^{\mathrm{T}} - N < 0. \tag{14}$$

Условие (14) гарантирует устойчивость матрицы динамики фильтра $\left(A_i - K_i S_i A_i\right)$.

4. Оценка неизвестного входа

Оценки вектора неизвестного входа, предлагается вычислять с помощью метода наименьших квадратов и его модификаций [5–8]. Тогда оценку $\hat{f}(k)$ можно построить на основе минимизации дополнительного квадратичного критерия [5]:

$$I(f(k)) = \sum_{t=1}^{k} \left\{ \left\| y(t+1) - S_i \tilde{x}(t+1) \right\|_{\overline{W}}^2 + \left\| f(t) \right\|_{\overline{\overline{W}}}^2 \right\}, \tag{15}$$

где $\overline{W}$, $\overline{\overline{W}}$ – положительно определенные весовые матрицы, $\tilde{x}(t+1) = A_i \hat{x}(t) + B_i f(t)$. Минимизируя (15) получим оценки неизвестного входа:

$$\hat{f}^{\langle \mathrm{LSM} \rangle}(k) = \left[B_i^{\mathrm{T}} S_i^{\mathrm{T}} \overline{W} S_i B_i + \overline{\overline{W}} \right]^{-1} B_i^{\mathrm{T}} S_i^{\mathrm{T}} \overline{W} \left\{ y(k+1) - S_i A_i \hat{x}(k) \right\}, \tag{16}$$

$$\hat{f}^{\langle \mathrm{NP} \rangle}(k) = \left[B_i^{\mathrm{T}} S_i^{\mathrm{T}} \overline{W} S_i B_i + \overline{\overline{W}} \right]^{-1} B_i^{\mathrm{T}} S_i^{\mathrm{T}} \overline{W} \hat{\Omega}, \tag{17}$$

$$\hat{f}^{\langle SM \rangle}(k) = \left[B_i^{\mathrm{T}} S_i^{\mathrm{T}} \overline{W} S_i B_i + \overline{\overline{W}} \right]^{-1} B_i^{\mathrm{T}} S_i^{\mathrm{T}} \overline{W} \overline{\Omega} \tag{18}$$

В формуле (17) *j*-я компонента вектора $\hat{\Omega}(k)$ принимает вид:

$$\hat{\Omega}_j(k) = \frac{\sum_{t=1}^{k} \frac{\left[y(t+1) - S_i A_i \hat{x}(t) \right]_j}{\mu_j} K_r\left(\frac{k-t+1}{\mu_j} \right)}{\sum_{t=1}^{k} \frac{1}{\mu_j} K_r\left(\frac{k-t+1}{\mu_j} \right)},$$

где в соотношении (18), $K_r(\cdot)$ является ядерной функцией, и μ_j является коэффициентом сглаживания,

$$K_r(u) = \frac{\exp\left(\frac{-u^2}{2} \right)}{\sqrt{2\pi}}.$$

В формуле (18) $\overline{\Omega}$ – оценка неизвестного входа с помощью алгоритма скользящего среднего:

$$\bar{\Omega}_j(k) = \sum_{t=k-l}^{k} \frac{\left[y(t+1) - S_i A_i \hat{x}(t) \right]_j}{l+1},$$

где $l+1$ – ширина окна сглаживания. Тогда, учитывая (16)–(18) можно представить алгоритм вычисления оценок для модели объекта со скачкообразными параметрами и неизвестным входом:

$$\hat{x}(k+1) = A_i\hat{x}(k) + B_i\hat{f}(k) + K_i(k)\left[y(k+1) - S_i\left(A_i\hat{x}(k) + B_i\hat{f}(k)\right)\right],\ \hat{x}(0) = \overline{x}_0,$$

где $K(k)$ вычисляется по формуле (4).

4. Результаты моделирования

Рассмотрим решение задачи для следующих исходных данных

$$A_1 = \begin{pmatrix} 0.85 & 0.1 \\ -0.05 & 0.94 \end{pmatrix},\ A_2 = \begin{pmatrix} 0.89 & 0.05 \\ -0.02 & 0.45 \end{pmatrix},\ B_1 = B_2 = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix},\ S_1 = S_2 = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix},$$

$$Q_1 = \begin{pmatrix} 0.03 & 0 \\ 0 & 0.04 \end{pmatrix},\ Q_1 = \begin{pmatrix} 0.05 & 0 \\ 0 & 0.03 \end{pmatrix},\ V_1 = \begin{pmatrix} 0.6 & 0 \\ 0 & 0.2 \end{pmatrix},\ V_2 = \begin{pmatrix} 0.7 & 0 \\ 0 & 0.3 \end{pmatrix},$$

$$p_{1,1} = 0.8,\ p_{1,2} = 0.2,\ p_{2,1} = 0.2,\ p_{2,2} = 0.8,$$

$$R_1 = \begin{pmatrix} 0.1 & 0 \\ 0 & 0.15 \end{pmatrix},\ R_2 = \begin{pmatrix} 0.12 & 0 \\ 0 & 0.13 \end{pmatrix},\ \overline{W} = \begin{pmatrix} 0.1 & 0 \\ 0 & 0.1 \end{pmatrix},\ \overline{\overline{W}} = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}.$$

Результаты моделирования представлены в таблицах. Усреднение производилось по 100 реализациям. Среднеквадратические ошибки вычислялись по формулам:

$$\sigma_{f,i} = \sqrt{\frac{\sum_{k=1}^{N}\left(r_i(k) - \hat{r}_i(k)\right)^2}{N-1}},\ \sigma_{x,i} = \sqrt{\frac{\sum_{k=1}^{N}\left(x_i(k) - \hat{x}_i(k)\right)^2}{N-1}}.$$

Таблица 1

Среднеквадратические отклонения оценок вектора состояния

номер компоненты	**Неизвестный вход полностью известен**	МНК	МНК и скользящее среднее	МНК и непараметрическое сглаживание
1	0.167	0.603	0.572	0.37
2	0.219	0.317	0.277	0.257

Таблица 2

Среднеквадратические отклонения оценок вектора неизвестного входа

номер компоненты	МНК	МНК и скользящее среднее	МНК и непараметрическое сглаживание
1	0.761	0.439	0.118
2	0.271	0.254	0.081

Из табл. 1, 2 видно, что применение для оценок неизвестного входа МНК и непараметрического сглаживания позволяет существенно уменьшить среднеквадратические отклонения оценок фильтрации вектора состояния и среднеквадратических отклонений оценок неизвестного входа.

Заключение

Получено решение задачи синтеза нестационарного и стационарного фильтра для линейной дискретной модели с неизвестным входом и скачкообразными параметрами. Показано, что применение алгоритмов непараметрического сглаживания при оценивании неизвестного входа позволяет повысить точность оценивания.

ЛИТЕРАТУРА

1. *Blair W.P., Sworder D.D.* Feedback-control of a class of linear discrete systems with jump parameters and quadratic cost criteria // Int. J. of Control. – 1975. – V. 21. – Is. 5. – P. 833–841.

2. *Costa E.F., De Saporta B.* Linear minimum mean square filters for Markov jump linear systems // IEEE Trans. On Automatic Control. – 2017. – V. 62. – Is. 7. – P. 3567–3572.

3. *Mariton M.* Jump linear systems in automatic control. – New York: Marcel Dekker, 1990.

4. *Terra M.H., Ishihara J.Y., Jesus G., Cerri J.P.* Robust estimation for discrete-time Markovian jump linear systems // IEEE Transactions on Automatic Control. – 2013. – V. 58. – No. 8. – P. 2065-2071.

5. *Janczak D., Grishin Yu.* State estimation of linear dynamic system with unknown input and uncertain observation using dynamic programming // Control and Cybernetics. – 2006. – No 4. – P. 851–862.

6. *Смагин В.И., Смагин С.В.* Фильтрация в линейных дискретных нестационарных системах с неизвестными возмущениями // Вестник Томского государственного университета. Управление, вычислительная техника и информатика. – 2011. – № 3(16). – С. 43–50.

7. *Koshkin G.M., Smagin V.I.* Filtering and prediction for discrete systems with unknown input using nonparametric algorithms // Proc. 10th International Conference on Digital Technologies. Zilina. Slovakia. – 2014. – P. 120–124.

8. *Witczak M.* Fault diagnosis and fault-tolerant control strategies for non-linear systems. Chapter 2. Unknown input observers and filters // Lecture Notes in Electrical Engineering. Springer International Publishing, Switzerland. – 2014. – P. 19–56.

9. *Смагин В.И.* Оценивание состояний нестационарных дискретных систем при неизвестном входе с использованием компенсаций // Известия высших учебных заведений. Физика. – 2015. – Т. 58. – № 7. – С. 122–127.

10. *Smagin V.I., Koshkin G.M.* Kalman filtering and conrol algorithms for systems with unknown disturbances and parameters using nonparametric technique // Proc. 20th International Conference on Methods and Models in Automation and Robotics (MMAR 2015). Miedzyzdroje. Poland. – 2015. – P. 247–251.

11. *Athans M.* The matrix minimum principle // Inf. and Contr. – 1968. – V. 11. – P. 592–606.

СИНТЕЗ АДАПТИВНЫХ СИСТЕМ УПРАВЛЕНИЯ ПРИ НЕПОЛНОЙ ИНФОРМАЦИИ

Г.Н. Решетникова, А.В. Володенок

Томский государственный университет

Введение

При решении задач управления приходится иметь дело с объектами, математические модели которых описываются динамическими стохастическими системами, где помехи, действующие на объект, рассматриваются как случайные величины. Кроме того, информация о состоянии объекта чаще всего бывает неполной и искаженной случайными ошибками [1,2]. В связи со сложностью таких задач большое значение приобретает разработка программ, позволяющих осуществлять синтез следящих систем адаптивного управления для стационарных и нестационарных стохастических моделей объектов различных порядков при полном и неполном измерениях с ошибками.

В настоящей работе предлагается программа для синтеза адаптивных следящих систем управления при неполной информации на языке программирования Python с использованием библиотек Matplotlib, NumPy, Tkinter. Приводятся результаты моделирования системы управления для нестационарной модели судна при изменении курса [3].

1. Алгоритмы синтеза адаптивной следящей системы управления при неполной информации

Рассматривается задача формирования управляющих воздействий с использованием локального квадратичного критерия при неполном измерении. Для оценивания состояния и параметров модели объекта используются фильтры Калмана.

1.1. Построение стационарных и нестационарных дискретных стохастических моделей

Пусть модель объекта управления описывается системой линейных стохастических дифференциальных уравнений вида

$$\dot{x}(t) = \overline{A}(t)x(t) + \overline{B}(t)u(t) + \overline{F}(t)q(t)\,,\ \ x(t_0) = x_0\,, \qquad (1)$$

где $x(t) \in R^n$, $u(t) \in R^m$, $q(t) \in R^{l_1}$ – векторы, задающие состояние, управление и внешние возмущения. При этом предполагается, что внешние возмущения являются аддитивными и задаются вектором, компоненты которого являются гауссовскими случайными величинами с заданными характеристиками

$$M\{q(t)\} = \overline{q}(t),\ M\left\{\left(q(t) - \overline{q}(t)\right)\left(q(\tau) - \overline{q}(\tau)\right)^T\right\} = Q(t)\delta(t-\tau).$$

Матрицы $\overline{A}(t)$, $\overline{B}(t)$, $\overline{F}(t)$ описывают динамические свойства, влияние управляющих воздействий и внешних возмущений.

При использовании цифровых технологий возникает необходимость в построении дискретной модели объекта. Пусть Δt – шаг дискретизации, совпадающий с периодом квантования управляющего сигнала. Предполагая, что $u(t)$ и $q(t)$ являются кусочно-постоянными непрерывными слева функциями, т.е.

$$u(t) = u(t_k),\ q(t) = q(t_k),\ t_k \le t < t_{k+1},\ t_{k+1} = t_0 + k\Delta t,\ k = \overline{0, N},$$

разностное уравнение, соответствующее (1), имеет вид

$$x(k+1) = A(k)x(k) + B(k)u(k) + F(k)q(k),\ x(0) = x_0,$$

где $A(k) = I_n + \Delta t\overline{A}(t_k)$, $B(k) = \Delta t\overline{B}(t_k)$, $F(k) = \sqrt{\Delta t}\overline{F}(t_k)$, I_n – единичная матрица порядка n, $q(k)$ – векторы, компоненты которых являются последовательностью гауссовских случайных величин с характеристиками

$$M\{q(k)\} = \overline{q}(k),\ M\left\{\left(q(k) - \overline{q}(k)\right)\left(q(j) - \overline{q}(j)\right)^T\right\} = Q(k)\delta_{k,j}.$$

Модель объекта является стационарной, если все элементы матриц $\overline{A}(t)$, $\overline{B}(t)$, $\overline{F}(t)$ не зависят от времени. Если среди элементов этих матриц есть зависящие от времени, то модель объекта является нестационарной. Построение такой модели осуществляется следующим образом. Задается число моментов времени τ_i, $i = \overline{1, s}$, для которых элементы матриц $\overline{A}(t)$, $\overline{B}(t)$, $\overline{F}(t)$ зависят от времени. Для каждого зависящего от времени элемента строится вектор, размерности s. Таким образом, получают векторы a_{ja}, b_{jb}, f_{jf}, содержащие s элементов. При этом в каждой из матриц $\overline{A}(t)$, $\overline{B}(t)$, $\overline{F}(t)$ содержится конкретное число зависящих от времени элементов. Таким образом значения ja, jb, jf являются различными, т.е. для каждой из матриц $\overline{A}(t)$, $\overline{B}(t)$, $\overline{F}(t)$ задается свой набор векторов, число которых определяется значениями ja, jb, jf. Конкретный элемент матриц в момент t_k $k = \overline{1, N}$, вычисляется с помощью многочлена Лагранжа по значениям τ_i, z_i, $i = \overline{1, s}$,

$$L_s(t, \tau, z) = \sum_{i=1}^{S} z_i \prod_{\substack{j=1 \\ j \ne i}}^{S} \frac{t - \tau_j}{\tau_i - \tau_j},$$

где z_i – элемент одного из векторов a_{ja}, b_{jb}, f_{jf} для конкретного значения ja, jb, jf.

1.2. Описание математических моделей измерительного комплекса

При синтезе управляющих воздействий информация о состоянии объекта поступает с измерительного комплекса, который достаточно часто содержит ошибки измерений. Будем предполагать, что математическая модель измерительного комплекса имеет вид

$$y(k) = x(k) + r(k), \tag{2}$$

где $y(k) \in R^l$. В (2) $r(k)$ – гауссовские последовательности с характеристиками

$$M\{r(k)\} = 0,\ M\left\{r(k)r^T(j)\right\} = R\delta_{k,j},\ M\left\{q(k)r^T(j)\right\} = 0.$$

Информация, поступающая с измерительного комплекса, может быть неполной, что соответствует, например, отказу некоторых датчиков. В этом случае математическую модель измерительного комплекса запишем в виде

$$y(k) = Hx(k) + Hr(k),$$

где H – матрица канала измерений размерности $l \times n$, состоящая из нулей и единиц, нулевые столбцы которой соответствуют компонентам вектора состояния, недоступным измерению.

1.3. Оценивание состояния и параметров модели объекта

В реальных условиях управление объектом осуществляется по результатам измерений, имеющих погрешности, причём могут существовать компоненты вектора состояния, недоступные измерению. Кроме того, необходимо учитывать наличие в модели неизвестных параметров, т.е. математическая модель объекта задается в виде

$$x(k+1) = A(k,\theta(k))x(k) + B(k,\theta(k))u(k) + F(k)q(k),\ x(0) = x_0, \qquad (3)$$

где $\theta(k)$ – вектор неизвестных параметров размерности N_θ.

В связи с этим для формирования управляющих воздействий будем строить оценки векторов состояния и параметров $\hat{x}(k)$, $\hat{\theta}(k)$ в дискретные моменты времени по текущей информации $y(k)$. При этом предполагается, что априорные распределения векторов начальных условий x_0 и $\theta(0)$ являются гауссовскими:

$$M\{x_0\} = \overline{x}_0,\ M\left\{(x_0 - \overline{x}_0)(x_0 - \overline{x}_0)^T\right\} = P_x,$$

$$M\{\theta(0)\} = \overline{\theta}(0),\ M\left\{(\theta(0) - \overline{\theta}(0))(\theta(0) - \overline{\theta}(0))^T\right\} = P_\theta,$$

где P_x, P_θ – ковариационные матрицы ошибок начальных условий. Для построения оценок будем использовать дискретные параллельные фильтры Калмана.

Рекуррентный алгоритм для оценивания вектора состояния имеет вид:

$$\begin{aligned}
&\hat{x}(k+1) = \hat{x}(k+1/k) + K(k)\left[y(k+1) - H\hat{x}(k+1/k)\right],\\
&\hat{x}(k+1/k) = A\left(k,\hat{\theta}(k)\right)\hat{x}(k) + B\left(k,\hat{\theta}(k)\right)u(k) + F(k)\overline{q}(k),\ \hat{x}(0) = \overline{x}_0,\\
&K(k) = P_x(k+1/k)H^T\left[HP_x(k+1/k)H^T + HRH^T\right]^{-1},\\
&P_x(k+1/k) = A\left(k,\hat{\theta}(k)\right)P_x(k)A^T\left(k,\hat{\theta}(k)\right) + F(k)Q(k)F^T(k),\\
&P_x(k+1) = \left[I_n - K(k)H\right]P_x(k+1/k),\ P_x(0) = P_{x_0}.
\end{aligned} \qquad (4)$$

В (4) $\hat{x}(k+1/k)$ – экстраполированная оценка вектора состояния $x(k+1)$, $P_x(k+1/k)$ – прогноз дисперсионной матрицы $P_x(k)$ на один шаг вперёд.

Рекуррентный алгоритм для оценивания параметров модели имеет вид:

$$\begin{aligned}
&\hat{\theta}(k+1) = \hat{\theta}(k) + L(k)\left[y(k+1) - H\Phi\left(\hat{x}(k),u(k)\right)\hat{\theta}(k) - Hf\left(\hat{x}(k),u(k)\right)\right],\ \hat{\theta}(0) = \overline{\theta}_0,\\
&L(k) = P_\theta(k)\Phi^T\left(\hat{x}(k),u(k)\right)M^{-1},\\
&M(k) = H\Phi\left(\hat{x}(k),u(k)\right)P_\theta(k)\Phi^T\left(\hat{x}(k),u(k)\right)H^T + HF(k)Q(k)F^T(k)H^T + HRH^T,\\
&P_\theta(k+1) = \left[I_{N_R} - L(k)H\Phi\left(\hat{x}(k),u(k)\right)\right]P_\theta(k),\ P_\theta(0) = P_{\theta_0}.
\end{aligned} \qquad (5)$$

В (5) матрица $\Phi(\cdot) \in R^{n \times N_R}$ и вектор $f(\cdot) \in R^n$ получаются при представлении системы (3) в виде

$$x(k+1) = \Phi\left(x(k),u(k)\right)\theta(k) + f\left(x(k),u(k)\right) + F(k)q(k),\ x(0) = x_0.$$

1.4. Формирование управляющих воздействий

Управляющие воздействия $u(k)$ будем строить при минимизации математического ожидания локального квадратичного критерия [3]

$$J(k) = M\left\{\left(x(k+1) - x_z(k)\right)^T C\left(x(k+1) - x_z(k)\right) + u^T(k) D u(k)\right\}, \quad (6)$$

где C – неотрицательно определённая, а D – положительно определённая весовые матрицы, $x_z(k)$ – отслеживаемое состояние, которое может быть как постоянным, так и изменяющимся во времени.

Из условия оптимальности критерия (6) адаптивное управление для модели (3) будет иметь вид:

$$u(k) = -\left[B^T\left(k, \hat{\theta}(k)\right) C B\left(k, \hat{\theta}(k)\right) + D\right]^{-1} B^T\left(k, \hat{\theta}(k)\right) C\left[A\left(k, \hat{\theta}(k)\right)\hat{x}(k) + F(k)\overline{q}(k) - x_z(k)\right],$$

где $\hat{x}(k)$, $\hat{\theta}(k)$ – оценки состояния и параметров, полученные с помощью фильтров Калмана (4), (5), в матрицах $A\left(k, \hat{\theta}(k)\right)$, $B\left(k, \hat{\theta}(k)\right)$ вместо точных значений параметров указываются их оценки.

Качество функционирования системы управления зависит от весовых матриц критерия (6). Весовые матрицы C и D локального критерия определяются решением алгебраического уравнения Риккати, которое решается по итерационной схеме

$$S_{i+1} = \left[A^T(0) S_i A(0) - A^T(0) S_i B(0)\left(B^T(0) S_i B + D\right)^{-1} B^T(0) S_i A(0) + C - S_i\right]\Delta t + S_i,$$

$$i = 0, 1, 2, \ldots, \; S_0 = C,$$

и при выполнении условия $\dfrac{\|S_{i+1} - S_i\|}{\|S_{i+1}\|} \le \varepsilon$ матрица C полагается равной C_{i+1}.

При синтезе управлений на величину управляющего воздействия обычно накладываются ограничения. Достаточно часто эти ограничения задаются в форме неравенств: $U_{\min i}(k) \le u_i(k) \le U_{\max i}(k)$, $i = \overline{1, m}$. Тогда, учитывая ограничения, будем полагать:

$$\overline{u}_i(k) = \begin{cases} U_{\min i}(k), & u_i(k) < U_{\min i}(k), \\ u_i(k), & U_{\min i}(k) \le u_i(k) \le U_{\max i}(k), \\ U_{\max i}(k), & u_i(k) > U_{\max i}(k). \end{cases}$$

Ограничения, задаваемые для вектора состояния, указывают максимальную величину отклонения от отслеживаемого состояния и определяют качество функционирования управляемого объекта. Если все компоненты вектора состояния удовлетворяют заданным ограничениям, то это говорит о том, что управление обеспечивает требуемое качество функционирования объекта.

2. Описание программного обеспечения

В результате анализа предметной области можно сделать вывод, что программное обеспечение должно обладать возможностью обработки различных типов данных, иметь интерфейс и возможность вывода результатов работы программы. Оптимальным решением является разделение множества всех задач на несколько независимых компонент, каждая из которых содержит реализацию конкретных методов.

2.1. Используемые технологии

Программное обеспечение реализовано на языке Python 3.6 с использованием библиотек Matplotlib, NumPy, Tkinter. Библиотека Matplotlib используется для вывода графиков, NumPy предоставляет реализацию вычислительных алгоритмов в виде функций и операторов, оптимизированных для работы с многомерными массивами, Tkinter нужна для реализации графического интерфейса.

2.2. Численные результаты

Для иллюстрации работоспособности и качества разработанной программы используется моделирование движения судна при изменении курса. Математическая модель объекта задана в виде (1) и является нестационарной. Результаты моделирования приведены на рис. 1.

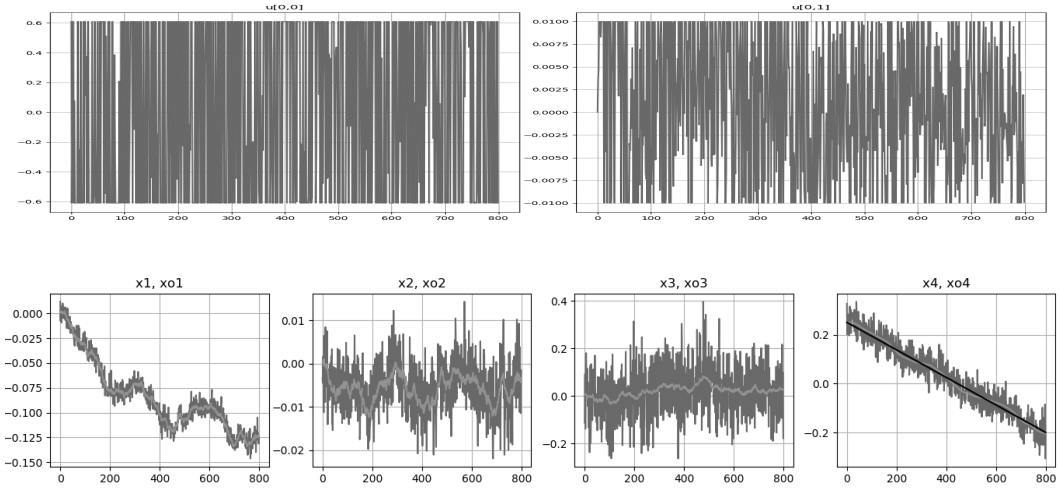


Рис. 1. u1 – отклонение угла перекладки руля, u2 – отклонение режима работы главного двигателя, x1 – отклонение угла дрейфа и оценка, x2 – отклонение угловой скорости дрейфа и оценка, x3 – отклонение скорости движения и оценка, x4 – отклонение угла курса и оценка, заданный курс.

Приведённые результаты иллюстрируют качество отслеживания заданного курса в процессе функционирования управляемого объекта, при этом видно, что управляющее действие руля приводит к слежению за заданным курсом.

Заключение

В настоящей работе рассмотрен синтез адаптивной следящей системы управления при неполной информации. Для реализации этой задачи разработана программа на языке Python, которая позволяет работать со стационарными и нестационарными стохастическими моделями различных размерностей при неполном измерении состояния.

Полученные результаты могут быть использованы для определения минимального набора датчиков измерительного комплекса, при котором функционирование объекта будет осуществляться без существенного ухудшения качества.

ЛИТЕРАТУРА

1. *Решетникова Г.Н.* Адаптивные системы: учеб.пособие. – Томск: Издательский Дом Томского государственного университета, 2016. – 112 с.

2. *Решетникова Г.Н.* Моделирование систем: учеб.пособие. –Томск: Томский гос. ун-т систем упр. и радиоэлектроники, 2007. – 441 с.

3. *Reshetnikova G.N., Khabibulina N. Yu., Kotcubinskiy V.P., Polonskaya M.S.* Servo systems with incomplete information // Proc. 2017 International Siberian Conference on Control and Communications (SIBCON-2017) Astana, Kazakhstan, IEEE, 2017.

МОДЕЛЬ УПРАВЛЕНИЯ ЗАПАСАМИ ОДНОРОДНОЙ ПРОДУКЦИИ С РЕЛЕЙНЫМ УПРАВЛЕНИЕМ ТЕМПОМ ПРОИЗВОДСТВА И ММР-ПОТОКОМ МОМЕНТОВ ПРОДАЖ

К.И. Лившиц, Т.А. Поволоцкая
Томский государственный университет

Введение

Одной из классических задач теории управления запасами является задача производства и сбыта однородной продукции. Систематическое исследование моделей управления запасами началось, по-видимому, еще в 1950-е гг. К настоящему времени опубликовано огромное количество работ, посвященных данной тематике, в которых используется либо чисто детерминистский подход к решению задачи, требующий полной информации о процессе реализации продукции [1,2], либо рассматриваются стохастические модели. Из работ последнего времени, в которых рассматриваются стохастические модели, отметим, например, работы [3–5].

Целью данной работы является определение вероятностных характеристик модели управления запасами с релейным управлением темпом производства, ММР-потоком моментов продаж с двумя состояниями и экспоненциальным распределением величин покупок.

1. Постановка задачи

В настоящей работе задача производства и сбыта продукции рассматривается при следующих предположениях. Пусть $S(t)$ количество продукции в момент времени t. Будем считать, что продукция производится со скоростью $C(S)$, так, что за время Δt поступает $C(S(t))\Delta t$ единиц продукции. Накопленная продукция непрерывно реализуется. Величины покупок являются независимыми одинаково распределенными случайными величинами с плотностью распределения $\varphi(x)$. Моменты продаж образуют ММР-поток с интенсивностью $\lambda(t)$. Процесс $\lambda(t)$ является однородной цепью Маркова с непрерывным временем и двумя состояниями $\lambda(t) = \lambda_1, \lambda_2$. Такое предположение представляется естественным, т.к. продажи происходят в принципе в случайные моменты времени, а интенсивность потока продаж с течением времени случайным образом может изменяться. Переход из состояния в состояние задается матрицей инфинитезимальных характеристик $Q = \left[q_{ij} \right]$. Финальные вероятности состояний λ_1 и λ_2 равны соответственно $\pi_1 = \dfrac{q_{22}}{q_{11} + q_{22}}$, $\pi_2 = \dfrac{q_{11}}{q_{11} + q_{22}}$.

Различные варианты выбора функции $C(S)$ приводят к различным моделям производства и сбыта продукции. В простейших случаях, которые и будут в дальнейшем рассматриваться, управление $C(S)$ является релейным с различными вариантами выбора точек переключения управления. Приведем несколько возможных вариантов. Пусть

$$C(S) = \begin{cases} C_0, S < S_0, \\ C_1, S \geq S_0, \end{cases} \tag{1}$$

где S_0 – пороговое значение желаемого запаса продукции, $C_1 < \lambda_0 a$, $C_0 > \lambda_0 a$ и $\lambda_0 = \sum_{i=1}^{n} \lambda_i \pi_i$ – средняя интенсивность потока покупок. Выбор управления вида (1) га-

рантирует в стационарном режиме стабилизацию уровня запаса продукции $S(t)$ около желаемого значения S_0. Отрицательные значения запаса $S(t)$ интерпретируются как неудовлетворенный спрос (накопленные заказы подлежат немедленному исполнению при пополнении запаса) [6].

При выборе функции $C(S)$ в виде

$$C(S) = \begin{cases} C, S \le S_0, \\ 0, S > S_0 \end{cases} \tag{2}$$

величина S_0 может интерпретироваться как максимально допустимый уровень запаса. Отметим, что при детерминированной постановке задачи управления производством и сбытом продукции релейное управление вида (2) является оптимальным [2].

Наконец, возможен вариант

$$C(S) = \begin{cases} C, 0 \le S \le S_0, \\ 0, S < 0 \text{ и } S > S_0, \end{cases} \tag{3}$$

который и будет рассматриваться подробно в данной работе. В этом случае неудовлетворенный спрос не учитывается, неудовлетворенные заказы теряются.

2. Функция распределения количества продукции в стационарном режиме

Обозначим $P_i(s,t) = P\{S(t) < s; \lambda(t) = \lambda_i\}$, $i = \overline{1,n}$. Можно показать, что функции $P_i(s,t)$ удовлетворяют системе уравнений [5]

$$\frac{\partial P_i(s,t)}{\partial t} + C(s)\frac{\partial P_i(\mathrm{s},t)}{\partial s} = -\lambda_i P_i(s,t) + \sum_{j=1}^{n} q_{ji} P_j(s,t) + +\lambda_i \int_0^{\infty} P_i(s+x,t)\varphi(x)dx .$$

Рассмотрим далее стационарный случай. Обозначим $P_i(s) = \lim\limits_{t\to\infty} P_i(s,t)$. Функции $P_i(s)$ удовлетворяют уравнениям

$$C(s)\dot{P}_i(s) = -\lambda_i P_i(s) + \sum_{j=1}^{n} q_{ji} P_j(s) + \lambda_i \int_0^{\infty} P_i(s+x,t)\varphi(x)dx \tag{4}$$

с вытекающими из их определения условиями нормировки

$$\lim_{s\to\infty} P_i(s) = \pi_i . \tag{5}$$

Функция распределения количества продукции $P(S)$ в стационарном режиме будет, очевидно, равна

$$P(s) = \sum_{k=1}^{n} P_i(s) . \tag{6}$$

Рассмотрим случай, когда функция $C(S)$ определяется соотношением (3). Система уравнений (4) тогда перепишется в виде

$$\begin{aligned} &-\lambda_i P_i(s) + \sum_{j=1}^{n} q_{ji} P_j(s) + \lambda_i \int_0^{\infty} P_i(s+x)\varphi(x)dx = 0, \ s > S_0, \\ &C\dot{P}_i(s) = -\lambda_i P_i(s) + \sum_{j=1}^{n} q_{ji} P_j(s) + \lambda_i \int_0^{\infty} P_i(s+x)\varphi(x)dx, \ 0 \le s \le S_0. \end{aligned} \tag{7}$$

Решения уравнений (7) с учетом граничных условий (5) и соотношений (2) имеют, очевидно, вид $P_i(s) = \pi_i$, $s > S_0$, откуда при $s \le S_0$

$$C\dot{P}_i(s) = -\lambda_i P_i(s) + \sum_{j=1}^{n} q_{ji} P_j(s) + \lambda_i \int_0^{S_0 - s} P_i(s+x)\varphi(x)dx + \lambda_i \pi_i \int_{S_0 - s}^{\infty} \varphi(x)dx , \ s \le S_0$$

с граничными условиями $P_i(0) = 0$.

3. Распределение количества продукции для случая потока с двумя состояниями и экспоненциального распределения величин покупок

Пусть величины покупок имеют экспоненциальное распределение $\varphi(x) = \frac{1}{a} e^{-\frac{x}{a}}$. Решение системы уравнений (7), которая перепишется как

$$C\dot{P}_k(s) = -\lambda_k P_k(s) + \sum_{j=1}^{2} q_{jk} P_j(s) + \frac{\lambda_k}{a} \int_0^{S_0 - s} P_k(s+x) e^{-\frac{x}{a}} dx + \lambda_k \pi_k e^{\frac{s - S_0}{a}}, \ 0 \le s \le S_0, \ k = 1,2, \quad (8)$$

будем искать в виде

$$P_m(s) = \sum_{j=1}^{\infty} A_{jm} e^{\gamma_j s}, \ m = 1,2. \quad (9)$$

Подставляя соотношения (9) в уравнения (8) и приравнивания коэффициенты при линейно независимых функциях $e^{\gamma_i s}$ и $e^{\frac{s}{a}}$, получим систему соотношений

$$\begin{aligned} &\sum_{k=1}^{\infty} A_{k1} \frac{1}{a\gamma_k - 1} e^{\gamma_k S_0} + \pi_1 = 0, \\ &CA_{k1}\gamma_k = -\lambda_1 A_{k1} + q_{11} A_{k1} - q_{22} A_{k2} - \lambda_1 A_{k1} \frac{1}{a\gamma_k - 1} \end{aligned} \quad (10)$$

при $m = 1$ и

$$\begin{aligned} &\sum_{k=1}^{\infty} A_{k2} \frac{1}{a\gamma_k - 1} e^{\gamma_k S_0} + \pi_2 = 0, \\ &CA_{k2}\gamma_k = -\lambda_2 A_{k2} - q_{11} A_{k1} + q_{22} A_{k2} - \lambda_2 A_{k2} \frac{1}{a\gamma_k - 1} \end{aligned} \quad (11)$$

при $m = 2$.

Относительно пар входящих в уравнения (10) и (11) коэффициентов A_{k1} и A_{k2} имеем однородную систему уравнений

$$\begin{aligned} &\left(C\gamma_k + \frac{\lambda_1 a\gamma_k}{a\gamma_k - 1} - q_{11} \right) A_{k1} + q_{22} A_{k2} = 0, \\ &q_{11} A_{k1} + \left(C\gamma_k + \frac{\lambda_2 a\gamma_k}{a\gamma_k - 1} - q_{22} \right) A_{k2} = 0. \end{aligned} \quad (12)$$

Для существования ненулевого решения её определитель

$$\Delta(\gamma_k) = \gamma_k \left[\gamma_k \left(C - \frac{\lambda_1 a}{1 - a\gamma_k} \right) \left(C - \frac{\lambda_2 a}{1 - a\gamma_k} \right) - (q_{11} + q_{22}) \left(C - \frac{\lambda_0 a}{1 - a\gamma_k} \right) \right]$$

должен равняться нулю. Отсюда $\gamma_1 = 0$, а остальные γ_k должны являться решениями уравнения

$$f(z) = z \left(C - \frac{\lambda_1 a}{1 - az} \right) \left(C - \frac{\lambda_2 a}{1 - az} \right) - (q_{11} + q_{22}) \left(C - \frac{\lambda_0 a}{1 - az} \right) = 0. \quad (13)$$

Покажем, что уравнение имеет три различных корня.

Очевидно, что должно выполняться условие $C - \lambda_0 a > 0$, т.е. $C = (1 + \theta)\lambda_0 a$, $\theta > 0$, т.к. темп производства должен быть выше среднего количества продукции, реализуемого в единицу времени, иначе всегда будет дефицит товара. Поэтому $f(0) > 0$, т.к.

$q_{11}+q_{22}<0$. Пусть k – корень уравнения $C-\dfrac{\lambda_0 a}{1-az}=0$, т.е. $k=\dfrac{C-\lambda_0 a}{Ca}=\dfrac{\theta}{(1+\theta)a}$. Тогда $f(k)=k\dfrac{(\lambda_1-\lambda_0)(\lambda_2-\lambda_0)a^2}{(1-ak)^2}<0$, т.к. либо $\lambda_1-\lambda_0<0$, либо $\lambda_2-\lambda_0<0$. Поэтому имеется положительный корень γ_2 и $0<\gamma_2<k<1/a$.

При $z\to\dfrac{1}{a}$ главный член в $f(z)$ имеет вид $f(z)\square z\dfrac{\lambda_1\lambda_2 a^2}{(1-az)^2}\to+\infty$, поэтому имеется положительный корень γ_2 уравнения (13) и $k<\gamma_2<\dfrac{1}{a}$.

При $z\to-\infty$ главный член в $f(z)\square z$ и $f(z)\to-\infty$. Поэтому существует корень $\gamma_4<0$ уравнения (13).

Наконец, $f(z)$ после приведения к общему знаменателю есть отношение многочлена третьей степени к многочлену второй степени. Поэтому других корней у функции $f(z)$ нет.

Таким образом. разложения (9) включают в себя всего четыре слагаемых с различными показателями γ_j.

Из системы (12) получаем, что

$$A_{k2}=A_{k1}\omega_k, \tag{14}$$

где

$$\omega_k=-\frac{1}{q_{22}}\left(C\gamma_k-\frac{\lambda_1 a\gamma_k}{1-a\gamma_k}-q_{11}\right), \tag{15}$$

и с учетом граничных условий $P_i(0)=0$ получаем систему соотношений на определение коэффициентов A_{k1}

$$\sum_{k=1}^{4}A_{k1}\frac{1}{a\gamma_k-1}e^{\gamma_k S_0}=-\pi_1, \tag{16}$$

$$\sum_{k=1}^{4}A_{k1}\omega_k\frac{1}{a\gamma_k-1}e^{\gamma_k S_0}=-\pi_2, \tag{17}$$

$$\sum_{k=1}^{4}A_{k1}=0,\ \sum_{k=1}^{4}\omega_k A_{k1}=0. \tag{18}$$

Таким образом, нахождение решения (9) уравнений (8) сводится к нахождению корней уравнения (13), решению системы уравнений (16)–(18) и использованию соотношений (14) и (15).

На рис. 1 приведены решения (9) $P_1(s)$ и $P_2(s)$ уравнений (8) и безусловные функция распределения количества продукции $P(s)$ (6) при $\lambda_1=15$, $\lambda_2=5$, $a=1$, $S_0=20$.

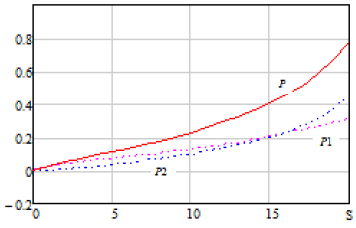


Рис. 1. Функция распределения количества продукции

Заключение

В работе получены выражения для распределения количества однородной продукции при релейном управлении темпом производства и ММР-потоке моментов продаж продукции при дополнительном предположении, что распределение величин покупок является экспоненциальным. Предлагаемая методика расчета статистических характеристик может быть использована для анализа более сложных вариантов управления темпом производства и интенсивностью потока продаж, например, для одновременного учета зависимости интенсивности потока продаж от меняющейся продажной цены продукции.

ЛИТЕРАТУРА

1. *Горский А.А., Локшин Б.Я.* Математическая модель процесса производства и продажи для управления и планирования производства // Фундаментальная и прикладная математика. – 2002. – Т. 8. – № 1. – С. 34–45.

2. *Параев Ю.И.* Решение задачи об оптимальном производстве, хранении и сбыте товара // Известия РАН. Теория и системы управления. – 2000.– № 2.– С. 103–107.

3. *Livshits K., Ulyanova E.* Switch-Hysteresis Control Of the Production Process in a Model with Perishable Goods// Information technologies and mathematical modeling: Proceeding of the 15 th International Scientific Conference, ITMM 2016 named after A.F. Terpugov. – CCIS. – Springer. – 2016. – V. 638. – P. 192–206.

4. *Livshits K., Ulyanova E.* Switch-hysteresis Control of the Selling Times Flow in a Model with Perishable Goods // Information technologies and mathematical modeling: Proceeding of the 14 th International Scientific Conference, ITMM 2015 named after A.F. Terpugov. – CCIS. – Springer. – 2015. – V. 564. – P. 263–274.

5. *Лившиц К.И., Ульянова Е.С.* Модель управления запасами однородной продукции с релейным управлением темпом производства и ММР-потоком моментов продаж // Вестник ТГУ. Управление, вычислительная техника и информатика.– 2018. – № 44. – С. 50 – 61.

6. *Карлин С.* Основы теории случайных процессов. – М.: Издательство «Мир», 1971. – 536 с.

РОЗНИЧНАЯ ПРОДАЖА МНОГОНОМЕНКЛАТУРНОЙ ПАРТИИ ПРОДУКЦИИ С ОГРАНИЧЕННЫМ СРОКОМ ГОДНОСТИ ПРИ ММР-ПОТОКЕ МОМЕНТОВ ПРОДАЖ

К.И. Лившиц, Л.В. Чупрасова, Е.С. Ульянова

Томский государственный университет

Введение

Одной из классических задач теории управления запасами является проблема управления запасами, имеющими ограниченный срок годности (the newsvendor problem) [1,2]. Первоначально исследовалась задача реализации единственного товара, имеющего фиксированную цену, в предположении, что поток покупателей является пуассоновским с фиксированной интенсивностью, затем аналогичные задачи начали рассматриваться для многономенклатурных моделей и более сложных моделях потока покупателей [3,4,5]. Целью данной работы является определение оптимального размера много – номенклатурной партии товара, имеющего ограниченный срок реализации, в предположении, что моменты продажи товара образуют ММР-поток с двумя состояниями. Такое предположение представляется довольно естественным, т.к. интенсивность потока покупателей в любой торговой точке изменяется с течением времени случайным образом.

1. Постановка задачи

Продавец средствами в размере S_0 приобретает партию из n товаров $\xi = \left[\xi_1, \xi_2, \ldots, \xi_n\right]^{\mathrm{T}}$ по оптовой цене $d = \left[d_1, d_2, \ldots d_n\right]^{\mathrm{T}}$ и перепродает ее по розничной цене $c = \left[c_1, c_2, \ldots c_n\right]^{\mathrm{T}}$. Предполагается, что время реализации ограничено. По истечении этого времени T товар не может быть продан, а продавец несет дополнительные затраты $b = \left[b_1, b_2, \ldots, b_n\right]^{\mathrm{T}}$, связанные с утилизацией нереализованной части товара. Необхо-

димо определить оптимальный размер партии товара ξ, который обеспечивает продавцу максимальную среднюю прибыль.

Предполагается, что моменты покупок образуют ММР-поток с интенсивностью $\lambda(t)$. Интенсивность $\lambda(t)$ является однородной цепью Маркова с непрерывным временем и двумя состояниями $\lambda(t)=\lambda_i$, $i=1,2$. Переход из состояния в состояние задается матрицей инфинитезимальных характеристик $Q=\left[q_{ij}\right]$, где $q_{ij}>0$ при $i\neq j$ и $q_{j1}+q_{j2}=1$. Таким образом, переход из состояния i в состояние j за малое время Δt имеет вероятность $P_{ij}=q_{ij}\Delta t+o(\Delta t)$, $i\neq j$, $P_{ii}=1+q_{ii}\Delta t+o(\Delta t)$, $i=j$.

Финальные вероятности состояний λ_i будут равны $\pi_1=\dfrac{q_{22}}{q_{11}+q_{22}}$, $\pi_2=\dfrac{q_{11}}{q_{11}+q_{22}}$.

Далее предполагается, что покупатели покупают товар независимо друг от друга. Объем покупки – случайная величина $z=\left[z_1,z_2,\ldots,z_n\right]^{\mathrm{T}}$ с многомерной плотностью распределения $\psi(z)$ и моментами $M\left\{z_i\right\}=a_i$, $M\left\{z_iz_j\right\}=r_{ij}$.

2. Асимптотическая плотность вероятностей спроса

Пусть $S(t)=\left[S_1(t),S_2(t),\ldots,S_n(t)\right]^{\mathrm{T}}$ – количество товара, проданного к моменту времени t, $\lambda(t)$ – значение интенсивности потока покупателей в момент времени t. Обозначим $P(S,k,t)ds=P\left\{S<S(t)\le S+ds,\lambda(t)=\lambda_k\right\}$. Пусть в момент времени t $S(t)=S$, $\lambda(t)=\lambda_k$. Пусть прошло время Δt. За время Δt могли произойти следующие события:

1. Покупки не производились, интенсивность потока покупателей не изменилась. Вероятность этого события равна $\left(1+q_{kk}\Delta t\right)\left(1-\lambda_k\Delta t\right)+o(\Delta t)$.

2. Произведена покупка величины x, интенсивность потока покупателей не изменилась. Вероятность этого события равна $\left(1+q_{kk}\Delta t\right)\lambda_k\Delta t\psi(x)dx+o(\Delta t)$.

3. Покупки не происходили, интенсивность потока покупателей изменилась с λ_k на λ_j. Вероятность этого события равна $\left(1-\lambda_k\Delta t\right)q_{kj}\Delta t$.

4. Происходили покупки и интенсивность потока изменилась. Вероятность этого события равна $o(\Delta t)$.

Используя формулу полной вероятности, получим

$$P(S,k,t+\Delta t)=\left(1-\lambda_k\Delta t+q_{kk}\Delta t\right)P(S,k,t)+\sum_{j\neq k}q_{jk}P(S,j,t)+\lambda_k\Delta t\int_0^S P(S-x,k,t)\psi(x)dx\,.$$

Переходя к пределу при $\Delta t\to 0$, будем иметь

$$\frac{\partial P(S,k,t)}{\partial t}=-\lambda_kP(S,k,t)+\sum_{j=1}^{2}q_{jk}P(S,j,t)+\lambda_k\int_0^S P(S-x,k,t)\psi(x)dx\,. \qquad (1)$$

Если начальные вероятности состояний λ_k совпадают с финальными вероятностями π_k, то имеют место начальные условия

$$P(S,k,0)=\pi_k\delta(S)\,. \qquad (2)$$

Плотность распределения спроса к моменту времени t: $P(S,t)=\sum_{k=1}^{2}P(S,k,t)$.

Обозначим $g_k(\omega,t)=g_k\left(\omega_1,\ldots,\omega_n,t\right)=\int_0^\infty e^{-\sum_{i=1}^{n}\omega_is_i}P(S,k,t)ds_1ds_2\ldots ds_n$ – частичные производящие функции процесса $S(t)$. Производящая функция процесса $S(t)$ будет,

очевидно равна $g(\omega,t)=g_1(\omega,t)+g_2(\omega,t)$. Применяя преобразование Лапласа к уравнениям (1), получим, что функции $g_k(\omega,t)$ удовлетворяют системе уравнений

$$\frac{\partial g_k(\omega,t)}{\partial t}=-\lambda_k g_k(\omega,t)+\sum_{j=1}^{2} q_{jk} g_k(\omega,t)+\lambda_k g_k(\omega,t)Y(\omega), \qquad (1)$$

где

$$Y(\omega)=\int_0^\infty \psi(s_1,s_2,\ldots,s_n)e^{-\sum_{i=1}^{n}\omega_i s_i} ds_1 ds_2 \ldots ds_n,\ g_k(0,t)=\int_0^\infty P(S,k,t)dS=P\{\lambda(t)=\lambda_k\}=\pi_k .$$

Обозначим $\mu_j(k,t)=M\{S_j(t)\,|\,\lambda(t)=\lambda_k\}$, $k=1,2$, $j=\overline{1,n}$ – частичные моменты первого порядка процесса $S(t)=[S_1(t),S_2(t),...,S_n(t)]^{\mathrm{T}}$. Дифференцируя уравнения (3) и положив $\omega=0$, получим систему уравнений относительно $\mu_j(k,t)$: $\dot{\mu}_j(k,t)=\lambda_k\pi_k a_j+\sum_{i=1}^{2} q_{ik}\mu_j(k,t)$, $\mu_j(k,0)=0$, откуда

$$\mu_i(k,t)=\lambda_0\pi_k a_i t+\frac{(\lambda_0-\lambda_k)\pi_k a_i}{q_{11}+q_{22}}\left(1-e^{(q_{11}+q_{22})t}\right),$$

и безусловные моменты первого порядка

$$\mu_i(t)=\lambda_0 a_i t, \qquad (4)$$

где $\lambda_0=\lambda_1\pi_1+\lambda_2\pi_2$.

Обозначим

$$R_{i,j}(k,t)=\int_0^\infty S_i S_j P(S,k,t)dS \qquad (5)$$

– частичные моменты второго порядка процесса $S(t)=[S_1(t),S_2(t),...,S_n(t)]^{\mathrm{T}}$. Как следует из соотношений (3), величины $R_{i,j}(k,t)$ удовлетворяют системе уравнений

$$\frac{\partial R_{i,j}(k,t)}{\partial t}=\sum_{m=1}^{2} q_{m,k}R_{i,j}(m,t)+\lambda_k\mu_i(k,t)a_j+\lambda_k\mu_j(k,t)a_i+\lambda_k\pi_k r_{i,j},$$

откуда для безусловных моментов второго порядка $R_{ij}(t)$ получаем систему уравнений

$$\frac{\partial R_{i,j}(t)}{\partial t}=\lambda_0 r_{i,j}+\sum_{k=1}^{2}\left(\lambda_k\mu_i(k,t)a_j+\lambda_k\mu_j(k,t)a_i\right),$$

решение которой имеет вид

$$R_{i,j}(t)=\lambda_0 r_{i,j}t+a_i a_j\lambda_0{}^2 t^2-2a_i a_j\frac{\sigma_\lambda{}^2}{q_{11}+q_{22}}t+2a_i a_j\frac{\sigma_\lambda{}^2}{(q_{11}+q_{22})^2}\left(e^{(q_{11}+q_{22})t}-1\right),$$

где $\sigma_\lambda{}^2=\sum_k(\lambda_k-\lambda_0)^2\pi_k$.

Обозначим $V_k(\omega,u)=\int_0^\infty e^{-ut}g_k(\omega,t)dt$. Применяя к уравнению (3) преобразование Лапласа по t получим, учитывая (2), имеем

$$\sum_{j=1}^{2}\left[q_{jk}-\left(u+\lambda_k(1-Y(\omega))\delta_{jk}\right)V_j(\omega,u)\right]=-\pi_k, \qquad (6)$$

откуда функция $V(\omega,u)=V_1(\omega,u)+V_2(\omega,u)=\int_0^\infty g(\omega,t)e^{-ut}dt$ будет равна

$$V(\omega,u)=-\frac{q_{11}+q_{22}-u-(1-Y(\omega))(\lambda_2\pi_1+\lambda_1\pi_2)}{\det(\omega,u)},$$

где определитель матрицы системы уравнений (6)

$$\det(\omega,u)=$$

$$=u^2-u\left[q_{11}+q_{22}-\left(\lambda_1+\lambda_2\right)(1-Y(\omega))\right]-(1-Y(\omega))\left(q_{11}\lambda_2+q_{22}\lambda_1\right)+\lambda_1\lambda_2(1-Y(\omega))^2.$$

Функция $g(\omega,t)$ – производящая функция плотности распределения спроса к моменту времени *t* определится как обратное преобразование Лапласа

$$g(\omega,t)=\frac{1}{2\pi j}\int_{o-j\infty}^{\sigma+j\infty}V(\omega,u)e^{ut}du\,. \tag{7}$$

Однако вычислить интеграл (7) не удается даже в простейших случаях. Поэтому ниже находится асимптотическая плотность распределения спроса в предположении, что длительность торговой сессии $t>>1$.

Уравнение $\det(\omega,u)=0$ имеет корни

$$u_{1,2}=\frac{q_{11}+q_{22}-\left(\lambda_1-\lambda_2\right)(1-Y(\omega))\pm\sqrt{\left[\left(\lambda_1-\lambda_2\right)(1-Y(\omega))+\left(q_{22}-q_{11}\right)\right]^2+4q_{11}q_{22}}}{2}, \tag{8}$$

поэтому по теореме о вычетах [6] $g(\omega,t)=\frac{A_1\left(\omega,u_1(\omega)\right)}{u_1(\omega)-u_2(\omega)}e^{u_1(\omega)t}+\frac{A_2\left(\omega,u_2(\omega)\right)}{u_1(\omega)-u_2(\omega)}e^{u_2(\omega)t}$, где

$$A_1\left(\omega,u_1(\omega)\right)=-\left[q_{11}+q_{22}-u_1(\omega)-(1-Y(\omega))\left(\lambda_2\pi_1+\lambda_1\pi_2\right)\right],$$

$$A_2\left(\omega,u_2(\omega)\right)=q_{11}+q_{22}-u_2(\omega)-(1-Y(\omega))\left(\lambda_2\pi_1+\lambda_1\pi_2\right).$$

Раскладывая выражения (8) в ряд Тейлора по степеням ω, можно показать, что

$$u_1(\omega)=\ q_{11}+q_{22}+\sum_{i=1}^{n}\left[-a_i\frac{\left(q_{11}\lambda_1+q_{22}\lambda_2\right)}{q_{11}+q_{22}}\right]\omega_i+$$

$$+\frac{1}{2}\sum_{i,j=1}^{n}\left[\frac{2\pi_1\pi_2\left(\lambda_1-\lambda_2\right)^2a_ia_j}{q_{11}+q_{22}}+r_{ij}\frac{\left(\lambda_1q_{11}+\lambda_2q_{22}\right)}{q_{11}+q_{22}}\right]\omega_i\omega_j,$$

$$u_2(\omega)=-\sum_{i=1}^{n}a_i\lambda_0\omega_i+\frac{1}{2}\sum_{i,j=1}^{n}\left[\lambda_0r_{ij}-\frac{2\pi_1\pi_2\left(\lambda_1-\lambda_2\right)^2a_ia_j}{q_{11}+q_{22}}\right]\omega_i\omega_j.$$

Введем случайные величины $z_i=\frac{s_i(t)-\mu_i(t)}{\sigma_i(t)}$, где $\mu_i(t)$ определяются соотношениями (4) и

$$\sigma_i^2(t)=\lambda_0r_{ii}t-2\frac{a_i^2\sigma_\lambda^2}{q_{11}+q_{22}}t\,. \tag{9}$$

Их производящая функция

$$g_z(\omega,t)=e^{\sum_{i=1}^{n}\frac{\mu_1(t)\omega_i}{\sigma_i(t)}}g\left(\frac{\omega_i}{\sigma_i(t)},t\right). \tag{10}$$

Подставляя в (10) все необходимые выражения, заменив ω на *j*ω и переходя к пределу при $t\to\infty$, получим $g_z(\omega,t)\to e^{-\sum_{i,j=1}^{n}\frac{R_{ij}}{\sqrt{R_{ii}R_{jj}}}\omega_i\omega_j}$, где $R_{ij}=\lambda_0r_{ij}-2\frac{\sigma_\lambda^2}{q_{11}+q_{22}}a_ia_j$.

Таким образом, случайная величина *z* при $t>>1$ имеет асимптотически стандартное нормальное распределение и, следовательно, при $t>>1$ количество товара $S(t)$, проданного к моменту времени *t* имеет асимптотически многомерное нормальное распределение с вектором средних значений (5) и дисперсиями $\sigma_i^2(t)$ (9).

Пример. Для оценки скорости сходимости рассмотрим случай, когда продается два товара и плотность распределения $\psi(S_1, S_2) = \dfrac{1}{b_1 b_2}$, $0 \le S_1 \le b_1$, $0 \le S_2 \le b_2$.

На рис. 1 приведены зависимости $f(\omega) = \ln g_z(j\omega_1, j\omega_2, t)$, вычисленные по формуле (10) при фиксированном $\omega_2 = 0$ и $\omega_2 = 5$ для различных значений *t*. Параметры $q_{11} = -1$, $q_{22} = -0.5$, $b_1 = 4$, $b_2 = 7$, $\lambda_1 = 1.1$, $\lambda_2 = 0.6$.

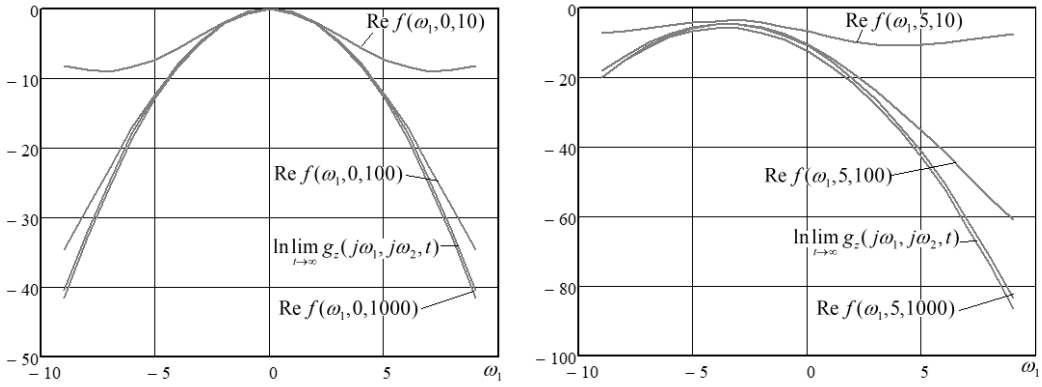


Рис. 1. Зависимость $f(\omega) = \ln g_z(j\omega_1, j\omega_2, t)$ от длительности *t* торговой сессии для различных значений ω_1, ω_2

На рис. 2 приведен трехмерный вид функции $f(\omega) = \ln g_z(j\omega_1, j\omega_2, t)$ для различных значений *t*.

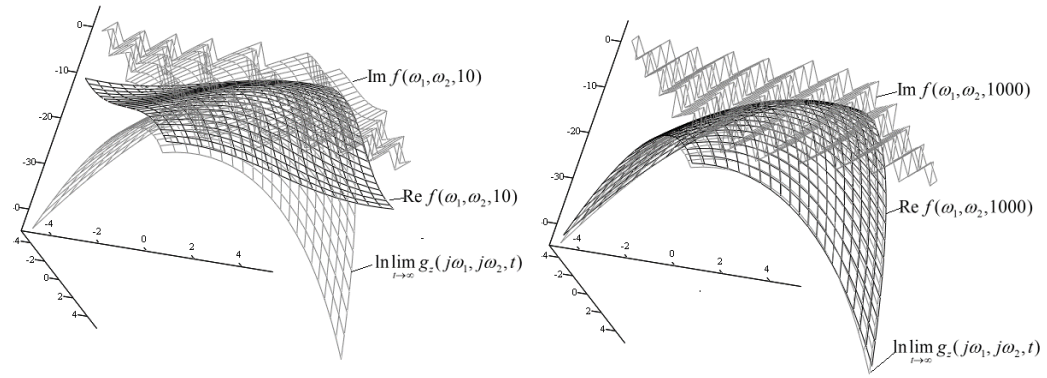


Рис. 2. Зависимость $f(\omega) = \ln g_z(j\omega_1, j\omega_2, t)$ от длительности *t*

В табл. 1 приведены значения абсолютной и относительной погрешностей аппроксимации $\mathrm{Re}\, f(\omega) = \mathrm{Re} \ln g_z(j\omega_1, j\omega_2, t)$ нормальным распределением f^*.

Таблица 1

Значения погрешностей аппроксимации

	$t = 10$	$t = 100$	$t = 1000$	$t = 10000$	$t = 100000$	$t = 1000000$
Абсолютная погрешность $\max\left\|f - f^*\right\|$	113.322	59.476	11.049	5.018	4.96	4.954
Относительная погрешность $\max\frac{\left\|f - f^*\right\|}{f}$	0.948	0.423	0.079	0.036	0.035	0.035

2. Определение средней прибыли

Обозначим через Wсреднюю прибыль продавца за время торговой сессии T. На приобретение ξ_i единиц i-го товара затрачивается $d_i\xi_i$ средств, а выручка составляет $c_i\xi_i$, если была продана вся партия i i-го товара и $c_i x - b_i\left(\xi_i - x\right)$, если партия размера x. Таким образом, средняя прибыль есть

$$W(\xi) = -\sum_{i=1}^{n} d_i\xi_i + \sum_{i=1}^{n} c_i\xi_i\int_{\xi_i}^{\infty} P_i(x,T)dx + \sum_{i=1}^{n}\int_0^{\xi_i}\left(c_i x - b_i\left(\xi_i - x\right)\right)P_i(x,T)dx\,, \qquad (11)$$

где $P_i(x,T)$ – плотность распределения спроса на i-й товар за время T. Если денежные средства, вкладываемые в приобретение товара неограниченны, то оптимальный размер партии товара определится из условия $\partial W(\xi)/\partial\xi_i = 0$. Дифференцируя (11) по ξ_i, получим

$$\int_0^{\xi_i} P_i(x,T)dx = 1 - \frac{b_i + d_i}{b_i + c_i}\,. \qquad (12)$$

Учитывая, что $d_i < c_i$, а интеграл монотонно возрастает по ξ_i в интервале от 0 до 1, получаем, что уравнение (12) имеет и причем единственное решение. При этом

$$W(\xi_{op}) = \sum_{i=1}^{n}\left(c_i + b_i\right)\int_0^{\xi_{i.op}} xP_i(x,T)dt\,.$$

Если $\chi_i(z)$ – функция обратная $\int_0^z P_i(S,T)dS$, то $\xi_{i,op} = \chi_i\left(1 - \frac{b_i + d_i}{b_i + c_i}\right)$.

При $T >> 1$ распределение $P_i(S,T)$ может быть аппроксимировано нормальным распределением. Поэтому при $T >> 1$ оптимальное значение партии товара $\xi_{i,op}$ определится соотношением $\xi_{i,op} = \mu_i(T) + \sigma_i(T)\mathrm{X}\left(1 - \frac{b_i + d_i}{b_i + c_i}\right)$, где $\mu_i(T)$ и $\sigma_i(T)$ определяются (4) и (9), а $\mathrm{X}(z)$ – функция, обратная к функции стандартного нормального распределения.

Пусть теперь количество средств на приобретение товара S_0 ограничено. Тогда функция $W(\xi)$ должна максимизироваться при дополнительных ограничениях

$$\xi_i \ge 0\,,\ \sum_{i=1}^{n} d_i\xi_i \le S_0\,. \qquad (13)$$

Получающуюся нелинейную оптимизационную задачу можно приближенно свести к задаче квадратичного программирования. Т.к. функция $W(\xi)$ сепарабельная, то ее можно представить в виде $W(\xi)=\sum_{i=1}^{n}W_i(\xi_i)$.

Раскладывая $W(\xi)$ в ряд Тейлора около точки ξ_{op}, получим, ограничившись первыми тремя членами разложения,

$$W(\xi)=W(\xi_{op})+\sum_{i=1}^{n}\frac{\partial W_i(\xi_{i,op})}{\partial\xi_i}(\xi_i-\xi_{i,op})+\frac{1}{2}\sum_{i=1}^{n}\frac{\partial^2 W_i(\xi_{i,op})}{\partial\xi_i^2}(\xi_i-\xi_{i,op})^2 .$$

Т.к. $\partial^2 W_i(\xi_{i,op})/\partial^2\xi_i=-(b_i+c_i)P_i(\xi_{i,\mathrm{op}},T)<0$, то исходную оптимизационную задачу можно приближенно заменить на задачу $\sum_{i=1}^{n}\frac{\partial^2 W_i(\xi_{i,op})}{\partial\xi_i^2}(\xi_i-\xi_{i,op})^2=\max$ при ограничениях (13).

Пример. Пусть имеется три товара, для которых: оптовая цена приобретения товаров – $d_1=5$, $d_2=2$, $d_3=1$; розничная цена продажи – $c_1=15$, $c_2=11$, $c_3=2$; плата за утилизацию $b_1=b_2=b_3=1$. Параметры $T=1000$, $\lambda_1=1.1$, $\lambda_2=0.6$, средние значения $a_1=2$, $a_2=3.5$, $a_3=2.5$, $q_{11}=-1$, $q_{22}=-0.5$. Тогда $\xi_{1,op}=1583$, $\xi_{2,op}=2857$, $\xi_{3,op}=1868$. Результаты решения оптимизационной задачи при различных S_0 приведены в табл. 2.

Таблица 1

Результаты решения оптимизационной задачи

	ξ_1	ξ_2	ξ_3
$S_0=1000$	0	475.689	48.622
$S_0=10\,000$	888.428	2125	1308
$S_0=100\,000$	1583	2857	1868

Как видно из табл. 2, при установленном ограничении S_0 наиболее выгодным во всех случаях считается товар ξ_2, у которого разница между ценой покупки и продажи в 5.5 раз.

Заключение

В работе рассмотрена задача розничной продажи многономенклатурной продукции, имеющей ограниченный срок годности, при условии, что моменты продаж продукции образуют ММР-поток с двумя состояниями. Найдена асимптотическая плотность распределения спроса на продукцию, определено оптимальное значение количества товара, предназначенного для перепродажи при условии, что продавец располагает ограниченными средствами для закупки товара. Отметим, что тем же путем могут быть исследованы и более сложные модели, например, представляет интерес исследование ситуации, когда продукция не только имеет ограниченный срок годности, но и постоянно портится в течение времени реализации. Модель реализации постоянно портящейся продукции рассматривалась, например, в [7,8].

ЛИТЕРАТУРА

1. *Petruzzi N.C., Dada M.* Pricing and the newsvendor problem: A review with extensions // Operations Research. – 1999.– V. 47(2). – P. 183–194.

2. *Chen X., Simchi-Levi D.* Joint Pricing and Inventory Management // Oxford handbook of pricing management, edited by Ö. Özer and R. Phillips. – 2012. – London: Oxford University Press.

3. *Turken N., Tan Y., Vakharia A.J., Wang L.* The Multi-Product Newsvendor Problem: Review, Extension, and Directions for Future Research // Handbook of Newsvendor Problems: Models, Extensions and Applications. – 2012. – V. 176. – P. 3–39.

4. *Kitaeva A., Livshits K., Ulyanova E.* The multiproduct Newboy Problem with Price- Depended Demand and Fast Moving Items // Information technologies and mathematical modeling: Proceeding of the 16 th International Scientific Conference, ITMM 2017 named after A.F. Terpugov. – CCIS. – V. 800. – Springer, 2017. – P. 297–311.

5. *Kitaeva A.V., Livshits K.I., Ulyanova E.S.* Estimating the demand parameters for single period problem, markov-modulated poisson demand, large lot size, and unobserved lost sale// IFAC-PapersOnline. – 2018. – V. 51. – № 11. – P. 882 –887.

6. *Лаврентьев М.А., Шабат Б.В.* Методы теории функций комплексного переменного. – М.: Наука, 1973. – 736 с.

7. *Livshits K., Ulyanova E.* Switch-Hysteresis Control Of the Production Process in a Model with Perishable Goods// Information technologies and mathematical modeling: Proceeding of the 15 th International Scientific Conference, ITMM 2016 named after A.F. Terpugov. – CCIS. – V. 638. – Springer, 2016. – P. 192–206.

8. *Livshits K., Ulyanova E.* Switch-hysteresis Control of the Selling Times Flow in a Model with Perishable Goods // Information technologies and mathematical modeling: Proceeding of the 14 th International Scientific Conference, ITMM 2015 named after A.F. Terpugov. – CCIS. – V. 564. – Springer, 2015. – P. 263–274.

ОПТИМАЛЬНОЕ РАСПРЕДЕЛЕНИЕ РЕСУРСОВ В ДВУХСЕКТОРНОЙ МОДЕЛИ ЭКОНОМИКИ С ГОСУДАРСТВЕННЫМ РЕГУЛИРОВАНИЕМ

К.О. Полуэктова, Т.И. Грекова
Томский государственный университет

Введение

Наиболее актуальными макроэкономическими задачами представляются задачи оптимального управления. В таких задачах требуется распределить имеющиеся ограниченные ресурсы (национальный доход, инвестиции, произведенный секторами экономики продукт) таким образом, чтобы обеспечить устойчивый экономический рост.

Использование государственного регулирования при распределении ограниченных ресурсов в экономике направлено на стабилизацию экономических процессов, адаптацию экономической системы к изменяющимся условиям. Выбор методов и способов воздействия государства на экономическую систему является актуальной задачей, а анализ влияния государственного регулирования – важной частью экономического анализа.

В данной работе рассматривается задача оптимального управления для двухсекторной модели экономики с учётом использования государственного бюджета.

Задачи анализа влияния налоговой политики на экономику были рассмотрены в работах [1–2]. Задачи оптимального управления в экономике рассматривались многими авторами как для односекторной модели экономики [3–4], так и для двухсекторной [5–12]. Основным отличием задачи оптимального управления, рассматриваемой в данной статье, от задач, рассматриваемых в работах [5–12], является наличие государственного регулирования: государство распределяет сформированный за счёт налогообложения бюджет в соответствии с потребностями секторов.

В [5–12] ВВП между секторами распределяется без учёта налоговых отчислений, трудовые ресурсы либо постоянны, либо не поставлена задача их распределения. В рассматриваемой работе предполагается, что трудовые ресурсы изменяются во времени и распределяются между секторами в соответствии с оптимальной нормой распределения трудовых ресурсов.

1. Постановка задачи

Рассматривается двухсекторная модель экономики на конечном интервале времени $[0,T]$, в которой первый сектор производит средства производства (основные средства), а второй – потребительские блага. В качестве ресурсов в каждом из секторов исполь-

зуются капитал $K_i(t)$, $i=1,2$ и труд $L_i(t)$, $i=1,2$, таким образом $K(t)=K_1(t)+K_2(t)$, $L(t)=L_1(t)+L_2(t)$.

Трудовые ресурсы изменяются в соответствии с экспоненциальным законом $L(t)=L_0e^{\lambda t}$, где λ – темп изменения трудовых ресурсов. Они распределяются между секторами в соответствии с нормой распределения трудовых ресурсов q.

Выпуск в каждом из секторов происходит в соответствии с производственными функциями неоклассического типа $F_i\left(K_i(t),L_i(t)\right)$, $i=1,2$.

Государственный бюджет формируется за счет сбора налогов с производственных секторов в соответствии с пропорциональной налоговой ставкой τ:

$$G_i(t)=\tau F_i\left(K_i(t),L_i(t)\right),\ i=1,2,\ G(t)=G_1(t)+G_2(t).$$

В рассматриваемой задаче предполагается, что государственный бюджет полностью используется для стимулирования производственных секторов. В каждый из секторов направляется доля государственного бюджета $G(t)$, которая определяется в соответствии с нормой распределения государственного бюджета m_1. Таким образом, валовой продукт каждого из секторов $Y_i(t)=F_i\left(K_i(t),L_i(t)\right)+m_iG(t)$, $i=1,2$, $m_2=1-m_1$, а общий валовой продукт $Y(t)=Y_1(t)+Y_2(t)$.

После уплаты налога валовой продукт первого сектора

$$\tilde{Y}_1(t)=(1-\tau)F_1\left(K_1(t),L_1(t)\right)+m_1G(t) \tag{1}$$

распределяется между первым и вторым секторами в соответствии с нормой распределения основных средств s: в первый сектор направляется величина $(1-\tau)s\tilde{Y}_1(t)$, во второй – $(1-\tau)(1-s)\tilde{Y}_1(t)$. Валовой продукт после уплаты налога второго сектора

$$\tilde{Y}_2(t)=(1-\tau)F_2\left(K_2(t),L_2(t)\right)+m_2G(t) \tag{2}$$

полностью направляется на потребление $C(t)$.

Изменение во времени основных фондов каждого из секторов описывается следующими дифференциальными уравнениями:

$$\begin{aligned}\dot{K}_1(t)&=s\tilde{Y}_1(t)-\mu K_1(t),\\ \dot{K}_2(t)&=(1-s)(1-\tau)\tilde{Y}_1(t)-\mu K_2(t).\end{aligned} \tag{3}$$

С учетом (1), (2) уравнения (3) принимают вид

$$\begin{aligned}\dot{K}_1(t)&=s\left[(1-\tau)F_1\left(K_1(t),L_1(t)\right)+m_1G(t)\right]-\mu K_1(t),\\ \dot{K}_2(t)&=(1-s)\left[(1-\tau)F_1\left(K_1(t),L_1(t)\right)+m_1G(t)\right]-\mu K_2(t),\end{aligned} \tag{4}$$

где μ – одинаковая для обоих секторов норма амортизации основных средств.

Естественная цель любой экономической системы – повысить благосостояние населения, т.е. максимизировать потребление на заданном временном интервале $[0,T]$:

$$J=\int_0^T C(t)e^{-\delta t}dt\to\max_{s,q}, \tag{5}$$

где δ – коэффициент дисконтирования.

Основная задача: найти такие норму распределения основных средств s^* и норму распределения трудовых ресурсов q^*, чтобы величина потребления $C(t)$ на временном интервале $[0,T]$ достигала максимального значения.

Введем удельные переменные: $y_i(t)=\dfrac{Y_i\left(K_i(t),L_i(t)\right)}{L_i(t)}$, $i=1,2$ – объем выпуска с учетом дотаций, приходящийся на единицу трудовых ресурсов; $f_i(t)=\dfrac{F_i\left(K_i(t),L_i(t)\right)}{L_i(t)}$,

$i = 1,2$ – средняя производительность труда; $k_i(t) = \dfrac{K_i(t)}{L_i(t)}$, $i = 1,2$ – фондовооруженность i-го сектора; $g_i(t) = \dfrac{m_i G(t)}{L_i(t)}$, $i = 1,2$ – объем государственных дотаций, направляемый в каждый из секторов, приходящийся на единицу трудовых ресурсов; $c(t) = \dfrac{C(t)}{L(t)}$ – объем потребления, приходящийся на единицу трудовых ресурсов.

С учетом введенных переменных изменение во времени фондовооруженности каждого сектора представляется дифференциальными уравнениями:

$$\dot{k}_1(t) = s(1-\tau) f_1\left(k_1\right) + s m_1 \tau \left[f_1\left(k_1\right) + \frac{(1-q)}{q} f_2\left(k_2\right) \right] - \nu k_1, \tag{6}$$

$$\dot{k}_2(t) = \frac{q(1-s)(1-\tau)}{(1-q)} f_1\left(k_1\right) + (1-s) m_1 \tau \left[\frac{q f_1\left(k_1\right)}{(1-q)} + f_2\left(k_2\right) \right] - \nu k_2, \tag{7}$$

$$k_1(0) = k_1^0,\ k_2(0) = k_2^0,\ k_1(T) = k_1^T,\ k_2(T) = k_1^T,\ k_1^T \geq 0,\ k_2^T \geq 0, \tag{8}$$

где введено обозначение $\nu = \mu + \lambda$, $\nu > 0$.

Таким образом, объем потребления, приходящийся на единицу трудовых ресурсов, определяется выражением

$$c(t) = \frac{C(t)}{L(t)} = \frac{\tilde{Y}_2(t)}{L(t)} = \frac{(1-\tau) F_2\left(K_2(t), L_2(t)\right) + G(t)\left(1-m_1\right)}{L(t)} =$$

$$= (1-\tau)(1-q) f_2\left(k_2\right) + \left(1-m_1\right)\tau \left[q f_1\left(k_1\right) + (1-q) f_2\left(k_2\right) \right]$$

и функционал (5) принимает вид

$$J = \int_0^T c(t) e^{-\delta t} dt = \int_0^T \left((1-\tau)(1-q) f_2\left(k_2\right) + \left(1-m_1\right)\tau \left[q f_1\left(k_1\right) + (1-q) f_2\left(k_2\right) \right] \right) e^{-\delta t} dt. \tag{9}$$

Итак, требуется найти такие значения нормы распределения основных средств $0 < s(t) < 1$ и нормы распределения трудовых ресурсов $0 < q(t) < 1$, которые при ограничениях (6)–(8) доставляли бы максимум функционалу (9). Значения m_1, τ определяются государственной политикой. Переменные $s(t)$ и $q(t)$ являются управляющими параметрами в задаче оптимального управления при определении максимума функционала (9).

2. Решение задачи оптимального распределения ресурсов

Задача оптимального распределения ресурсов с учетом государственного регулирования решается как задача оптимального управления с применением принципа максимума Понтрягина. Оптимальная траектория экономического роста для фиксированных значений налоговой ставки τ и нормы распределения государственного бюджета m_1 определяется из условия максимума функции Гамильтона

$$H\left(s, q, f_1, f_2, k_1, k_2, p_1, p_2\right) = (1-\tau)(1-q) f_2\left(k_2\right) + (1-m_1)\tau \left[q f_1\left(k_1\right) + (1-q) f_2\left(k_2\right) \right] +$$

$$+ p_1 \left(s(1-\tau) f_1\left(k_1\right) + s m_1 \tau \left[f_1\left(k_1\right) + \frac{(1-q)}{q} f_2\left(k_2\right) \right] - \nu k_1 \right) + p_2 \left(\frac{(1-s) q (1-\tau)}{(1-q)} f_1\left(k_1\right) + \right.$$

$$\left. + (1-s) m_1 \tau \left[\frac{q f_1(k_1)}{(1-q)} + f_2\left(k_2\right) \right] - \nu k_2 \right),$$

где $p_1(t) \geq 0$, $p_2(t) \geq 0$ – сопряженные переменные, для которых выполняются условия трансверсальности: $p_1(T) \cdot \left(k_1(T) - k_1^T\right) = 0$, $p_2(T) \cdot \left(k_2(T) - k_2^T\right) = 0$.

Т.к. функция Гамильтона линейно зависит от нормы распределения основных средств s и с учетом свойств производственной функции неоклассического типа получаем

$$s(t) = \arg\max_{0<s<1} \left\{ \begin{array}{l} p_1 \left(\left[(1-\tau) s f_1(k_1) \right] + m_1 \tau s \left[f_1(k_1) + \dfrac{(1-q)}{q} f_2(k_2) \right] \right) - \\ \qquad - p_2 \left(\dfrac{s(1-\tau) q f_1(k_1)}{(1-q)} + s m_1 \tau \left[\dfrac{q f_1(k_1)}{(1-q)} + f_2(k_2) \right] \right) \end{array} \right\}.$$

Таким образом, получаем условие, при котором начинает действовать особое управление s^*, которое является оптимальным значением нормы распределения основных средств.

В теории экономического роста доказано, что независимо от управляющего параметра s, экономика выходит на стационарную траекторию экономического роста [13]. Поставленная задача оптимального управления решается на стационарных траекториях, что означает $\dot{k}_1(t) = 0$, $\dot{k}_2(t) = 0$, $\dot{p}_1(t) = 0$, $\dot{p}_2(t) = 0$.

Необходимые условия максимума функции Гамильтона $\dfrac{\partial H}{\partial s} = 0$, $\dfrac{\partial H}{\partial q} = 0$ и необходимые условия для существования стационарного решения образуют систему уравнений:

$$s(1-\tau) f_1(k_1) + s m_1 \tau \left[f_1(k_1) + \frac{(1-q)}{q} f_2(k_2) \right] - \nu k_1 = 0, \tag{10}$$

$$\frac{q(1-s)(1-\tau)}{(1-q)} f_1(k_1) + (1-s) m_1 \tau \left[\frac{q f_1(k_1)}{(1-q)} + f_2(k_2) \right] - \nu k_2 = 0, \tag{11}$$

$$-f_1' \left[(1-m_1) \tau q + s p_1 \left[1 - \tau + m_1 \tau \right] + \frac{p_2 (1-s) q \left[1 - \tau + m_1 \tau \right]}{(1-q)} \right] + (\nu + \delta) p_1 = 0, \tag{12}$$

$$-f_2' \left[(1-q)(1-m_1\tau) + p_1 \frac{m_1 (1-q) \tau s}{q} + p_2 m_1 \tau (1-s) \right] + (\nu + \delta) p_2 = 0, \tag{13}$$

$$p_1 \left((1-\tau) f_1(k_1) + m_1 \tau \left[f_1(k_1) + \frac{(1-q)}{q} f_2(k_2) \right] \right) - \\ - p_2 \left[\frac{(1-\tau) q}{(1-q)} f_1(k_1) + m_1 \tau \left(\frac{q f_1(k_1)}{(1-q)} + f_2(k_2) \right) \right] = 0,$$

$$-(1-\tau) f_2(k_2) + (1-m_1) \tau \left[f_1(k_1) - f_2(k_2) \right] - \frac{p_1 s m_1 \tau}{q^2} f_2(k_2) + \\ + \frac{(1-s) p_2 f_1(k_1) (1 - \tau + m_1 \tau)}{(1-q)^2} = 0,$$

где f_1', f_2' – производные функций $f_1(k_1)$, $f_2(k_2)$ соответственно.

Будем называть магистралью траекторию экономического роста, для которой выполняются условия $k_1(t) = k_1^* = \text{const}$, $k_2(t) = k_2^* = \text{const}$.

Из уравнений (10) и (11) определяются оптимальные значения нормы распределения основных средств s^* и нормы распределения трудовых ресурсов q^*. Решение уравнений (12) и (13) существует, единственно в силу неоклассических условий [14] и определяет оптимальные значения фондовооруженности первого k_1^* и второго k_2^* секторов анализируемой экономики на траектории экономического роста.

3. Влияние государственного регулирования на решение задачи

Т.к. экономическая система при любой норме распределения основных средств выходит на стационарную траекторию экономического роста [13], то рассмотрим влияние государственного регулирования на магистраль.

Проведём анализ решения для наиболее часто используемой производственной функции Кобба-Дугласа, которая во многих практических ситуациях адекватно моделирует производственный процесс.

Пусть в течение некоторого достаточно продолжительного отрезка времени $t \in [0,50]$ налоговая ставка $\tau = 0.13$, норма распределения государственного бюджета $m_1 = 0.5$, технология производства не изменяется.

При начальных значениях фондовооруженностей секторов $k_1^0 = 3.2$, $k_2^0 = 2.9$ фондовооруженности первого и второго секторов увеличиваются до значений, соответствующих магистральному состоянию экономики k_1^*, k_2^* при оптимальных значениях нормы распределения основных средств и нормы распределения трудовых ресурсов $s^* = 0.237$, $q^* = 0.195$, соответствующих «золотому правилу накопления» [13].

Предположим, что для устранения дефицита государственного бюджета налоговая ставка увеличена до $\tau = 0.2$: траектории экономического роста переходят в новое стационарное состояние. Результаты моделирования приведены в табл. 1.

Таблица 1

Зависимость оптимальных значений и потребления от налоговой ставки

m_1	τ	k_1^*	k_2^*	s^*	q^*	J
0.5	0.13	3.92	3.064	0.237	0.195	43.694
0.5	0.20	3.928	2.814	0.199	0.151	41.003

Как видно из табл. 1, при налоговой ставке $\tau = 0.13$ большая часть основных средств, произведенных первым сектором, и большая часть трудовых ресурсов направляются во второй сектор. При увеличении налоговой ставки доли трудовых ресурсов и основных средств, направляемых во второй сектор, еще более увеличиваются. С увеличением доли трудовых ресурсов во втором секторе его оптимальный уровень фондовооруженности снижается, в то время как оптимальный уровень фондовооруженности первого сектора возрастает: такой эффект наблюдается в связи с увеличением оттока денежных средств из секторов в виде налогов и перераспределением трудовых ресурсов в секторах в соответствии с решением задачи оптимального распределения ресурсов. Эти факторы приводят к тому, что суммарное потребление J на отрезке времени $t \in [60,100]$ при увеличении налоговой ставки и приросте трудовых ресурсов во втором секторе уменьшается.

Предположим теперь, что налоговая ставка фиксирована $\tau = 0.13$, а норма распределения государственного бюджета $m_1 = 0.3$, т.е. в первый сектор направляется меньшая доля государственных дотаций. В некоторый момент времени государство перераспределяет часть своих дотаций в сектор производства средств производства для увеличения производственных мощностей, т.е. происходит шоковое увеличение нормы

распределения государственного бюджета до величины $m_1 = 0.8$. Результаты моделирования данной ситуации представлены в табл. 2.

Таблица 2

Зависимость оптимальных значений и потребления от нормы распределения государственного бюджета

m_1	τ	k_1^*	k_2^*	s^*	q^*	J
0.3	0.13	3.92	3.252	0.273	0.237	45.746
0.8	0.13	3.919	2.759	0.175	0.13	40.361

Как видно из данных в табл. 2, при $m_1 = 0.3$ (большая часть дотаций направляется во второй сектор) суммарное потребление J на временном отрезке $t \in [60,100]$ больше, чем при $m_1 = 0.8$ и большая часть основных средств и трудовых ресурсов направляются во второй производственный сектор. После перераспределения государственного бюджета в первый сектор оптимальные значение фондовооруженностей секторов уменьшаются (причем второго сектора – на 15%), во второй сектор направляется еще больше основных средств и трудовых ресурсов. Уменьшение дотаций во второй сектор, а также увеличение в секторе доли трудовых ресурсов является причиной снижения оптимального уровня фондовооруженности.

В результате анализа можно сделать вывод, что поставленная правительством цель стимулирования производства посредством фискальной и денежной политики в долгосрочной перспективе не приводит к желаемому результату. Возможное краткосрочное увеличение потребления при сокращении производственных ресурсов первого сектора в итоге ведёт к снижению благосостояния населения, определяемое объёмом производства второго сектора.

Заключение

В данной статье представлено решение задачи оптимального распределения ресурсов для двухсекторной модели рыночной экономики на конечном интервале времени с учетом государственного регулирования. В отличие от существующих работ, в которых рассматриваются задачи оптимального управления в экономике, в данной работе представлен анализ влияния государственного регулирования на выпуск производственных секторов и непроизводственное потребление.

Проведен анализ влияния налоговой политики и нормы распределения государственного бюджета на решение задачи оптимального распределения ресурсов. Полученные результаты показывают, каким образом государство при помощи налоговой политики и распределения дотаций, может регулировать выпуск секторов и, соответственно, уровень благосостояния населения.

При условии, что государственный бюджет полностью распределяется между секторами, установлено, что увеличение объема государственных дотаций, направляемых в первый сектор, приводит к уменьшению фондовооруженности второго сектора на магистрали и, следовательно, сокращению объёма производства потребительских благ и уменьшению уровня благосостояния населения. Показано, что государственное регулирование оказывает существенное влияние на норму распределения основных средств, и, следовательно, на уровень частных инвестиций в производство:

Таким образом, из рассматриваемых ситуаций можно сделать вывод, что любое вмешательство государства в экономику (изменение налоговой ставки или перераспределение государственного бюджета) оказывает существенное отрицательное влияние на экономическую систему в долгосрочной перспективе: норма распределения основных средств снижается, что ведет к вытеснению частных инвестиций, что, в свою очередь, тормозит развитие рыночных отношений в экономике и не является положительным фактором в теории экономического роста.

ЛИТЕРАТУРА

1. *Turnovsky S.J.* Fiscal policy and growth in a small open economy with elastic labour supply // The Canadian Journal of Economics. – 1999. – V. 32. –No 5. – P. 1191–1214.

2. *Been-Lon Chen.* Factor taxation and labor supply in a dynamic one-sector growth model // Journal of Economic Dynamics and Control. – 2007. – No 31 (12). – P. 3941–3964.

3. *Параев Ю.И., Грекова Т.И., Данилюк Е.Ю.* Аналитическое решение задачи оптимального управления односекторной экономикой на конечном интервале времени // Вестник Томского государственного университета. Управление, вычислительная техника и информатика. – 2011. – № 4 (17). – С. 5–15.

4. *Параев Ю.И., Грекова Т.И., Рощин А.В.* Решение терминальной задачи оптимального управления односекторной экономикой // Вестник Томского государственного университета. Управление, вычислительная техника и информатика. – 2012. – № 2 (19). – С. 14–19.

5. *Параев Ю.И.* Оптимальное управление двухсекторной экономикой // Вестник Томского государственного университета. Управление, вычислительная техника и информатика. – 2014. – № 3 (28). – С. 4–11.

6. *Kiselev Yu.N., Orlov M.V.* Optimal resource allocation program in a two-sector economic model with a Cobb-Douglas production function // Differential equations. – 2010. – V. 46. – No 12. – P. 1750–1766.

7. *Kiselev Yu.N., Orlov M.V.* Investigating a two-sector model of economic growth with the Cobb-Douglas production function // Moscow University Computational Mathematics and Cybernetics. – 2010. – V. 34. – No 2. – P. 66–73.

8. *Orlov M.V., Kiselev Yu.N., Orlov S.M.* Optimal control problem for two-sector economic model with Cobb–Douglas production function // 13th Viennese Workshop on Optimal Control and Dynamic Games. Research Unit ORCOS Vienna. – 2015. – P. 135.

9. *Kiselev Yu.N., Orlov M.V.* Optimal resource distribution program in a two-sector economic model with a Cobb-Douglas production function with distinct amortization factors // Differential equations. – 2012. – V. 48. – No 12. – P. 1607–1622.

10. *Kiselev Yu.N., Orlov M.V., Orlov S.M.* Optimal resource allocation program in a two-sector economic model with an integral type functional for various amortization factors // Differential equations. – 2015. – V. 51. – No 5. – P. 683–700.

11. *Грекова Т.И., Полуэктова К.О.* Оптимальное распределение ресурсов для двухсекторной модели экономического роста // Изв. вузов. Физика. – Т. 58. – № 11/2. – С. 264–268.

12. *Алонцева А.А.* Оптимальное распределение национального дохода в двухсекторной модели экономики // Материалы II Всероссийской молодежной научной конференции «Математическое и программное обеспечение информационных, технических и экономических систем». Томск, 16-17 мая 2014 г. – Томск : Изд-во Том. Ун-та. – 2014. – С. 167–170.

13. *Ромер Д.* Высшая макроэкономика. – М.: Издательский дом НИУ ВШЭ, 2015. – 855 с.

14. *Шараев Ю.В.* Теория экономического роста. М.: Издательский дом ГУ ВШЭ, 2006. – 254 с.

ПРОГРАММНОЕ ОБЕСПЕЧЕНИЕ ДЛЯ РЕШЕНИЯ ЗАДАЧИ ЭЛЕКТРОМАГНИТНОГО РАССЕЯНИЯ НА ТОНКИХ ПАРАЛЛЕЛЬНЫХ ИДЕАЛЬНО ПРОВОДЯЩЕМ И ДИЭЛЕКТРИЧЕСКОМ ЦИЛИНДРАХ

А.Г. Дмитренко, О.М. Балашова
Томский государственный университет

Введение

В работе [3] был представлен численный метод решения задачи электромагнитного рассеяния на структуре из двух тонких параллельных цилиндров, один из которых является диэлектрическим, а другой – идеально проводящим. В настоящей работе изложена основанная на этом методе программа расчета характеристик поля, рассеянного структурой, а также представлены результаты расчета этих характеристик для конкретной структуры.

1. Постановка задачи и её решение

Геометрия задачи показана на рис. 1. Будем рассматривать стационарную (зависимость от времени выбрана в виде $e^{-i\omega t}$) задачу дифракции электромагнитного поля $\vec{E}_0, \vec{H}_0$ на структуре, состоящей из двух тонких прямолинейных параллельных цилиндров, один из которых является диэлектрическим, а другой – идеально проводящим. Диэлектрический цилиндр имеет длину l_d, радиус r_d и характеризуется электродинамическими параметрами ε_i, μ_i. Идеально проводящий цилиндр имеет длину l_p, радиус r_p.

Для цилиндров выполняются условия $2r_d << \lambda$, $r_d << l_d$, $2r_p << \lambda$, $r_p << l_p$, где λ – длина падающей волны. Декартова система координат *Oxyz* выбрана таким образом, что её начало *O* совпадает с серединой осевой линии диэлектрического цилиндра, а ось *z* направлена вдоль осевой линии. Структура размещена в однородной среде D_e с электродинамическими параметрами ε_e,μ_e. Требуется найти рассеянное поле $\left\{\vec{E}_e, \vec{H}_e\right\}$ в среде D_e.

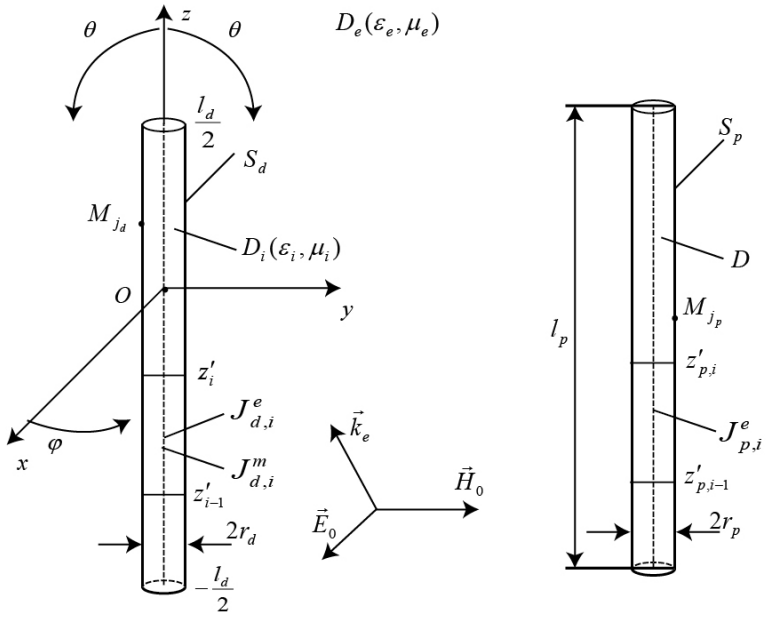


Рис. 1. Геометрия задачи

Решение сформулированной выше задачи, основанное на [1,2], получено в [3]. Оно имеет следующий вид:

$$E_{e,\theta}(M) = \sqrt{\frac{\mu_e}{\varepsilon_e}} H_{e,\varphi}(M) = \frac{e^{ik_e R}}{R} D_\theta(\theta,\varphi) + O\left(R^{-2}\right),$$

$$E_{e,\varphi}(M) = -\sqrt{\frac{\mu_e}{\varepsilon_e}} H_{e,\theta}(M) = \frac{e^{ik_e R}}{R} D_\varphi(\theta) + O\left(R^{-2}\right),$$

где $E_{e,\theta}(M), E_{e,\varphi}(M)$ и $H_{e,\theta}(M), H_{e,\varphi}(M)$ – компоненты векторов напряженности электрического и магнитного полей в сферической системе координат в произвольной точке *M* дальней зоны ($k_e R >> 1$), $k_e = \frac{2\pi}{\lambda}$ – волновое число для внешней среды D_e, *R* – расстояние от начала координат до точки *M*, $D_\theta(\theta,\varphi)$ и $D_\varphi(\theta)$ – компоненты диаграммы рассеяния, определяемые выражениями

$$D_\theta(\theta,\varphi) = -\frac{i\omega\mu_e}{4\pi}\sin\theta \times$$
$$\times\left[\sum_{i=1}^{N_d} J^e_{d,i}\int_{z'_{i-1}}^{z'_i} e^{-ik_e z'\cos\theta}dz' + e^{-ik_e(x_0\sin\theta\cos\varphi + y_0\sin\theta\sin\varphi)}\sum_{i=1}^{N_p} J^e_{p,i}\int_{z'_{p,i-1}}^{z'_{p,i}} e^{-ik_e z'_p\cos\theta}dz'_p\right], \quad (1)$$
$$D_\varphi(\theta) = \frac{ik_e}{4\pi}\sin\theta\sum_{i=1}^{N_d} J^m_{d,i}\int_{z'_{i-1}}^{z'_i} e^{-ik_e z'\cos\theta}dz'.$$

В (1) θ и φ – общепринятые сферические угловые координаты точки наблюдения M (см. рис. 1), $J^e_{d,i}$ и $J^m_{d,i}$ – элементы электрического и магнитного токов на оси диэлектрического цилиндра, $J^e_{p,i}$ – элементы электрического тока на оси идеально проводящего цилиндра, N_d – число элементов разбиения осевых токов диэлектрического цилиндра, N_p – число элементов разбиения осевого тока идеально проводящего цилиндра. Интегрирование в (1) выполняется вдоль i-го элемента разбиения, z'_{i-1} и z'_i – координаты начала и конца i-го элемента разбиения осевой линии диэлектрического цилиндра, $z'_{p,i-1}$ и $z'_{p,i}$ – координаты начала и конца i-го элемента разбиения осевой линии идеально проводящего цилиндра. Элементы токов $J^e_{d,i}$, $J^m_{d,i}$ $(i=\overline{1,N_d})$, $J^e_{p,i}$ $(i=\overline{1,N_p})$ определяются путём решения системы линейных алгебраических уравнений

$$\left[\vec{n}_i^{j_d},\left(\vec{E}_i^{j_d} - \vec{E}_e^{j_d}\right)\right] = \left[\vec{n}_i^{j_d},\vec{E}_0^{j_d}\right],\ \left[\vec{n}_i^{j_d},\left(\vec{H}_i^{j_d} - \vec{H}_e^{j_d}\right)\right] = \left[\vec{n}_i^{j_d},\vec{H}_0^{j_d}\right],\ j_d = \overline{1,L_d},$$
$$E^{j_p}_{e,z} = -E^{j_p}_{0,z},\ j_p = \overline{1,L_p}, \quad (2)$$

где $\vec{n}_i^{j_d}$ – единичный вектор нормали в точках коллокации M_{j_d} на поверхности диэлектрического цилиндра, $\vec{E}_e^{j_d},\vec{H}_e^{j_d}$, $\vec{E}_i^{j_d},\vec{H}_i^{j_d}$ и $\vec{E}_0^{j_d},\vec{H}_0^{j_d}$ – компоненты внешнего, внутреннего (для диэлектрического цилиндра) и возбуждающего полей в этих же точках, $\vec{E}^{j_p}_{e,z},\vec{E}^{j_p}_{0,z}$ – ориентированные вдоль оси z составляющие рассеянного и возбуждающего полей в точках коллокации M_{j_p} на поверхности идеально проводящего цилиндра.

Компоненты внешнего поля $\vec{E}_e^{j_d}$, $\vec{H}_e^{j_d}$, $\vec{E}^{j_p}_{e,z}$ определяются из выражений

$$\vec{E}_e(M) = \frac{i}{\omega\varepsilon_e}\operatorname{rotrot}\vec{\Pi}^e_d - \operatorname{rot}\vec{\Pi}^m_d + \frac{i}{\omega\varepsilon_e}\operatorname{rotrot}\vec{\Pi}^e_p,$$
$$\vec{H}_e(M) = \operatorname{rot}\vec{\Pi}^e_d + \frac{i}{\omega\mu_e}\operatorname{rotrot}\vec{\Pi}^m_d + \operatorname{rot}\vec{\Pi}^e_p,\ \vec{\Pi}^e_d = \vec{e}_z\sum_{i=1}^{N_d} J^e_{d,i}\int_{z'_{i-1}}^{z'_i}\frac{\exp\left(ik_e R_{MM'}\right)}{4\pi R_{MM'}}dz',$$
$$\vec{\Pi}^m_d = \vec{e}_z\sum_{i=1}^{N_d} J^m_{d,i}\int_{z'_{i-1}}^{z'_i}\frac{\exp\left(ik_e R_{MM'}\right)}{4\pi R_{MM'}}dz',\ \vec{\Pi}^e_p = \vec{e}_z\sum_{i=1}^{N_p} J^e_{p,i}\int_{z'_{p,i-1}}^{z'_{p,i}}\frac{\exp\left(ik_e \overline{R}_{MM'}\right)}{4\pi\overline{R}_{MM'}}dz'_p,$$

где $\vec{e}_z$ – единичный вектор, направленный вдоль оси z, $R_{MM'}$ – расстояние от точки z' на оси диэлектрического цилиндра до точки M в D_e, $\overline{R}_{MM'}$ – расстояние от точки z'_p на оси идеально проводящего цилиндра до точки M в D_e.

Компоненты поля $\vec{E}_i^{j_d},\vec{H}_i^{j_d}$ внутри диэлектрического цилиндра определяются из выражений

$$\vec{E}_i(M) = \frac{i}{\omega\varepsilon_i}\sum_{n=1}^{N_i}\operatorname{rotrot}\vec{\Pi}_{n,i},\ \vec{H}_i(M) = \sum_{n=1}^{N_i}\operatorname{rot}\vec{\Pi}_{n,i},$$

$$\vec{\Pi}_{n,i} = \Psi_i\left(M, M_{n,i}\right)\vec{p}_{\tau}^{\,n,i},\ \vec{p}_{\tau}^{\,n,i} = p_{\tau_1}^{n,i}\vec{e}_{\tau_1}^{\,n,i} + p_{\tau_2}^{n,i}\vec{e}_{\tau_2}^{\,n,i},\ M \in D_i,$$

в которых функция $\Psi_i\left(M, M_{n,i}\right) = \exp\left(ik_i R_{MM_{n,i}}\right)\Big/ 4\pi R_{MM_{n,i}}$, $R_{MM_{n,i}}$ – расстояние от точки $M_{n,i}$ на вспомогательной поверхности S_i (см. рис. 2) до точки M внутри диэлектрического цилиндра, $k_i = \omega\sqrt{\varepsilon_i\mu_i}$, N_i – число точек размещения пар вспомогательных диполей на S_i, $p_{\tau_1}^{n,i}$, $p_{\tau_2}^{n,i}$ ($n = \overline{1, N_i}$) – неизвестные дипольные моменты этих диполей, которые совместно с элементами токов $J_{d,i}^{e}$, $J_{d,i}^{m}$ ($i = \overline{1, N_d}$), $J_{p,i}^{e}$ ($i = \overline{1, N_p}$) определяются из системы (2).

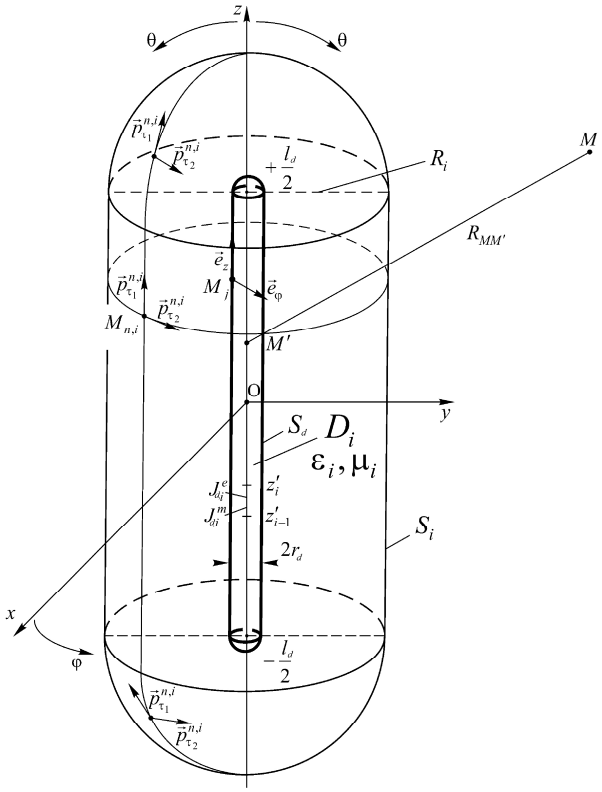


Рис. 2. Вспомогательная поверхность, охватывающая диэлектрический цилиндр

2. Описание компьютерной программы

Изложенные выше соотношения реализованы в виде программы для расчета компонент рассеянного поля и контроля точности получаемых результатов по критерию невязки граничных условий. Программа написана в среде Mathcad14, блок-схема программы представлена на рис. 3.

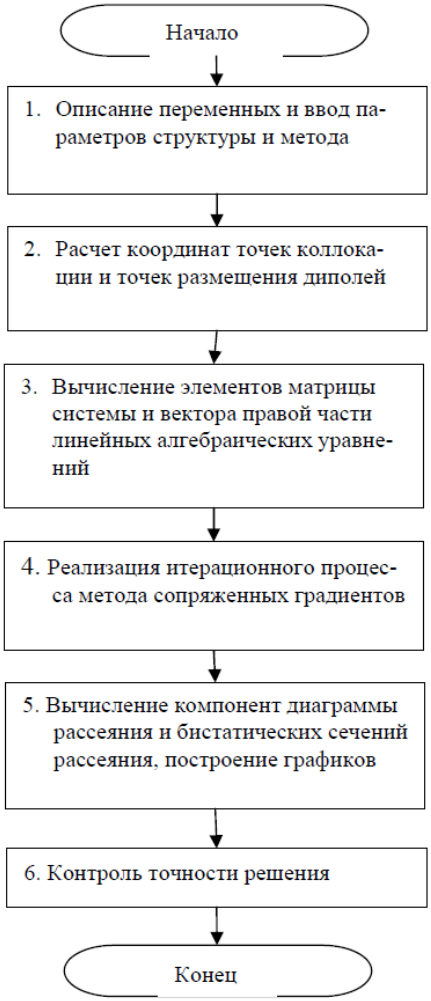


Рис. 3. Блок-схема компьютерной программы

Программа начинается с ввода в блоке 1 геометрических параметров структуры и параметров метода. К геометрическим параметрам относятся параметры l_p, r_p, и l_d, r_d, $\varepsilon_i' = \varepsilon_i / \varepsilon_e$, $\mu_i' = \mu_i / \mu_e$, характеризующие идеально проводящий и диэлектрический цилиндры, а также координаты x_0,y_0 осевой линии идеально проводящего цилиндра и координата z_0 середины его осевой линии. К параметрам метода относятся параметры N_d, N_p, R_i, $N_\varphi^{\text{цил}}$, N_z, $N_\varphi^{\text{сф}}$, $N_\theta^{\text{сф}}$, определяющие разбиение осевых линий цилиндров, количества точек коллокации на поверхностях цилиндров, выбор вспомогательной поверхности и точек размещения диполей на ней.

В блоке 2 выполняется расчёт координат точек коллокации и точек размещения диполей. В блоке 3 осуществляется вычисление элементов матрицы системы линейных алгебраических уравнений (2), а также вектора правой части системы.

В блоке 4 выполняется поиск решения системы линейных алгебраических уравнений (2) методом сопряженных градиентов. Алгоритм метода сопряженных градиентов следующий. Задается начальное приближение $\vec{p}^{(0)}$ и находится начальное значение вектора невязки $\vec{R}^{(0)} = A\vec{p}^{(0)} - \vec{f}$ и вектора направления $\vec{S}^{(0)} = -A^* \vec{R}^{(0)}$, где A – матрица системы (2), $\vec{p}$ – вектор, состоящий из неизвестных дипольных моментов и элементов

токов, A^* – матрица, транспонированная комплексно-сопряженная по отношению к матрице A. Далее выполняется следующая последовательность итераций:

$$t^{(k)} = \frac{\left\| A^* \vec{R}^{(k)} \right\|^2}{\left\| A\vec{S}^{(k)} \right\|^2},\ \vec{p}^{(k+1)} = \vec{p}^{(k)} + t^{(k)}\vec{S}^{(k)},\ \vec{R}^{(k+1)} = \vec{R}^{(k)} + t^{(k)} A\vec{S}^{(k)},$$

$$q^{(k)} = \frac{\left\| A^* \vec{R}^{(k+1)} \right\|^2}{\left\| A^* \vec{R}^{(k)} \right\|^2},\ \vec{S}^{(k+1)} = -A^*\vec{R}^{(k+1)} + q^{(k)}\vec{S}^{(k)},\ k = 0,1,2,\ldots.$$

Итерационный процесс останавливается при выполнении условия $\frac{\Phi^{(k)} - \Phi^{(k+1)}}{\Phi^{(k)}} < \varepsilon$, где $\Phi^{(k)} = \left\| A\vec{p}^{(k)} - f \right\|^2$, $\Phi^{(k+1)} = \left\| A\vec{p}^{(k+1)} - f \right\|^2$, ε – заданная точность (в данной программе $\varepsilon = 0.001$).

Блок 5 осуществляет вычисление компонент диаграммы рассеяния по формулам (1), а также бистатических сечений рассеяния по формуле

$$\frac{\sigma}{\lambda^2} = \frac{1}{\pi}\left(\left| D_\theta(\theta,\varphi) \right|^2 + \left| D_\varphi(\theta) \right|^2 \right). \qquad (3)$$

Результаты расчётов представляются на экране в виде таблиц и графиков. Задавая конкретный угол φ, при котором пользователя интересует рассеянное поле, вычисляют компоненты $D_\theta(\theta,\varphi)$ и $D_\varphi(\theta)$ при всех $\theta = \overline{0,\pi}$ с заданным шагом и строят график зависимости бистатического сечения рассеяния от θ.

Блок 8 осуществляет контроль точности решения задачи по величине относительной нормы невязки граничных условий

$$\Delta = \sqrt{\frac{\Phi'}{\Phi_0}}, \qquad (4)$$

где $\Phi' = \left\| A\vec{p} - f \right\|^2$, а $\Phi_0 = \left\| f \right\|^2$, на совокупности точек, промежуточных по отношению к точкам коллокации.

3. Численные результаты

Разработанная программа может быть использована для расчёта и анализа характеристик рассеяния широкого класса конкретных структур. Под характеристиками рассеяния понимаются компоненты диаграммы рассеяния (1), а также сечения рассеяния (3).

Ниже приводятся некоторые результаты для структуры, показанной на рис. 4.

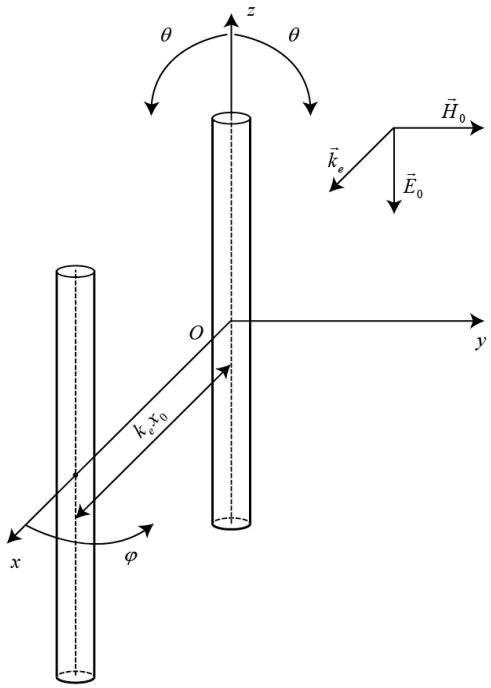


Рис. 4. Исследуемая структура

Структура состоит из диэлектрического цилиндра с осью, направленной вдоль оси *z*, и параллельного ему идеально проводящего цилиндра. Середины осевых линий цилиндров расположены на оси *x* на расстоянии $k_e x_0$ одна от другой. При выполнении вычислений, результаты которых представлены ниже, длины цилиндров $k_e l_d$ и $k_e l_p$ предполагались одинаковыми и равными 4.71 ($l_d = l_p = 0.75\lambda$), относительная магнитная проницаемость диэлектрического цилиндра предполагалась равной 1 (немагнитный цилиндр). Остальные параметры структуры менялись в зависимости от целей исследований. Во всех случаях предполагалось, что структура возбуждается плоской волной, падающей вдоль оси *x*, вектор $\vec{E}_0$ падающей волны ориентирован вдоль осей цилиндров. Параметры метода для диэлектрического цилиндра, определяющие положение вспомогательной поверхности, а также число и расположение диполей на ней во всех рассмотренных ниже случаях одинаковые. Радиус $k_e R_i$ вспомогательной поверхности выбран равным 2, число точек размещения диполей на вспомогательной поверхности равно 160. Эти точки размещены как на цилиндрической части вспомогательной поверхности (60 точек), так и на её сферических скруглениях (50 точек на каждом скруглении) и распределены следующим образом. Цилиндрическая часть рассечена десятью плоскостями $z = \text{const}$ ($N_z = 10$), отстоящими друг от друга на одинаковые расстояния Δz. В каждом сечении равномерно по азимутальному углу φ распределены шесть точек размещения диполей ($N_\varphi^{\text{цил}} = 6$). На каждом из сферических округлений в каждом из 5 полусечений $\varphi = \text{const}$ ($N_\varphi^{\text{сф}} = 5$), отстоящих друг от друга на угловое расстояние $\Delta\varphi = 72^\circ$, равномерно по углу θ выбраны 10 точек размещения диполей ($N_\theta^{\text{сф}} = 10$). Линия тока (осевая линия) диэлектрического цилиндра разбита на 30 участков ($N_d = 30$).

Число поперечных сечений $z = \text{const}$, в которых размещены точки коллокации на поверхности диэлектрического цилиндра, выбрано также равным 30. Эти сечения про-

ведены посередине каждого из участков разбиения осевой линии, в каждом сечении расположены 4 точки коллокации равномерно по азимутальному углу. Такой выбор параметров для диэлектрического цилиндра основан на результатах [4].

3.1. Исследование области применимости метода

Обычно под тонким цилиндром понимается цилиндрическое тело, поперечные размеры которого много меньше его длины и длины падающей волны. Однако такое определение является довольно расплывчатым. Цель данного раздела – уточнить значения величин радиусов цилиндров, при которых предлагаемый метод можно использовать, т.е. установить область применимости метода. Это можно сделать, наблюдая за поведением невязки граничных условий (4) при одновременном увеличении радиусов обоих цилиндров.

Такие исследования были проведены для описанной выше структуры; расстояние между цилиндрами $k_e x_0 = 1.884$ ($x_0 = 0.3\lambda$). Число элементов разбиения осевого тока идеально проводящего цилиндра выбрано равным 30. Результаты представлены на рис. 5. По оси абсцисс отложены значения радиусов цилиндров $k_e r_d = k_e r_p = k_e r$, по оси ординат – значения невязки граничных условий (4). Кривая 1 относится к случаю, когда относительная диэлектрическая проницаемость ε_i' диэлектрического цилиндра равна 4, а кривая 2 – к случаю, когда она равна 20.

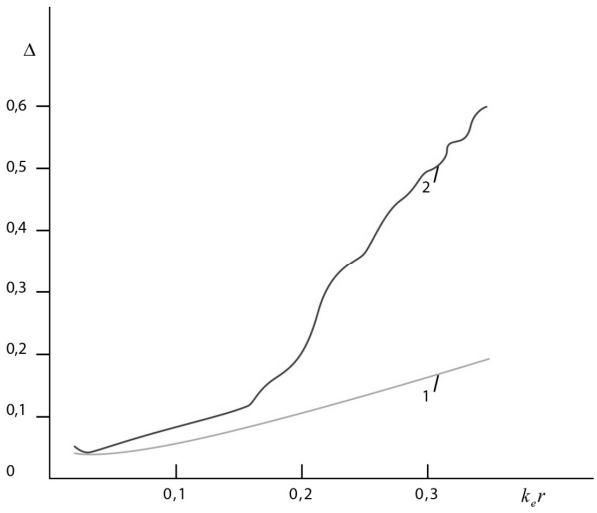


Рис. 5. Зависимость нормы невязки граничных условий от радиусов цилиндров при различных значениях относительной диэлектрической проницаемости диэлектрического цилиндра.
Кривая 1 – $\varepsilon_i' = 4$, кривая 2 – $\varepsilon_i' = 20$.

Как показывает рисунок, область применимости метода сужается с ростом относительной диэлектрической проницаемости диэлектрического цилиндра. Если считать приемлемым значение невязки, меньшее 0.1, то при $\varepsilon_i' = 4$ метод можно применять при $k_e r_d = k_e r_p < 0.2$, тогда как при $\varepsilon_i' = 20$ его можно применять только при $k_e r_d = k_e r_p < 0.13$.

3.2. Сравнение результатов расчета бистатических сечений рассеяния с результатами других авторов

Сравнение результатов расчёта, выполненных предлагаемым методом, с результатами других авторов имеет большое значение для оценки правильности метода и кор-

ректности работы программы. Поскольку решённая нами задача другими авторами ранее не решалась, мы не имеем возможности сравнить наши результаты с такими же результатами других авторов для рассматриваемой задачи. Однако при малых значениях относительной диэлектрической проницаемости диэлектрического цилиндра получаемые нами результаты должны быть очень близки соответствующим результатам для одиночного идеально проводящего цилиндра. Последняя задача является решённой, и результаты её решения содержатся, в частности, в [5]. Это обстоятельство и было использовано для контроля правильности получаемых нами результатов.

Результаты сравнения представлены на рис. 6. По оси абсцисс на этом рисунке отложен угол θ в градусах (см. рис. 4), по оси ординат – сечение рассеяния, нормированное на квадрат длины волны, определяемое выражением (3). Кривая 1 – сечение рассеяния, взятое из [5], кривые 2 и 3 – сечения рассеяния для рассматриваемой структуры при относительной диэлектрической проницаемости диэлектрического цилиндра $\varepsilon_i' = 4$ в полуплоскостях $\varphi = 0$ и $\varphi = \pi/2$ соответственно.

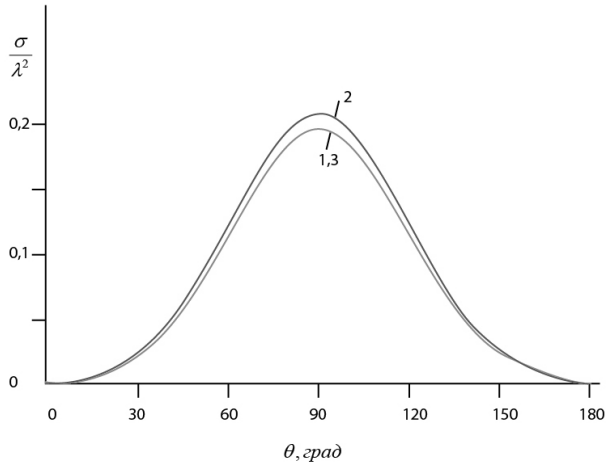


Рис. 6. Сравнение бистатических сечений рассеяния, полученных предлагаемым методом, с сечениями рассеяния, представленными в [5]. Кривая 1 – сечение рассеяния, взятое из [5], кривая 2 – сечение рассеяния в полуплоскости $\varphi = 0$, кривая 3 – сечение рассеяния в полуплоскости $\varphi = \pi/2$.

При получении кривых 2 и 3 радиус $k_e r_d$ диэлектрического цилиндра выбран равным 0.1, а радиус идеально проводящего цилиндра $k_e r_p$ выбран равным 0.03; расстояние между цилиндрами $k_e x_0 = 1.884$ ($x_0 = 0.3\lambda$). Уменьшение радиуса идеально проводящего цилиндра обусловлено стремлением приблизиться к радиусу, который выбран в [5] (в [5] $k_e r_p = 0.03$).

При анализе результатов, представленных на рис. 6, необходимо иметь в виду, что для одиночного идеально проводящего цилиндра бистатические сечения рассеяния во всех полуплоскостях $\varphi = \text{const}$ одинаковы. Для нашей задачи они в общем случае должны быть различными в силу отсутствия у структуры осевой симметрии. Однако, как показывают приведенные на рис. 6 кривые, при $\varepsilon_i' = 4$, как и следовало ожидать, различия сечений рассеяния для разных полуплоскостях $\varphi = \text{const}$ очень малы. Близость полученных нами сечений рассеяния к сечениям рассеяния, представленным в [5], говорит как о правильности самого метода решения задачи, так и о правильности выбора параметров метода.

3.3. Исследование зависимости бистатических сечений рассеяния от величины относительной диэлектрической проницаемости диэлектрического цилиндра

При увеличении относительной диэлектрической проницаемости диэлектрического цилиндра отличия бистатических сечений рассеяния структуры из двух тонких параллельных цилиндров, один из которых является диэлектрическим, а другой – идеально проводящим, от бистатических сечений рассеяния одиночного тонкого идеально проводящего цилиндре становятся более существенными. Поскольку исследуемая структура не обладает осевой симметрией относительно оси *z*, следует ожидать, что бистатические сечения рассеяния в различных полуплоскостях $\varphi = \text{const}$ будут различными.

Цель данного раздела – изучить влияние относительной диэлектрической проницаемости диэлектрического цилиндра на бистатические сечения рассеяния структуры в различных полуплоскостях $\varphi = \text{const}$.

При проведении соответствующих исследований радиусы цилиндров выбраны одинаковыми: $k_e r_d = k_e r_p = 0.1$; расстояние между цилиндрами $k_e x_0 = 1.884$ ($x_0 = 0.3\lambda$). Полученные результаты представлены на рис. 7–9.

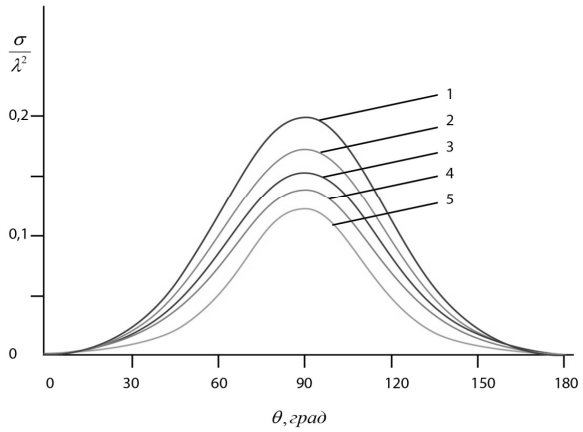


Рис. 7. Бистатические сечения рассеяния в полуплоскости $\varphi = 0$ при различных значениях относительной диэлектрической проницаемости диэлектрического цилиндра. Кривая 1 – $\varepsilon_i' = 4$, кривая 2 – $\varepsilon_i' = 10$, кривая 3 – $\varepsilon_i' = 15$, кривая 4 – $\varepsilon_i' = 20$, кривая 5 – $\varepsilon_i' = 30$.

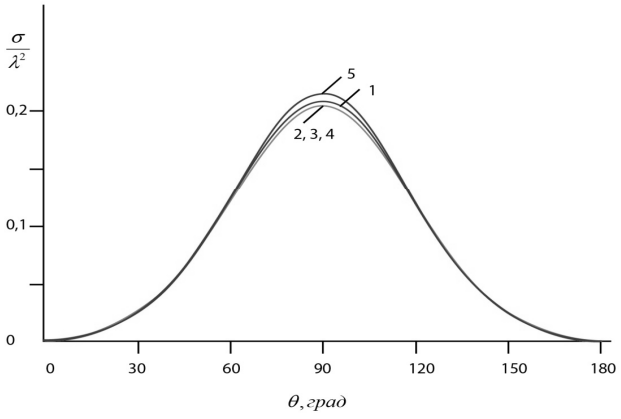


Рис. 8. Бистатические сечения рассеяния в полуплоскости $\varphi = \pi/2$

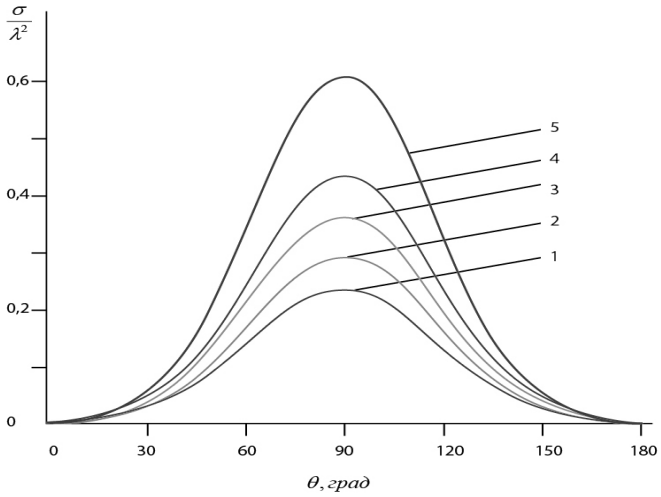


Рис. 9. То же, что на рис. 7, в полуплоскости $\varphi = \pi$

На рис. 7 приведены бистатические сечения рассеяния в полуплоскости $\varphi = 0$, на рис. 8 – в полуплоскости $\varphi = \pi/2$ и на рис. 9 – в полуплоскости $\varphi = \pi$. В полуплоскости $\varphi = 3\pi/2$ бистатические сечения рассеяния совпадают с сечениями рассеяния в полуплоскости $\varphi = \pi/2$ в силу симметрии структуры относительно плоскости *yz*. Кривые 1 на этих рисунках относятся к случаю, когда $\varepsilon_i' = 4$, кривые 2 – к случаю, когда $\varepsilon_i' = 10$, кривые 3 – к случаю, когда $\varepsilon_i' = 15$, кривые 4 – к случаю, когда $\varepsilon_i' = 20$ и кривые 5 – к случаю, когда $\varepsilon_i' = 30$.

Результаты, представленные на рис. 7–9, позволяют сделать следующие выводы. Величина относительной диэлектрической проницаемости диэлектрического цилиндра оказывает существенное влияние на бистатические сечения рассеяния. Характер этого влияния различен в различных полуплоскостях. В полуплоскости $\varphi = 0$, которая характеризует рассеяние в направлении падения плоской волны, наблюдается уменьшение сечений рассеяния с ростом величины относительной диэлектрической проницаемости, тогда как в полуплоскости $\varphi = \pi$, которая характеризует рассеяние в обратном направлении, наблюдается, наоборот, значительное увеличение сечений рассеяния. Неожиданным оказался тот факт, что в полуплоскостях $\varphi = \pi/2$ и $\varphi = 3\pi/2$, образующих плоскость, ортогональную плоскости поляризации падающей волны, бистатические сечения рассеяния очень слабо зависят от величины относительной диэлектрической проницаемости диэлектрического цилиндра и близки сечениям рассеяния одиночного идеально проводящего цилиндра.

Заключение

Таким образом, в данной работе приведено решение задачи электромагнитного рассеяния на структуре, состоящей из двух параллельных тонких цилиндров, один из которых является диэлектрическим, а другой – идеально проводящим, описана основанная на нём компьютерная программа. Созданная программа позволяет определять характеристики рассеяния широкого класса структур, отличающихся взаимным расположением цилиндров и их геометрическими и электродинамическими параметрами. Подробно исследована одна из возможных структур. Определена область применимости метода решения задачи, осуществлено сравнение полученных результатов с извест-

ными, исследована зависимость бистатических сечений рассеяния от значений относительной диэлектрической проницаемости диэлектрического цилиндра.

ЛИТЕРАТУРА

1. *Дмитренко А.Г., Гольцварт Е.П.* Решение задачи электромагнитного рассеяния на тонком диэлектрическом цилиндре методом вспомогательных источников // Радиотехника и электроника. – 2011. Т.56. – № 5. – С. 600–607.

2. *Дмитренко А.Г., Колчин В.А.* Численный метод исследования электромагнитного рассеяния структурами, содержащими тонкие проводники // Радиотехника и электроника. – 2003. Т. 48. – № 5. – С. 545–551.

3. *Дмитренко А.Г., Балашова О.М.* Численный метод решения задачи электромагнитного рассеяния на тонких параллельных идеально проводящем и диэлектрическом цилиндрах // Математическое и программное обеспечение информационных, технических и экономических систем: материалы VI Международной молодежной научной конференции. Томск, 24–26 мая 2018 г. – Томск: Издательский Дом Томского государственного университета, 2018. – С. 55–61.

4. *Дмитренко А.Г., Полин Е.П.* Исследование особенностей применения метода вспомогательных источников к решению задачи электромагнитного рассеяния на тонком диэлектрическом цилиндре // Известия вузов. Физика. –2015. – Т. 58. – № 4/2. – С. 15–21.

5. *Bowman J.J., Senior T.B.A., Uslengli P.L.E.* Electromagnetic and acoustic scattering by simple shapes // Amsterdam: North-Holland Publ. Company, 1969. – P. 488.

ПРИМЕНЕНИЕ ГАРМОНИЧЕСКОГО АНАЛИЗА ДЛЯ ИЗУЧЕНИЯ ХАРАКТЕРИСТИК СИСТЕМ ВИДЕНИЯ ЧЕРЕЗ АТМОСФЕРУ

А.И. Давыдова, И.Ю. Гендрина
Томский государственный университет

Введение

При решении климатических задач встаёт необходимость определения тенденций изменения различных характеристик. Прежде чем проводить анализ реальных данных, необходимо научиться устанавливать некоторые закономерности на экспериментальных данных, например, выявлять периодичности и локальные особенности сигнала. Главным средством анализа реальных физических процессов служит гармонический анализ, в частности, анализ Фурье, а также его модификация, называемая вейвлет-анализом.

Под спектральным анализом будем понимать разложение сигнала на его частотные или спектральные составляющие, а также оценку или измерение их характеристик.

Будем рассматривать исходный сигнал как физический процесс, способный передавать информацию и описанный математически с помощью функций определенного типа или полученных в результате натурных, лабораторных или численных экспериментов.

Разложение сложных периодических сигналов на элементарные гармонические составляющие называют анализом Фурье, а обратное конструирование сигналов по его гармоническим составляющим – синтезом Фурье. Математической же основой анализа реальных физических процессов служит преобразование Фурье. С его помощью произвольный процесс можно разложить на простые гармонические колебания с различными частотами.

Приблизительно два десятилетия назад в мире сформировалось и подверглось глубокому и тщательному изучению новое научное направление, называемое вейвлет-анализом, которое на сегодняшний день нашло широкое применение в обработке сигналов и изображений. Вейвлет-анализ вызвал большой интерес в теоретической и прикладной сферах математики, вследствие чего он нашёл применение в самых разных областях. В частности, он является хорошим инструментом для анализа климатических данных и их долгосрочного прогнозирования. Например, используя вейвлет-анализ, можно провести реконструкцию прошлого климата. С этой целью необходимо выявить характерные периоды, присутствующие в косвенных источниках изменения климата (природных архивах, включающих в свою структуру климатический сигнал). Таким

образом, выявляя периоды климатических колебаний, можно сделать прогноз о поведении климатических систем в будущем. В частности, вейвлет-анализ можно использовать для характеристики временной изменчивости облачности в широтных зонах [1,2].

1. Постановка задачи

Целью данной работы является применение гармонического анализа для обработки тестовых сигналов и, как следствие, получение алгоритмов для исследования характеристик систем видения через атмосферу. Под системой видения будем понимать системы, включающие поверхность Земли, её атмосферу и оптические устройства на верхней границе атмосферы. Для решения поставленной задачи был проведён анализ тестовых временных сигналов с использованием Фурье и вейвлет-анализа.

2. Методы решения

Анализ Фурье. Всякий исходный сигнал может быть представлен совокупностью гармонических составляющих (гармоник), т.е. рядом Фурье. Для того, чтобы сигнал был представлен более точно, необходимо, чтобы число гармоник было достаточно большим.

Исследуемые сигналы были представлены рядом Фурье с учётом их периодических особенностей:

$$S(t) = \frac{a_0}{2} + \sum_{k=1}^{N} \left(a_k \cos(\omega k t) + b_k \sin(\omega k t) \right), \quad (1)$$

где $\omega = 2\pi v$ – частота повторения в герцах, k – номер гармоники.

Под спектром функции $s(t)$ понимается совокупность её гармоник, образующих ряд Фурье [3].

Для каждого тестового сигнала был проведён спектральный анализ, заключающийся в нахождении коэффициентов ряда Фурье a_k и b_k:

$$a_k = \frac{2}{T} \int_0^T s(t) \cos(\omega k t)\, dt, \quad (2)$$

$$b_k = \frac{2}{T} \int_0^T s(t) \sin(\omega k t)\, dt. \quad (3)$$

Вейвлет-анализ. В узком смысле вейвлеты – это общее название для семейств математических функций, получающихся путём масштабирования и сдвига по оси времени одной материнской (базовой) функции. В широком смысле вейвлеты – это функции, которые локальны во времени и по частоте, среднее значение которых равно нулю [4,5].

Представление одномерного сигнала в виде обобщенного ряда или интеграла Фурье по системе базисных функций $\psi(t) = \frac{1}{\sqrt{a}} \psi\left(\frac{t-b}{a} \right)$, построенных на основе материнского (исходного) вейвлета $\psi(t)$, называется вейвлет-преобразованием.

Материнский вейвлет имеет полезные свойства благодаря операциям изменения временного масштаба a и сдвига во времени b. Множитель $1/\sqrt{a}$ обеспечивает независимость нормы этих функций от масштабирующего числа a [6].

Для получения вейвлет-спектрограмм был реализован алгоритм прямого вейвлет-преобразования исследуемых сигналов $s(t)$ на основе МНАТ-вейвлета («мексиканская шляпа») $\mathrm{МНАТ}(t) = \left(1 - t^2\right) e^{-t^2/2}$.

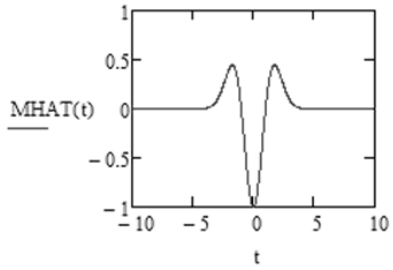


Рис. 1. MHAT-вейвлет («мексиканская шляпа»)

В ходе реализации были определены вейвлет-коэффициенты путём интегрирования скалярного произведения сигнала на вейвлет-функцию: $W(a,b)=\int_{-N}^{N}\psi(a,b,t)s(t)dt$, где вейвлет-функция $\psi(t)$ для заданных значений a и b представляет собой выражение вида $\psi(a,b,t)=\frac{1}{\sqrt{a}}\text{MHAT}\left(\frac{t-b}{a}\right)$.

Дискретные значения a и b заданы таким образом, чтобы представленные вейвлет-спектрограммы имели более детальное графическое представление [7].

3. Результаты моделирования

Пример. В качестве тестового сигнала была рассмотрена функция следующего вида: $s(t)=\sin(2t)+\sin(3t)$.

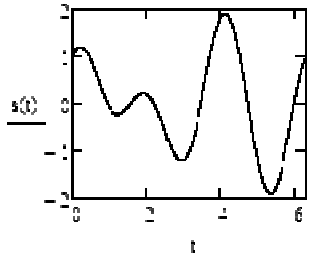


Рис. 2. Анализируемая функция

1. Рассмотрим случай, когда $T=2\pi$ – период и $\omega=\frac{2\pi}{T}$ – основная частота анализируемой функции. В ходе применения гармонического анализа, т.е. при разложении заданной на бесконечной прямой функции $s(t)$ в ряд Фурье, по формулам (2), (3) были получены следующие коэффициенты:

a_k =	b_k =
0	0
0	0
0	1
1	0

Далее был проведён гармонический синтез согласно (1) по $N=3$ гармоникам, т.е. получены колебания сложной формы путем суммирования их гармонических составляющих (гармоник).

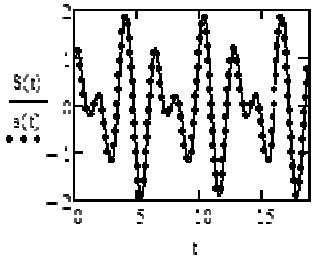
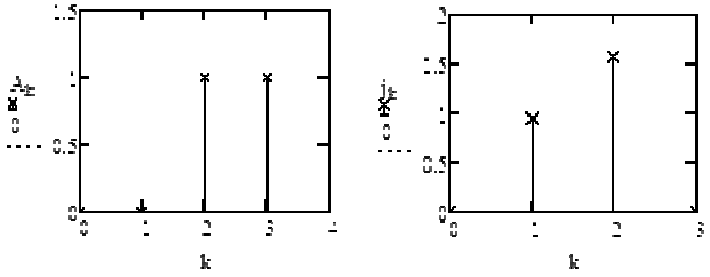


Рис. 3

Как видно из рис. 3 результаты синтеза полностью совпадают с графиком исходной функции. В этом случае, выявление истинного периода исходного сигнала не составит труда.

Наглядное представление о спектре дают спектральные диаграммы: амплитудная (рис. 4 а) и фазовая (рис. 4 б).



a) b)

Рис. 4. Амплитудная a) и фазовая b) спектральные диаграммы

2. Зададим период исследуемой функции $T = \frac{\pi}{2}$ – меньше, чем истинный период $T = 2\pi$, рассматриваемый в случае 1. С помощью аналогичных расчётов были получены следующие результаты:

a_k =
0.849
0.121
-0.015
$8.081 \cdot 10^{-3}$

b_k =
0
0.728
0.185
0.113

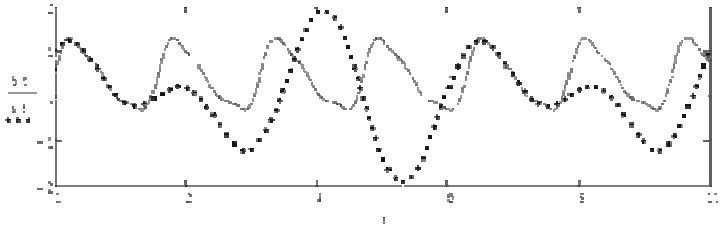
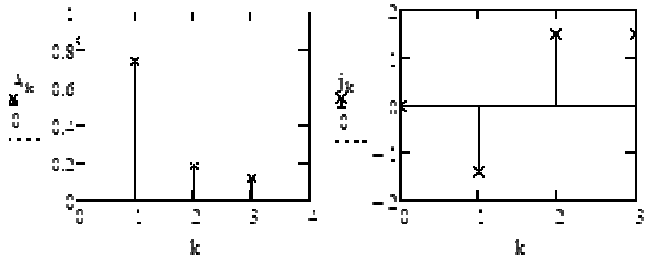


Рис. 5

Из рис. 5 видно, что результаты синтеза совпадают с графиком исходной функции только на промежутке $\left[0, \frac{\pi}{2}\right]$. В этом случае выявление истинного периода исходного сигнала невозможно.



a) b)

Рис. 6. Амплитудная a) и фазовая b) спектральные диаграммы

3. Пусть период исследуемой функции $T = \frac{5\pi}{2}$ – больше, чем истинный период $T = 2\pi$, рассматриваемый в случае 1. С помощью аналогичных расчётов по $N = 6$ гармоникам были получены следующие результаты:

0.17	0
0.212	-0.024
0.589	-0.063
-0.815	-0.189
0.453	0.657
0.024	0.146
$9.145 \cdot 10^{-4}$	0.087

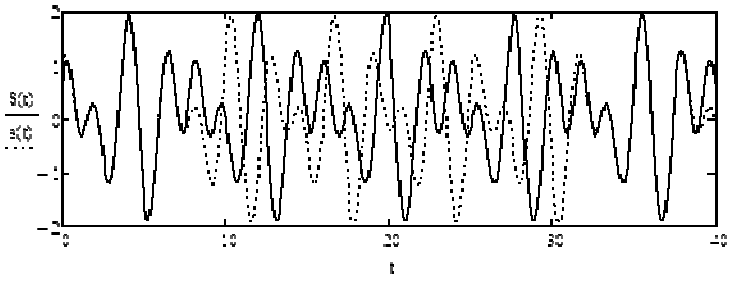
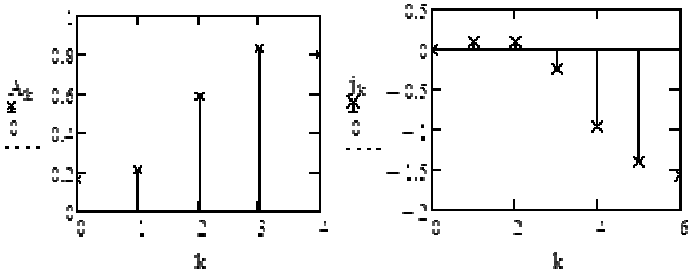


Рис. 7

Как видно из рис. 7, результаты синтеза совпадают с графиком исходной функции на промежутке $\left[0, \frac{5\pi}{2}\right]$. В этом случае выявление истинного периода исходного сигнала возможно.



a) b)

Рис. 8. Амплитудная a) и фазовая b) спектральные диаграммы

В ходе моделирования прямого вейвлет-преобразования были получены вейвлет-спектрограммы (диаграммы типа «уровень-время», а точнее – «уровень-индекс» для вейвлет-коэффициентов) и выведены в виде графика линий уровня, позволяющего проследить изменение интенсивности амплитуд вейвлет-преобразования на разных масштабах *a* и во времени *b*. Вдоль вертикальной оси отмечены пространственные масштабы *a*, а вдоль горизонтальной оси – локализация соответствующих компонент в пространстве *b*.

Легко убедиться в том, что исследуемый сигнал содержит достаточное количество перегибов при переходе производной через нуль в точках экстремумов. Вейвлет-спектрограммы хорошо отражают подобные особенности сигналов.

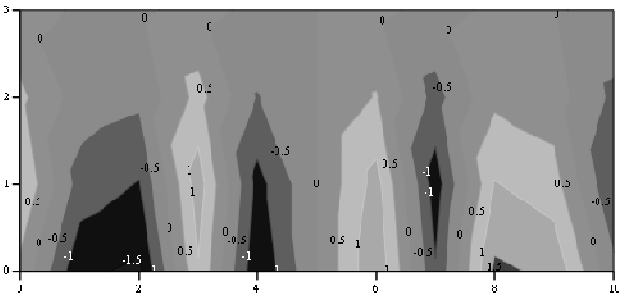


Рис. 9. Вейвлет-спектрограмма анализируемого сигнала

Середины тёмных областей на рис. 9 соответствуют локальным минимумам, а светлых – локальным максимумам исходной функции, т.е. чем резче выражена особенность сигнала, тем нагляднее она показана на спектрограмме и тем выше вейвлет-коэффициенты.

Проведя сравнение исходного сигнала и его вейвлет-спектограммы, можно сделать вывод о том, что на спектограмме хорошо отображена цикличность сигнала, а также его локальные и глобальные минимумы и максимумы.

Заключение

На основании проведенного исследования можно сделать вывод, что вейвлет-спектрограммы являются хорошим инструментом для обнаружения особенностей сигналов, поскольку они имеют более детальный вид по сравнению с Фурье-спектрограммами.

Большие уровни при вейвлет-анализе присущи только коэффициентам тех вейвлетов, которые находятся недалеко от различных локальных особенностей функции. Благодаря этому, вейвлет-спектрограммы способны легко обнаруживать локальные особенности функций с привязкой их ко времени или координатам пространства. В свою очередь, Фурье спектрограммы помогают обнаружить периодичность сигнала.

Уникальные математические свойства вейвлетов сделали их мощным и эффективным инструментом анализа и последующего синтеза различных сигналов, в том числе и сложных. Применения вейвлетов в различных областях, таких как медицина, астрофизика, криминалистика, телекоммуникации и т.п., многочисленны и приводят к важным практическим результатам.

Безусловно, преобразование Фурье является одним из самых известных инструментов для анализа сигналов, однако вейвлет-преобразование способно предоставить более гибкую технику их обработки. Существенным достоинством вейвлетов в сравнении с Фурье анализом является то, что они позволяют заметить хорошо локализованные изменения сигналов. Вследствие того, что вейвлет-преобразование хорошо анализирует нестационарные сигналы, оно превратилось в мощную альтернативу преобразованию Фурье.

ЛИТЕРАТУРА

1. *Нагорнов О.В., Никитаев В.Г., Простокишин В.М., Тюфлин С.А., Проничев А.Н., Бухарова Т.И., Чистов К.С., Кашафутдинов Р.З., Хоркин В.А.* Вейвлет-анализ в примерах. – М.: НИЯУ МИФИ, 2010. – 120 с.

2. *Чернокульский А.В.* Анализ глобального поля облачности и связанных с его вариациями климатических эффектов: дис. ... канд. физ.-мат наук: 25.00.29. – М., 2010. – 179 с.

3. *Латыпова Н.В., Тучинский Л.И.* Ряды Фурье: учеб.-метод. пособие. – Ижевск: Изд-во “Удмуртский университет”, 2011. – 80 с.

4. *Воскобойников Ю.Е.* Вейвлет-фильтрации сигналов и изображений (с примерами в пакете MathCAD). Новосибирск : НГАСУ (Сибстрин), 2015. – 188 с.

5. *Приоров А.Л., Волохов В.А., Апальков И.В.* Обработка сигналов на основе вейвлет-преобразования: методические указания. – Яросл. гос. ун-т им. П.Г. Демидова. Ярославль: ЯрГУ, 2011. – 44 с.

6. *Яковлев А.Н.* Введение в вейвлет-преобразования: Учеб. пособие. – Новосибирск: Изд-во НГТУ, 2003. – 104 С.

7. *Дьяконов В.П.* Вейвлеты. От теории к практике. 2-е изд. – М.: СОЛОН-Пресс, 2010. – 400 с.

II. СИСТЕМНЫЙ АНАЛИЗ И МАТЕМАТИЧЕСКОЕ МОДЕЛИРОВАНИЕ

ПОДХОД ЛИПОВЕЦКОГО К МОДИФИКАЦИИ МЕТОДА ЦЕНООБРАЗОВАНИЯ ШТОЦЕЛЯ

А.А. Андриевская, Ж.Н. Зенкова
Томский государственный университет

Введение

Вопрос о назначении цены – стратегическая задача каждого коммерческого предприятия, т.к. цена является базовым фактором при принятии потребителем решения о покупке. От конечной цены продажи напрямую зависит прибыль предприятия, а также уровень спроса на товары и услуги.

Любопытно отметить, что в современных учебниках по маркетингу очень мало информации о конкретных методах ценообразования, обычно в процессе выбора цены предприятие ориентируется на затраты и целевой уровень прибыли, при этом рассчитанная цена может показаться потребителю несправедливой, что весьма негативно сказывается на уровне продаж и итоговой прибыли. Ещё сложнее определить цену на товар-новинку [4,7], особенно если на рынке совсем нет аналогов, при этом у предприятия нет никакой информации о возможных уровнях цен, которые служат ориентиром в процессе ценообразования.

Для решения этой проблемы используются методы ценообразования, ориентированные на спрос, которые позволяют рассчитать предпочитаемый потребителями ценовой диапазон, например, метод ценообразования «Психологическая цена» Штоцеля [3]. Он достаточно прост, относительно дёшев и нашел широкое применение в зарубежной маркетинговой практике, однако редко применяется российскими маркетологами. В данной работе предложена модификация метода Штоцеля на основе подхода Стана Липовецкого [1,2], применившего модели логистической регрессии при модификации метода ценообразования PSM – Price Sensitivity Metter.

Метод апробировался на реальных данных о ценовых предпочтениях пользователей программного продукта. Произведено сравнение итоговых ценовых значений, полученных классическим методом Штоцеля, а также с помощью методики, предложенной С. Липовецким.

1. Метод ценообразования Штоцеля

Рассматриваемый метод был предложен французским психологом и социологом Жаном Штоцелем в середине XX-го века [3]. Этот метод очень прост и состоит в следующем. После того, как представители целевой аудитории пользовались товаром в течение некоторого периода времени, им предлагают ответить на следующие два вопроса:

1. Ниже какого уровня цены P_1 товар кажется вам настолько дешёвым, что возникает подозрение о его низком качестве и/или поддельности?

2. Начиная с какого уровня цены P_2 товар кажется настолько дорогим, что вопрос о покупке даже не стоит?

В результате получаем независимую двумерную выборку $\left(P_{i1}, P_{i2}\right)$, $i = \overline{1, N}$, где N – количество опрошенных потребителей. При этом цены P_j, $j = 1, 2$ рассматриваются как случайные величины (с.в.) с функциями распределения (ф.р.) $F^{(j)}(p) = P\left(P_j < p\right)$. Да-

лее каждый элемент выборки обрабатывается отдельно, по каждому вектору $\{P_{ij}\}$, $j = 1,2$ строится эмпирическая функция распределения (э.ф.р.) по формуле [5,6,9]:

$$F_N^{(j)}(p) = \frac{1}{N}\sum_{i=1}^{N} \mathrm{I}_{(0,p)}\left(P_{ij}\right), \tag{1}$$

где $\mathrm{I}_A(y) = \begin{cases} 0, & y \notin A, \\ 1, & y \in A \end{cases}$ – индикаторная функция. Э.ф.р. – кусочно-непрерывная функция, и в методе Ж. Штоцеля её сглаживают. Искомые цены $P_{\min}$ и $P_{\max}$ находят из уравнения $d(p) = F_N^{(1)}(p) - F_N^{(2)}(p) = 0.5$.

Пример реализации метода Штоцеля приведен на рис. 1.

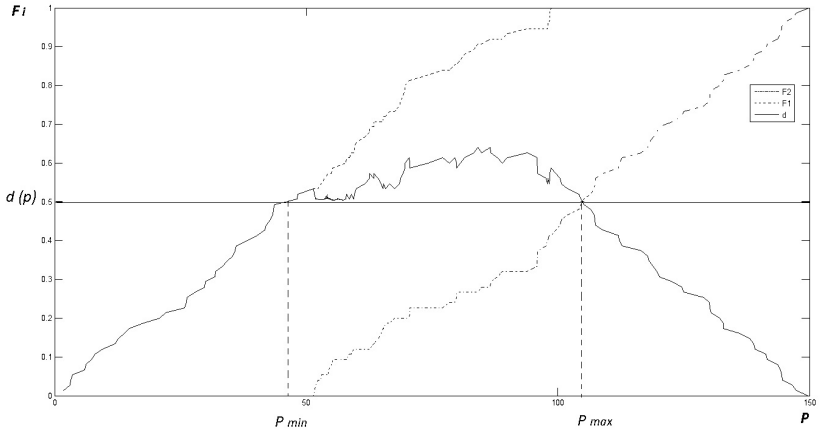


Рис 1. Пример реализации метода Штоцеля

Заметим, что назначать цену ниже $P_{\min}$ и больше $P_{\max}$ нецелесообразно – при слишком низкой цене основная масса потенциальных покупателей откажется от покупки, подозревая низкое качество, а при слишком высокой отказ от покупки произойдет в связи с неплатежеспособностью целевой аудитории. Если предприятие нацелено на широкий охват рынка и завоевание большей рыночной доли, то цену лучше выбрать ближе к $P_{\min}$, если же цель предприятия – завоевание репутации статусного товара, то цена должна быть высокой, ближе к $P_{\max}$.

2. Модификация метода Штоцеля с примененим подхода Липовецкого

Для модификации метода Штоцеля был использован подход, который применил С. Липовецкий при рассмотрении метода PSM – Price Sensitivity Metter. Суть метода PSM заключается в следующем. Респондентам предлагается ответить не на два, а на четыре основных вопроса:

1. Очень дёшево – Too Cheap (TCh): Ниже какого уровня цены X_1 товар кажется Вам настолько дешёвым, что начинают возникать сомнения в его качестве?

2. Дёшево – Cheap (Ch): Какая цена X_2 для Вас является приемлемой для покупки товара?

3. Дорого – Expensive (Ex): При какой цене X_3 Вы посчитаете, что товар стоит дороже, чем следует, но все же купите?

4. Очень дорого – Too Expensive (Tex): Начиная с какой цены X_4 товар кажется Вам слишком дорогим, настолько, что Вы не станете его покупать?

Как и в методе Штоцеля, цены P_j, $j=1,2$ рассматриваются как с.в. с ф.р. $F^{(j)}(p)=P\left(P_j<p\right)$. В результате из ответов формируется случайная выборка $\left(P_{i1},P_{i2},P_{i3},P_{i4}\right)$, $i=\overline{1,N}$, i-й элемент которой представляет собой четыре уровня цен, указанных i-м респондентом. По каждому вектору $\left\{P_{ij}\right\}$, $j=\overline{1,4}$, строится э.ф.р. по (1) для $j=1$ и 3 – оценки функции выживания $S_N^{(j)}(p)=1-F_N^{(j)}(p)$, далее эти четыре функции, называемые кривыми ценовой чувствительности, отображаются на одном графике.

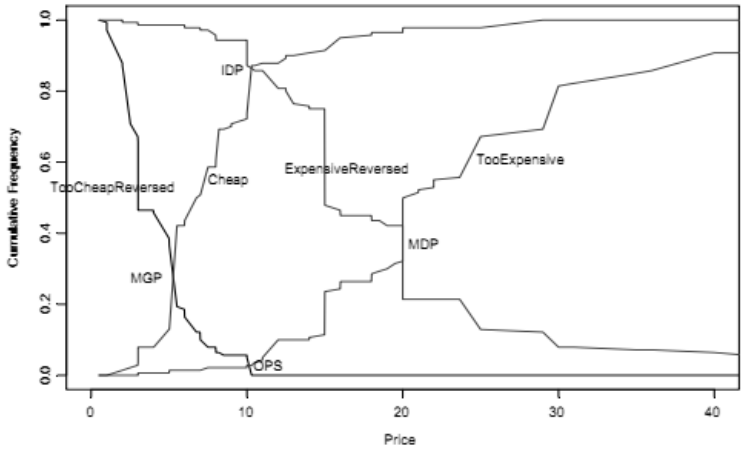


Рис. 2. Пример реализации метода PSM [1]

Обычно абсциссы точек пересечения интерпретируются на качественном уровне. Значение точки пересечения кривой «очень дёшево» и «дёшево» является нижней границей приемлемого диапазона цен, её называют точкой предельной дешевизны (MGP – Marginal Cheapness Point). Абсцисса точки пересечения кривой «очень дорого» и «недорого» является верхней границей приемлемого диапазона цен, её называют точкой предельной дороговизны (MDP – Marginal Expensiveness Point). Точка пересечения кривых «дорого» и «дёшево» называется точкой безразличия (IDP – Indifference Price Point). Одинаковое число респондентов считают цену, соответствующую данной точке, «дорогой» или «дешёвой». Точка пересечения кривой «очень дорого» и «очень дёшево» называется точкой оптимальной цены (OPS – Optimal Price Point). Соответствующее значение цены в этой точке считается одинаковым числом респондентов выходящим за пределы их приемлемого ценового диапазона. В качестве рекомендуемого диапазона рассматривается отрезок от оптимального до ожидаемого уровня цены. В [10,11] изучены статистические свойства оценок уровня цен, показана их асимптотическая нормальность и состоятельность.

Отметим, что рассматриваемые ценовые значения (точки пересечения кривых ценовой чувствительности) определяются из уравнений, которые для удобства можно представить в следующем виде (обозначения предложены С. Липовецким):

$$F_{\text{TCh}}+F_{\text{Ch}}=1,\ F_{\text{Ex}}+F_{\text{Tex}}=1,\ F_{\text{Ch}}+F_{\text{Ex}}=1,\ F_{\text{TCh}}+F_{\text{Tex}}=1,$$

где $F_{\text{TCh}}=1-F^{(1)}(p)$, $F_{\text{Ch}}=F^{(2)}(p)$, $F_{\text{Ex}}=1-F^{(3)}(p)$, $F_{\text{Tex}}=F^{(4)}(p)$.

Чтобы получить более согласованное представление данных, можно использовать другой подход. Представим не четыре, а пять ценовых интервалов восприятия цены: слишком дёшево TCh (ниже кривой F_{Tch}), приемлемая цена – Bargain или Br (от TCh до Ch, или между F_{Tch} и F_{Ch}), допустимая цена – Ok (между F_{Ch} и F_{Ex}), дорого – Premium или Pr (между F_{Ex} и F_{TEx}), и очень дорого – Too Expensive или TEx (выше F_{TEx}). Доли

покупателей, приходящиеся на каждое состояние, выражаются через исходные доли согласно формулам

$$Q_{\mathrm{TCh}} = 1 - F_{\mathrm{TCh}},\ Q_{\mathrm{Br}} = F_{\mathrm{TCh}} - F_{\mathrm{Ch}},\ Q_{\mathrm{Ok}} = F_{\mathrm{Ch}} - F_{\mathrm{Ex}},\ Q_{\mathrm{Pr}} = F_{\mathrm{Ex}} - F_{\mathrm{TEx}},\ Q_{\mathrm{TEx}} = F_{\mathrm{TEx}}.$$

Каждая из этих величин зависит от цены p, при этом выполняется равенство

$$Q_{\mathrm{TCh}}(p) + Q_{\mathrm{Br}}(p) + Q_{\mathrm{Ok}}(p) + Q_{\mathrm{Pr}}(p) + Q_{\mathrm{TEx}}(p) = 1.$$

Отношения для четырёх исходных распределений можно выразить в виде

$$F_{\mathrm{Tch}} = 1 - Q_{\mathrm{Tch}} = Q_{\mathrm{Br}} + Q_{\mathrm{Ok}} + Q_{\mathrm{Pr}} + Q_{\mathrm{TEx}},\ F_{\mathrm{Ch}} = 1 - \left(Q_{\mathrm{TCh}} + Q_{\mathrm{Br}}\right) = Q_{\mathrm{Ok}} + Q_{\mathrm{Pr}} + Q_{\mathrm{TEx}},$$

$$F_{\mathrm{Ex}} = 1 - \left(Q_{\mathrm{Tch}} + Q_{\mathrm{Br}} + Q_{\mathrm{Ok}}\right) = Q_{\mathrm{Pr}} + Q_{\mathrm{TEx}},\ F_{\mathrm{TEx}} = 1 - \left(Q_{\mathrm{Tch}} + Q_{\mathrm{Br}} + Q_{\mathrm{Ok}} + Q_{\mathrm{Pr}}\right).$$

С. Липовецкий предложил рассмотреть пять ценовых диапазонов в виде дискретных состояний с непрерывной ценовой переменной. Модель может быть представлена в виде системы линейных дифференциальных уравнений Чепмена – Колмогорова. При повышении цены переходы из одного состояния в другое можно выразить следующим образом:

Too Cheap → Bargain → Ok → Premium → Too Expensive

Уравнения Чемпена – Колмогорова, выражающие вероятность нахождения пребывания в любом из этих состоянии:

$$\begin{gathered} Q'_{\mathrm{Tch}} = -\lambda_1 Q_{\mathrm{TCh}},\ Q'_{\mathrm{Br}} = \lambda_1 Q_{\mathrm{TCh}} - \lambda_2 Q_{\mathrm{Br}},\ Q'_{\mathrm{Ok}} = \lambda_2 Q_{\mathrm{Br}} - \lambda_3 Q_{\mathrm{Ok}}, \\ Q'_{\mathrm{Pr}} = \lambda_3 Q_{\mathrm{Ok}} - \lambda_4 Q_{\mathrm{Pr}},\ Q'_{\mathrm{Tch}} = \lambda_4 Q_{\mathrm{TCh}}. \end{gathered} \tag{2}$$

Параметры $\lambda_j,\ j = \overline{1,4}$ обозначают интенсивность перехода между состояниями. Штрих обозначает производную по цене. Элементы с положительным знаком в правой части уравнения определяют переход к последующему состоянию из предыдущего, а отрицательный знак определяет выход из одного состояния в другое в случае повышения цены. Начальным условием для (2) является $Q_{\mathrm{TCh}} = 1$, при этом другие вероятности равны нулю.

Решение данной системы уравнений было найдено в [1,2] в виде $\widehat{F}_j(p) = \dfrac{1}{1 + e^{-(a_j + b_j \ln(p))}}$, где a_j и $b_j,\ j = \overline{1,4}$ – коэффициенты линейной регрессии, описывающей зависимость

$$\ln\left(\frac{F_j(p)}{1 - F_j(p)}\right) = a_j + b_j \ln(p), \tag{3}$$

что в итоге позволяет найти сглаженное представление функции $d(p)$:

$$\hat{d}(p) = \frac{1}{1 + e^{-(a_1 + b_1 \ln(p))}} - \frac{1}{1 + e^{-(a_4 + b_4 \ln(p))}} = 0.5. \tag{4}$$

Таким образом, ценовые диапазоны могут быть найдены из (4) без применения сглаживания исходных э.ф.р.

3. Определение цены нового программного продукта

Предложенная модификация метода Штоцеля апробировалась на реальных данных о результатах опроса потребителей программного продукта на рынке B2B (см. табл. 1, рис. 3,4), его потребители – фирмы и организации, оказывающие медицинские услуги. Программа была предоставлена фирмам в пользование на период один месяц, после чего производился опрос согласно методике PSM [8].

Таблица 1

Результаты опроса пользователей программного продукта по методу Штоцеля, у.е./шт.

№	P_1	P_2	P_3	P_4	№	P_1	P_2	P_3	P_4
1	181	294	374	472	27	138	262	335	384
2	209	314	430	685	28	151	266	350	418
3	259	323	474	838	29	130	262	333	382
4	165	286	367	439	30	173	293	373	467
5	290	343	581	1309	31	193	306	399	560
6	125	255	310	365	32	165	288	367	454
7	148	266	341	398	33	155	274	354	421
8	108	236	294	359	34	66	175	254	301
9	183	295	377	489	35	85	199	274	318
10	198	308	408	611	36	225	320	438	690
11	238	322	439	773	37	126	256	313	374
12	152	271	350	419	38	159	277	359	428
13	196	307	403	504	39	78	186	260	309
14	163	284	364	436	40	53	145	204	257
15	129	257	315	379	41	205	313	424	643
16	157	276	355	422	42	162	280	362	432
17	186	301	391	526	43	301	361	655	1350
18	200	311	408	620	44	288	336	508	1152
19	167	291	368	462	45	206	313	427	660
20	95	201	277	322	46	73	186	256	302
21	103	202	279	324	47	116	253	308	363
22	184	300	381	522	48	270	327	474	869
24	60	173	253	290	50	109	251	302	360
25	144	264	339	390	51	108	222	293	352
26	104	220	283	327					

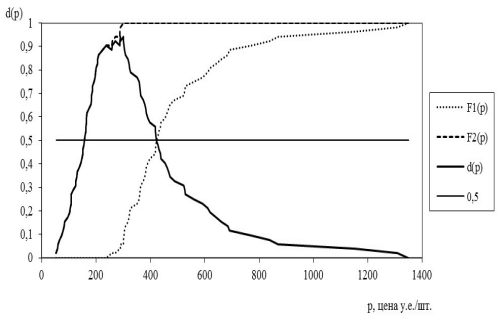


Рис 3. Реализация метода Штоцеля

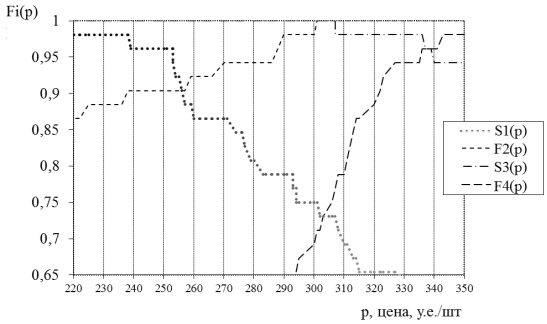


Рис 4. Реализация метода PSM

Далее, согласно методике С. Липовецкого, по реальным данным были построены модели линейной регрессии (3). В итоге из уравнения (4) получили, что $P_1 = 144.8$, $P_2 = 448.5$ у.е./шт.

Таким образом, можно сделать вывод, что устанавливаемая цена на товар должна лежать в интервале $p \in [144.8;448.5]$ у.е./шт., при этом полученный ценовой диапазон оказался наиболее широким из всех рассматриваемых.

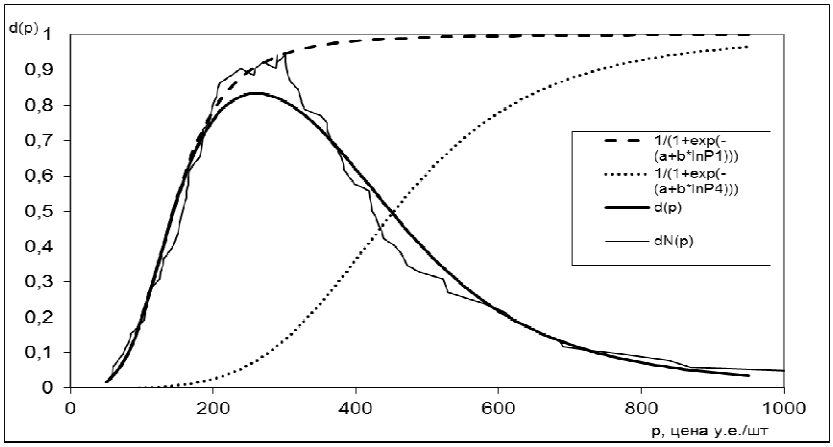


Рис. 5. Реализации модифицированного метода и метода Штоцеля по реальным данным

Заключение

В данной работе предложена модификация метода ценообразования Штоцеля с применением подхода С. Липовецкого к сглаживанию кривых ценовой чувствительности. Модифицированный метод использовался в процессе анализа ценовых предпочтений целевых потребителей нового программного продукта. Получен новый, более широкий ценовой диапазон, который может быть использован производителем товара-новинки как базовый при принятии окончательного решения о цене товара.

ЛИТЕРАТУРА

1. *Lipovetsky S.* Van Westendrop Price Sensitivity in Statistical Modeling // International Journal of Operations and Quantitative Management. – 2006. – Vol. 12. – № 2. – . P. 141–156.

2. *Lipovetsky S., Magnan S., Polzi A.Z.* Pricing Models in Marketing Research // Intelligent Information Management. – 2011. № 3.– P. 167–174.

3. *Stoetzel J.* Psychological/Sociological Aspects of Price // Bernard Taylor and Gordon Wills (Eds.). Pricing Research, Princeton, NJ: Brandon Systems Press. – 1970. P. 70–74.

4. *Zenkova Zh., Krakovetckaia I.* Marketing of New Bio-Energy Drinks/ Atiner conference paper series No: SME2013-0739. – 2013.

5. *Андриевская А.А., Зенкова Ж.Н.* Метод ценообразования Штоцеля для случайно цензурированных данных // Математическое и программное обеспечение информационных, технических и экономических систем: материалы VI Международной молодежной научной конференции. Томск, 24–26 мая 2018 г. – Томск: Издательский Дом Томского государственного университета, 2018. – С. 107–112.

6. *Боровков А.А.* Математическая статистика. – Новосибирск: Наука, 1997. – 772 с.

7. *Зенкова Ж.Н., Краковецкая И.В.* Непараметрическая оценка Тёрнбулла для интервально-цензурированных данных в маркетинговом исследовании спроса на биоэнергетические напитки // Вестник Томского государственного университета. Управление, вычислительная техника и информатика. – 2013. – № 3(24). – С. 64–69.

8. *Зенкова Ж.Н., Копнова Е.Е., Бараксанов Д.Н.* Применение метода PSM и алгоритма Тёрнбулла при определении цены нового программного продукта // Высокие технологии, фундаментальные исследования, инновации: сборник статей Семнадцатой международной научно-практической конференции "Фундаментальные и прикладные исследования, разработка и применение высоких технологий в промышленности и экономике". 22–23 мая 2014 года. Санкт-Петербург, Россия. – Санкт-Петербург: Изд-во Политехн. Ун-та Санкт-Петербурга, 2014. – С. 109–112.

9. *Зенкова Ж.Н., Цыбульникова Н.Р.* Исследование цены спроса на продукцию ЗАО «Натуральное мыло» методом Штоцеля по цензурированным данным // Математическое и программное обеспечение информационных, технических и экономических систем: материалы IV Международной молодежной научной конференции. Томск, 20–21 мая 2016 г. – Томск: Издательский Дом Томского государственного университета, 2016. – С. 62–65.

10. *Дмитриев Ю.Г., Зенкова Ж.Н., Зенков А.Г.* Статистические свойства оценок метода измерения ценовой чувствительности PSM // Новые информационные технологии в исследовании сложных структур. Материалы Двенадцатой конференции с международным участием. 4–8 июня 2018 г. – Томск: Издательский Дом Томского государственного университета, 2018. – С. 111–112.

11. *Дмитриев Ю.Г., Зенкова Ж.Н., Зенков А.Г.* Статистические свойства оценок метода измерения ценовой чувствительности PSM // Вестник Томского государственного университета. Управление, вычислительная техника и информатика. – 2019. – № 47. С. 30–40.

РАЗРАБОТКА ПРОГРАММНОЙ НАДСТРОЙКИ ДЛЯ ABC-XYZ-АНАЛИЗА ПО СЛУЧАЙНО ЦЕНЗУРИРОВАННЫМ ДАННЫМ В СРЕДЕ MS EXCEL

А.С. Охрименко
Томский государственный университет

Введение

Для анализа товарного ассортимента в зависимости от его прибыльности и стабильности спроса практически каждое производственное и торговое предприятие использует ABC-XYZ анализ [1–5], базирующийся на данных об осуществленных продажах за некоторый период. При этом очевидно, что данные, полученные после проведения ABC-XYZ-анализа, не анализируют и не отображают реального спроса на товар, т.к. не учитывают, например, такие ситуации, когда все имеющиеся в наличии изделия были проданы и некоторым покупателям товар просто не достался. В этом случае аналитики фактически работают с цензурированными справа данными, которые требуют особого подхода при их статистической обработке, обычные алгоритмы при этом приводят к серьезным неточностям [6–11], а значит, к прямым убыткам.

В данной работе рассматривается процедура разработки программной надстройки в среде MS Excel для автоматизации ABC-XYZ-анализа на случай присутствия в исходных данных случайного цензурирования справа. При оценивании среднего уровня и коэффициента вариации объемов продаж применялся метод подстановки [12], в качестве оценки функции распределения использовалась непараметрическая оценка Каплана-Мейера [7]. В работе также представлена поэтапная структура процесса выполнения и внешний вид интерфейса программы.

1. ABC-XYZ анализ для случайно цензурированных данных

Метод ABC-анализа базируется на принципе Парето и позволяет классифицировать товарный ассортимент на три группы в зависимости от степени важности товаров для предприятия, критерием классификации при этом может выступать прибыль или выручка от продажи M товаров R_i, $i=\overline{1,M}$ за рассматриваемый период [1–5].

Алгоритм классического ABC-анализа состоит из четырех этапов:

1. Исходные данные сортируются в порядке убывания: $R_1 \geq R_2 \geq \ldots \geq R_M$,

2. Для $i=\overline{1,M}$ вычисляется доля каждого показателя в общем итоге

$$d_i = \frac{R_i}{\sum_{i=1}^{M} R_i}.$$

3. Для $i=\overline{1,M}$ рассчитывается кумулятивная доля $S_i = S_{i-1} + d_i$, $S_0 = 0$.

4. Принятие решения:

– если $S_i \leq 0.8$, то товар принадлежит группе A – наиболее значимой для предприятия, дефицит товаров этой группы ведет к прямым убыткам, потому она требует к себе максимально пристального внимания;

– если $0.8 < S_i \leq 0.95$, то товар относят к менее значимой группе B, подходы к управлению которой менее затратны, чем к A;

– если $S_i > 0.95$, то это товар наименее значимой группы C, который может быть даже исключен из ассортимента.

Методология XYZ- анализа основана на коэффициенте вариации:

$$CV = \frac{S}{\overline{X}} \cdot 100\%, \qquad (1)$$

где

$$\overline{X} = \frac{1}{N} \cdot \sum_{i=1}^{N} X_i \tag{2}$$

– выборочное среднее, $X = (X_1, X_2, ..., X_N)$ – выборка, состоящая из объема продаж продукта, N – размер выборки,

$$S = \sqrt{\frac{1}{N} \cdot \sum_{i=1}^{N} \left(X_i - \overline{X}\right)^2} \tag{3}$$

– выборочное среднеквадратическое отклонение.

Коэффициент вариации (1) рассчитывается для каждого товара, и после этого весь ассортимент группируется следующим образом:

– если $CV \le 10\%$, то товар принадлежит группе X – группе со стабильным, практически постоянным спросом;

– если $10\% < CV \le 30\%$, то товар принадлежит группе Y, где спрос менее стабилен;

– если $CV > 30\%$, то товар принадлежит группе Z. Здесь часто оказывается спрос с ярко выраженной сезонностью.

Заметим, что границы групп Y и Z могут меняться, например, 30% могут быть изменены до 25% в зависимости от мнения экспертов.

Рассмотрим случайную величину (с.в.) $\xi \ge 0$ – уровень продаж товара с функций распределения (ф.р.) $F(x) = P\{\xi < x\}$, которая является неизвестной. Её оценка по полной выборке – эмпирическая функция распределения (э.ф.р) – позволяет путем метода подстановки [12] получить оценки (2) и (3).

Здесь рассмотрим случайно цензурированную справа выборку (ц.в.) $(X, I) = \{(X_1, I_1), .., (X_N, I_N)\}$ объема N, где для $i = \overline{1, N}$

$$I_i = \begin{cases} 0, & X_i - \text{полное наблюдение}, \\ 1, & X_i - \text{наблюдение до цензурирования}. \end{cases}$$

Тогда в качестве оценки ф.р. $F(x)$ можно использовать $F_N^C(x) = 1 - S_N^C(x)$, где

$$S_N^C(x) = \prod_{j=2,...,N:\, x \in (X_{(j-1)}; X_{(j)}]} \left(\frac{N - R_{(j)}}{N - R_{(j)} + 1} \right)^{1 - I_{(j)}}$$

– оценка функции выживания Каплана – Мейера, $X_{(1)} \le X_{(2)} \le ... \le X_{(N)}$ – упорядоченные значения ц.в., $R_{(j)}$ – ранг $X_{(j)}$, $j = \overline{1, N}$, $S_N^C(x) = 1$ для $x \le X_{(1)}$. Известно, что $S_N^C(x)$ – непараметрическая и состоятельная оценка функции выживания $S(x) = 1 - F(x)$ [7].

Методом подстановки найдем оценки математического ожидания и среднеквадратического отклонения для случайно цензурированной выборки:

$$\overline{X}^C = \int_0^{\infty} x dF_N^C(x) = \sum_{i=1}^{N} X_i \overline{I}_i h_i, \; S^C = \sqrt{\sum_{i-1}^{N} \left(X_i - \overline{X}^C\right)^2 \overline{I}_i h_i},$$

где для $j = \overline{2, N}$ $h_j = S_N^C\left(X_{(j-1)}\right) - S_N^C\left(X_{(j)}\right)$.

Оценка $\overline{X}^C$ используется при перерасчете итоговой выручки от каждого товара, для которого наблюдалось цензурирование, следующим образом: $R^C = N \cdot \overline{X}^C$, полученные значения выручки, рассматриваются как базовые для повторной группировки согласно описанному выше алгоритму ABC-анализа.

Для модификации XYZ-анализа используется коэффициент вариации, рассчитанный для случайно цензурированных данных по формуле $CV = \frac{S^C}{\overline{X}^C} \cdot 100\%$. Принцип группировки при этом остается прежним.

2. Разработка программной надстройки для ABC-XYZ анализ для случайно цензурированных данных в среде MS Excel

В рамках MS Excel программная надстройка реализуется через язык программирования VBA (Visual Basic for Applications). Язык VBA – это слегка упрощенная версия языка программирования Visual Basic, которую поддерживают приложения Microsoft Office, он позволяет автоматизировать задачи Excel, является мощным дополнением к приложению. Можно создать VBA-код, повторяющий действия пользователя, и таким образом автоматизировать простые задачи. VBA позволяет создавать формы для общения с пользователем, более поздние версии позволяют использовать элементы объектно-ориентированного программирования.

Общий дизайн надстройки включает в себя три листа MS Excel, на которых пользователь будет проводить модифицированный ABC-XYZ-анализ. Предполагается, что пользователь вводит исходные данные, начиная с ячейки A2 Листа 1. Данные должны располагаться в определенном порядке наименований, ограничений на форматирование текста нет. Основной функционал вызывается нажатием сочетания клавиш «Ctrl+r, Ctrl+q,Ctrl+w,Ctrl+e» на листе и представляет собой процедуру, а не функцию, т.е., не задает функциональных связей (не возвращает значения), и каждое обращение к процедуре может быть инициировано только нажатием сочетания указанных клавиш – автоматически пересчет не происходит.

Функционал представляет собой показ процессной логики ABC-XYZ-анализа. Используемые переменные объявляются в начале процедуры: *i* и *j* используются в циклах для обозначения строк и столбцов, соответственно; *n*, *m* – вспомогательные переменные. Логика процесса представлена на схеме (рис. 1).

На рис. 2, 3 представлены результаты применения программы к реальным данным о продажах крупного томского производственного предприятия за 2015 г. Из представленных таблиц можно увидеть существенные различия между группами классического и модифицированного ABC-XYZ-анализа. Множество товаров из группы *Y* перешли в группу *Z* .Это значит, что в срочном порядке следует пересмотреть систему заказов по данным позициям, т.к. при высоком товарообороте эти товары отличаются низкой прогнозируемостью объема продаж и попытка обеспечить гарантированное наличие всех товаров только за счет избыточного страхового запаса приведет к тому, что средний товарный запас компании значительно увеличится, при этом оборачиваемость может снизиться, и весьма значительно, что негативно скажется на прибыльности предприятия.

Заключение

Разработанный инструмент является программной надстройкой, которая позволяет автоматизировать проведение классического и модифицированного методов ABC-XYZ-анализа средствами MS Excel, что позволяет ему быть доступным для малых и средних предприятий. Автоматизация ускоряет процесс анализа, освобождая пользователя от необходимости вводить все процедуры, которые лежат в основе программы, вручную, а также обеспечивает высокую степень точности результатов, так как данная надстройка использует средства высококачественной и стабильной программной платформы MS Excel.

На реальных данных о продажах томского производственного предприятия было показано, что учет дефицита, т.е., цензурирования, приводит к существенным различи-

ям в итогах группировки. Новые методы следует рассматривать как более точные, т.к. используемая в расчетах среднего уровня продаж и среднеквадратического отклонения оценка Каплана-Мейера позволяет получить меньшие отклонения от истинных значений.

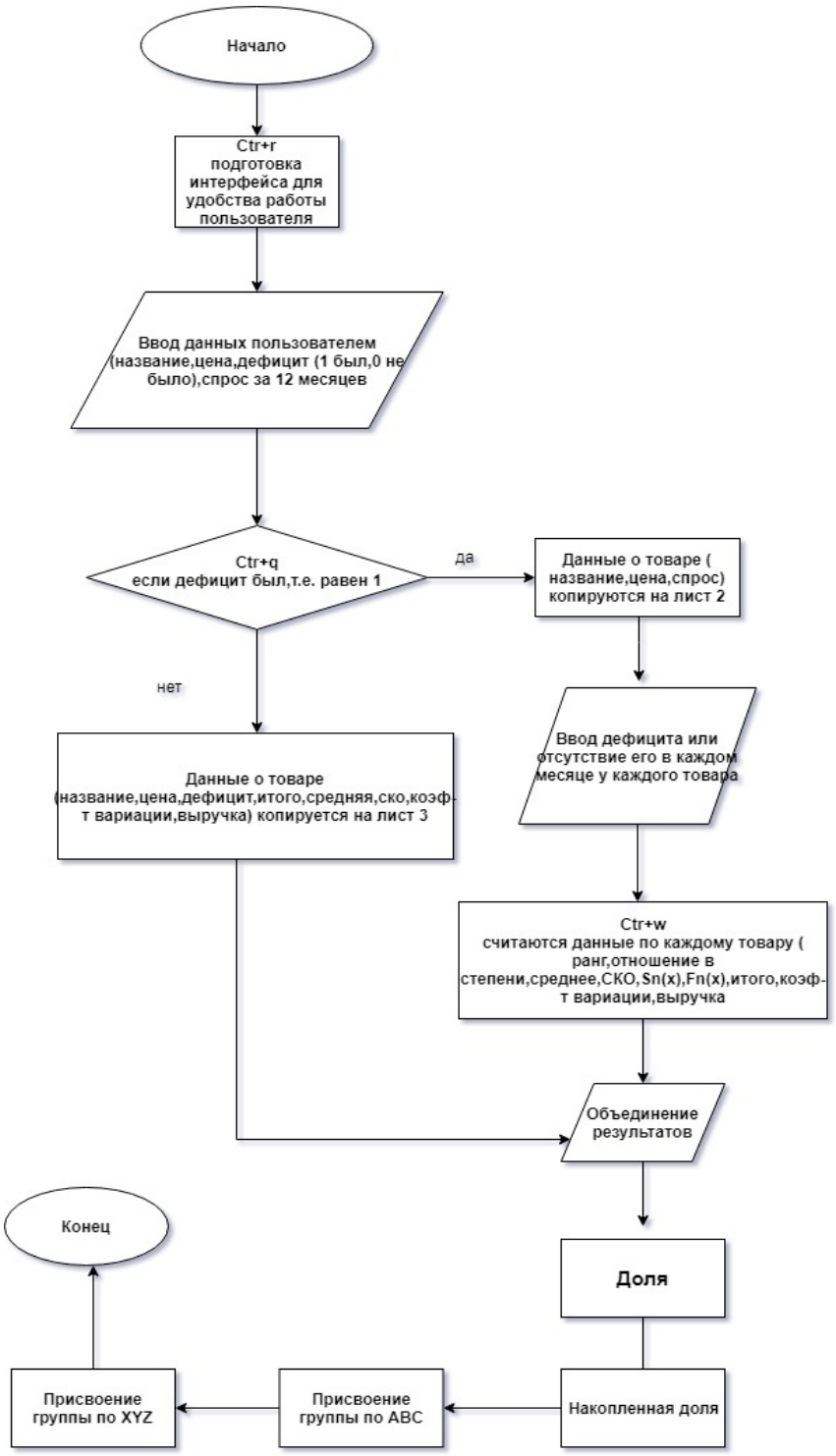


Рис. 1. Схема последовательности выполнения программы

	A	B	C	D	E	F	G	H	I	J	K	L
1	Наименование	Цена	Дефицит	Итого	Средняя продажа шт/мес	СКО	Коэф-т вариации	Выручка	Доля	Накопленная доля выручки в общем итоге	ABC	XYZ
2	*Товар 1*	42,00	1	35334,34	2944,528	455,7269	0,1547708	1484042	0,020913	0,91188107	B	Y
3	*Товар 2*	43,00	1	35230,78	2935,899	514,3486	0,1751929	1514924	0,02	0,89	B	Y
4	*Товар 3*	54,50	0	31363,7	2613,642	432,5883	0,1655117	1709321,7	0,02	0,85	B	Y
5	*Товар 4*	56,20	0	29748,75	2479,063	357,6018	0,1442488	1671880	0,02	0,87	B	Y
6	*Товар 5*	67,00	1	30008,9	2500,741	397,0271	0,1587637	2010596	0,03	0,80	A	Y
7	*Товар 6*	68,00	0	31253,56	2604,463	355,5811	0,1365276	2125242,0	0,03	0,77	A	Y
8	*Товар 7*	81,00	1	82709,25	6892,438	1133,342	0,1644327	6699449	0,09	0,67	A	Y
9	*Товар 8*	83,00	1	90668,33	7555,694	1386,271	0,1834736	7525471	0,11	0,57	A	Y
10	*Товар 9*	20,00	1	65998,93	5499,911	1128,586	0,2052008	1319979	0,018601	0,930482116	B	Y
11	*Товар 10*	21,00	1	18084,88	1507,073	323,0314	0,2143436	379782,4	0,005352	0,989683151	C	Y
12	*Товар 11*	21,00	1	10243,24	853,6032	208,5705	0,2443413	215108	0,003031	0,992714438	C	Y
13	*Товар 12*	43,50	1	116487,5	9707,295	1756,684	0,1809653	5067208	0,07	0,74	A	Y
14	*Товар 13*	45,50	1	27199,64	2266,637	447,1901	0,1972923	1237584	0,01744	0,947922059	B	Y
15	*Товар 14*	45,50	1	13235,77	1102,98	237,058	0,214925	602227,3	0,008487	0,984331284	C	Y
16	*Товар 15*	85,00	1	388895	32407,92	4291,365	0,1324172	33056077	0,47	0,47	A	Y
17	*Товар 16*	89,00	1	20733,1	1727,758	336,0461	0,1944983	1845246	0,03	0,82	B	Y
18	*Товар 17*	89,00	0	12321,35	1026,779	167,2182	0,162857	1096600	0,015453	0,963375273	C	Y
19	*Товар 18*	21,00	1	3058,541	254,8784	66,97068	0,2627554	64229,36	0,000905	0,99876736	C	Y
20	*Товар 19*	21,00	0	2389,858	199,1548	51,96292	0,2609173	50187,01	0,000707	0,999474592	C	Y
21	*Товар 20*	21,00	1	1775,445	147,9537	132,3278	0,8943866	37284,34	0,000525	1	C	Z
22	*Товар 21*	44,00	0	4269,395	355,7829	77,2477	0,2171203	187853,4	0,002647	0,995361654	C	Y
23	*Товар 22*	44,00	1	4032,918	336,0765	59,06462	0,1757475	177448,4	0,002501	0,997862244	C	Y
24	*Товар 23*	22,00	1	40221,17	3351,765	862,4679	0,2573176	884865,8	0,012469	0,97584474	C	Y
25								70962606				

Рис. 2 Результаты классического ABC-XYZ-анализа и числовые характеристики данных

	A	B	C	D	E	F	G	H	I	J	K	L
1	Наименование	Цена	Дефицит	Итого	Средняя продажа шт/мес	СКО	Коэф-т вариации	Выручка	Доля	Накопленная доля выручки в общем итоге	ABC	XYZ
2	*Товар 1*	42,00	1	39285,05	3273,754	3268,568	1,0015867	1649972	0,021763	0,90995748	B	Z
3	*Товар 2*	43,00	1	39109,43	3259,119	3257,525	1,0004893	1681706	0,022182	0,866142067	B	Z
4	*Товар 3*	54,50	0	31363,7	2613,642	432,5883	0,1655117	1709321,7	0,022546	0,843960279	B	Y
5	*Товар 4*	56,20	0	29748,75	2479,063	357,6018	0,1442488	1671880	0,022052	0,888194255	B	Y
6	*Товар 5*	67,00	1	33796,26	2816,355	2818,626	0,9991945	2264350	0,029867	0,766366256	A	Z
7	*Товар 6*	68,00	0	31253,56	2604,463	355,5811	0,1365276	2125242,0	0,028032	0,794398315	A	Y
8	*Товар 7*	81,00	1	92503,56	7708,63	7721,627	0,9983167	7492788	0,09883	0,662219966	A	Z
9	*Товар 8*	83,00	1	102185,8	8515,48	8573,504	0,9932322	8481419	0,11187	0,563389678	A	Z
10	*Товар 9*	20,00	1	73025,62	6085,469	6083,816	1,0002717	1460512	0,019264	0,929221719	B	Z
11	*Товар 10*	21,00	1	20164,59	1680,383	1684,854	0,9973464	423456,4	0,005585	0,989514758	C	Z
12	*Товар 11*	21,00	1	11446,26	953,8553	972,1347	0,9811966	240371,5	0,003171	0,992685272	C	Z
13	*Товар 12*	43,50	1	129459,1	10788,26	10820,42	0,9970279	5631470	0,074279	0,73649936	A	Z
14	*Товар 13*	45,50	1	30523,67	2543,639	2554,876	0,9956017	1388827	0,018319	0,94754042	B	Z
15	*Товар 14*	45,50	1	14698,93	1224,911	1227,208	0,9981284	668801,4	0,008822	0,983929345	C	Z
16	*Товар 15*	85,00	1	402727	33560,59	31815,65	1,0548452	34231799	0,451519	0,451519303	A	Z
17	*Товар 16*	89,00	1	23013,52	1917,794	1922,84	0,9973758	2048204	0,027016	0,821414232	B	Z
18	*Товар 17*	89,00	0	12321,35	1026,779	167,2182	0,162857	1096600	0,014464	0,962004639	C	Y
19	*Товар 18*	21,00	1	3422,776	285,2313	288,1314	0,9899349	71878,29	0,000948	0,998717511	C	Z
20	*Товар 19*	21,00	0	2389,858	199,1548	51,96292	0,2609173	50187,01	0,000662	0,999379481	C	Y
21	*Товар 20*	21,00	1	2240,214	186,6845	225,1681	0,8290892	47044,48	0,000621	1	C	Z
22	*Товар 21*	44,00	0	4269,395	355,7829	77,2477	0,2171203	187853,4	0,002478	0,997769433	C	Y
23	*Товар 22*	44,00	1	4490,925	374,2438	381,2762	0,9815556	197600,7	0,002606	0,995291636	C	Z
24	*Товар 23*	22,00	1	45155,16	3762,93	3793,463	0,9919513	993413,5	0,013103	0,975107818	C	Z
25								75814696				

Рис. 3. Результаты ABC-XYZ-анализа с учетом цензурирования

ЛИТЕРАТУРА

1. *Гаджинский А.М.* Логистика: учеб. 17-е изд., перераб. и доп. – М.: Дашков и К, 2008. – 484 с.

2. *Зенкова Ж.Н., Стебловская А.С.* ABC-, XYZ-анализ спроса на слабоалкогольные напитки // EDUCATIO, Новосибирск. – 2016. – № 7 (25) – 8 (26). – С. 26–29.

3. *Зенкова Ж.Н., Гаман М.И.* Модификация методов ABC-, XYZ-анализа на случай цензурированных данных // Материалы VII Международной конференции «Логистические системы в глобальной экономике», 16-17 марта 2017 г., Красноярск. – Красноярск. – 2017. – С. 150–153.

4. *Zenkova Z.N., Kabanova T.V.* The ABC-XYZ Analysis Modified for Data with Outliers // Proceeding GOL'2018 The 4th IEEE International Conference on Logistics Operations Management. April 10-12. – Le Havre, France. – 2018. – P. 63–68.

5. *Dmitriev Y., Zenkova Z., Musoni W.* Statistical Estimation with a Known Quantile and Its Application in a Modified ABC-XYZ Analysis // 8th International Conference on Risk Analysis and Design of Experiments. Vienna, 23-27 Apr., 2019. Book of Abstracts. – P. 143–144.

6. *Зенкова Ж.Н., Краковецкая И.В.* Моделирование по неполным данным в логистике и маркетинге // Логистические системы в глобальной экономике: материалы Междунар. науч.-практ. конф. (14-15 марта 2013 г., Красноярск): в 2 ч. Ч. 1. Научно-исследовательский сектор / Сиб. гос. аэрокосмич. ун-т. – Красноярск. – 2013. – С. 98–105.

7. *Скрипник В.М., Назин А.Е., Приходько Ю.Г., Благовещенский Ю.Н.* Анализ надежности технических систем по цензурированным выборкам. – М.: Радио и связь, 1988. – 184 с.

8. *Зенкова Ж.Н., Краковецкая И.В.* Непараметрическая оценка Тёрнбулла для интервально-цензурированных данных в маркетинговом исследовании спроса на биоэнергетические напитки // Вестник томского государственного университета. Управление, вычислительная техника и информатика. – 2013. – № 3(24). – С. 64–69.

9. *Дмитриев Ю.Г., Зенкова Ж.Н.* Об оценивании симметричного распределения по цензурированной выборке // Труды X юбилейного симпозиума по непараметрическим и робастным статистическим методам в кибернетике. Томск. 2002. – Томск: Изд-во НТЛ, 2004. – С. 22–30.

10. *Дмитриев Ю.Г., Зенкова Ж.Н.* Оценка симметричной функции распределения по цензурированной выборке // Теория вероятностей, случайные процессы, математическая статистика и их приложения: сб. науч. ст. междунар. конф., посвящ. 70-летию профессора, д-ра физ.-мат. наук Г.А. Медведева. Минск, 21-25 февр. 2005 г. / редкол.: Н.Н. Труш (отв. ред.) [и др.]. – Мн.: БГУ, 2005. – С. 78–86.

11. *Дмитриев Ю.Г., Зенкова Ж.Н.* Оценки симметричной функции распределения по усеченным и цензурированным выборкам // Вестник ТГУ. – 2005. – № 14. – С. 291–296.

12. *Боровков А.А.* Математическая статистика. – Новосибирск: Наука. 1997. – 772 с.

РАСЧЕТ ЦЕНЫ СПРОСА НА ПРОГРАММНЫЙ ПРОДУКТ МЕТОДОМ PSM ДЛЯ ЗАВИСИМЫХ НАБЛЮДЕНИЙ

Н.Р. Цыбульникова

Томский государственный университет

Введение

В процессе вывода нового продукта на рынок очень важно определить ту цену, с которой производитель готов выйти на массового потребителя. В связи с этим возникает ряд проблем, т.к. он не обладает достаточной информацией о возможном уровне цен, потому что предприятие не имеет накопленной информации по ранее совершённым продажам. Это вынуждает маркетологов пользоваться историческим методом, когда в основу анализа ложатся данные о продажах подобного, уже выпускаемого товара.

Цель данного исследования состояла в определении предельно допустимой цены, которую потребители готовы отдавать за товар-новинку – программный продукт, выводимый томским предприятием на рынок B2B. При этом использовались ценовые значения, полученные в процессе маркетингового исследования, проведенного с целью получения данных для метода измерения ценовой чувствительности PSM. В ходе исследования была выдвинута гипотеза о том, что полученные уровни цен могут линейно зависеть друг от друга, в частности, они связаны с минимальной из четырёх рассматриваемых цен.

В работе исследовался вопрос, как изменятся результаты применения метода PSM, а также метода Ван Вестендорпа и Штоцеля, если при расчётах привлечь информацию о наличии зависимости ценовых значений.

1. Методы ценообразования, ориентированные на спрос

Существует три основные метода, ориентированные на потребительские предпочтения, которые используются маркетологами при определении цены товара-новинки [1-4,5,9].

Метод PSM – Price Sensitivity Meter

Одним из самых распространенных методов является метод PSM – Price Sensitivity Meter [3], предложенный Питером Ван Вестендорпом. Данный метод заключается в том, что представителям целевой аудитории после того, как они имели возможность пользоваться товаром в течение некоторого периода времени, предлагается ответить на четыре вопроса:

1) Ниже какого уровня цены P_1 товар кажется Вам настолько дешевым, что начинают возникать подозрения о том, что товар некачественный или поддельный?
2) Какая цена P_2 для Вас является приемлемой для покупки товара?
3) Какая цена P_3 кажется Вам высокой, однако Вы еще рассматриваете вопрос о покупке?
4) Начиная с какого уровня цены P_4 товар кажется Вам настолько дорогим, что вопрос о покупке даже не ставится?

В результате получаем выборку объема N, каждый элемент которой представляет собой четыре значения цен $\left(P_{i1}, P_{i2}, P_{i3}, P_{i4}\right)$, $i = \overline{1,N}$. По каждому вектору $\left(P_{ij}\right)$, $j = \overline{1,4}$, строится вариационный ряд, а затем эмпирическая функция распределения (э.ф.р.) по формуле $F_N^{(j)}(p) = \frac{1}{N}\sum_{i=1}^{N} I_{(0,p)}\left(P_{ij}\right)$, где $I_A(y) = \begin{cases} 0, & y \notin A, \\ 1, & y \in A, \end{cases}$ – индикаторная функция.

Для первой и третьей последовательности цен строится оценка функции выживания $S_N^{(j)}(p) = 1 - F_N^{(j)}(p)$. Далее эти функции отображаются на одном графике. В данном методе они называются кривыми ценовой чувствительности (Price Sensitivity Curve).

В качестве рекомендуемого диапазона цен рассматривается отрезок, правый конец которого – абсцисса точки пересечения $S_N^{(3)}(p)$ и $F_N^{(4)}(p)$, левый – значение точки p, при которой пересекаются и $S_N^{(1)}(p)$ и $F_N^{(2)}(p)$. Назначать цену ниже минимально возможной $P_{\text{мин}}$ нет смысла, т.к. потребители будут воспринимать товар как некачественный. Если назначить цену выше максимально возможной $P_{\text{макс}}$, то потребители будут считать его слишком дорогим. Стоит отметить, что если в результате опросов цена безразличия $P_{\text{безр}}$ (абсцисса точки пересечения $S_N^{(3)}(p)$ и $F_N^{(2)}(p)$) будет меньше, чем оптимальная $P_{\text{опт}}$ (абсцисса точки пересечения $S_N^{(1)}(p)$ и $F_N^{(4)}(p)$), то это значит, что в исследовании участвовало достаточно много представителей нецелевой аудитории, которые мало что знают о данном товаре и его ценности, поэтому дают несправедливые, слишком заниженные или завышенные значения цен, иногда с большим разбросом.

Для того, чтобы наиболее точно определить цену, с которой производитель будет выводить готовый продукт на рынок, нужно сравнить значение потенциального объема рынка в значении оптимальной цены и в значении ожидаемой цены.

Метод Ж. Штоцеля «Психологическая цена»

Является в некотором смысле частным случаем метода PSM. Он был предложен французским психологом и социологом Жаном Штоцелем в середине XX в. В основу расчётов берутся цены P_1 и P_4, полученные методом PSM. Искомые цены $P_{\min}$ и $P_{\max}$, которые являются границами диапазона, в рамках которого, в дальнейшем и рекомендуется назначать цену на товар, находят из уравнения $d(p) = F_N^{(1)}(p) - F_N^{(4)}(p) = 0.5$.

Метод Ван Вестендорфа

Ещё одна, максимально простая версия метода PSM. Метод даёт точечную оценку цены, которая находится из уравнения $S_N^{(1)}(p) = F_N^{(4)}(p)$.

Заметим, что по сравнению с вышеописанными методами, метод PSM считается более точным, т.к. диапазон рекомендованных цен, полученных с помощью этого метода, значительно сужается, относительно ценового интервала, полученного методом Штоцеля.

2. Модификация методов ценообразования с учетом линейной зависимости

Рассматривается ситуация, когда ценовые предпочтения потребителя линейно зависимы от минимального уровня цен, т.е., имеет место равенство: $P_i = a_i + b_i P_1,\ i = \overline{2,4}$, где a_i и $b_i,\ i = \overline{2,4}$ – некоторые коэффициенты. Тогда для $i = \overline{2,4}$

$$F^{(i)}(p) = P\left(P_i < p\right) = P\left(a_i + b_i P_1 < p\right) = P\left(P_1 < \frac{p - a_i}{b_i}\right) = F^{(1)}\left(\frac{p - a_i}{b_i}\right). \qquad (1)$$

Формула (1) позволяет модифицировать методику расчёта кривых ценовой чувствительности, заменив $F_N^{(j)}(p)$ и $S_N^{(j)}(p)$ на $F_N^{(i)}(p) = F_N^{(1)}\left(\frac{p - a_i}{b_i}\right)$, $i = \overline{2,4}$, $F_N^{(3)}(p) = F_N^{(1)}\left(\frac{p - a_3}{b_3}\right)$.

Однако в связи с тем, что истинные значения коэффициентов a_i и $b_i,\ i = \overline{2,4}$, неизвестны, на практике возникает необходимость в их оценивании, например, с помощью метода наименьших квадратов, что, фактически, приводит к модели линейной регрессии.

3. Определение цены нового программного продукта с помощью модифицированных методов

Пусть требуется определить цену на новый программный продукт, выводимый на рынок товаров производственного назначения B2B. Товар был предоставлен респондентам в пользование на период один месяц, после чего производился опрос по методу PSM[7]. Результаты опроса приведены в табл. 1.

Таблица 1

Результаты опроса потребителей программного продукта по методу PSM, у.е./шт.

№	P_1	P_2	P_3	P_4	№	P_1	P_2	P_3	P_4	№	P_1	P_2	P_3	P_4
1	53	145	204	257	19	130	262	333	382	37	186	301	391	526
2	58	172	239	287	20	138	262	335	384	38	186	303	394	529
3	60	173	253	290	21	144	264	339	390	39	193	306	399	560
4	66	175	254	301	22	148	266	341	398	40	196	307	403	591
5	73	186	256	302	23	151	266	350	418	41	198	308	408	611
6	78	186	257	305	24	152	271	350	419	42	200	311	408	620
7	82	186	260	309	25	155	274	354	421	43	205	313	424	643
8	85	199	274	318	26	157	276	355	422	44	206	313	427	660
9	95	201	277	322	27	159	277	359	428	45	209	314	430	685
10	103	202	279	324	28	162	280	362	432	46	225	320	438	690
11	104	220	283	327	29	163	284	364	436	47	238	322	439	773
12	108	222	293	352	30	165	286	367	439	48	259	323	474	838
13	108	236	294	359	31	165	288	367	454	49	270	327	474	869
14	109	251	302	360	32	167	291	368	462	50	288	336	508	1152
15	116	253	308	363	33	173	293	373	467	51	290	343	581	1309

№	P_1	P_2	P_3	P_4	№	P_1	P_2	P_3	P_4	№	P_1	P_2	P_3	P_4
16	125	255	310	365	34	181	294	374	472	52	301	361	655	1350
17	126	256	313	374	35	183	295	377	489					
18	129	257	315	379	36	184	300	381	522					

Факт наличия линейной зависимости между случайными величинами подтверждается высокими значениями коэффициентов корреляции (см. табл. 2) [6,8,10]. Также в табл. 2 представлены значения коэффициента детерминации и соответствующие уравнения линейной регрессии.

Таблица 2

Результаты выявления зависимости между выборками

	Зависимость P_2 от P_1	Зависимость P_3 от P_1	Зависимость P_4 от P_1
Коэффициент корреляции	0.95	0.97	0.91
Коэффициент детерминации	0.903	0.948	0.829
Уравнение линейной зависимости	$P_2 = 142.92 + 0.79*P_1$	$P_3 = 147.81 + 1.34P_1$	$P_4 = -58.58 + 3.56P_1$

Полученные результаты позволяют сделать вывод об очень сильной зависимости между данными выборками. Воспользуемся данной информацией для получения новых ценовых значений с помощью рассматриваемых методов. На рис. 1, 2 приведены графики кривых ценовой чувствительности метода PSM для классической и модифицированной методик расчёта.

При классическом подходе получаем следующие ценовые значения: $P_1 = 202$, $P_2 = 290$, $P_3 = 307$, $P_4 = 381$ у.е./шт. При модифицированном подходе: $P_1 = 206.6$, $P_2 = 223.3$, $P_3 = 296.67$, $P_4 = 396.7$ у.е./шт.

Таким образом, при применении классического метода PSM получаем ценовой диапазон [202;381] в рамках которого, согласно методу, можно назначать цену на товар, а при применении модифицированного метода получаем ценовой диапазон [206.6;396.7], который смещен относительно диапазона, полученного классическим методом.

Заметим также, что при модификации (рис. 2) возникли недопустимые пересечения кривых ценовой чувствительности, что является недопустимым.

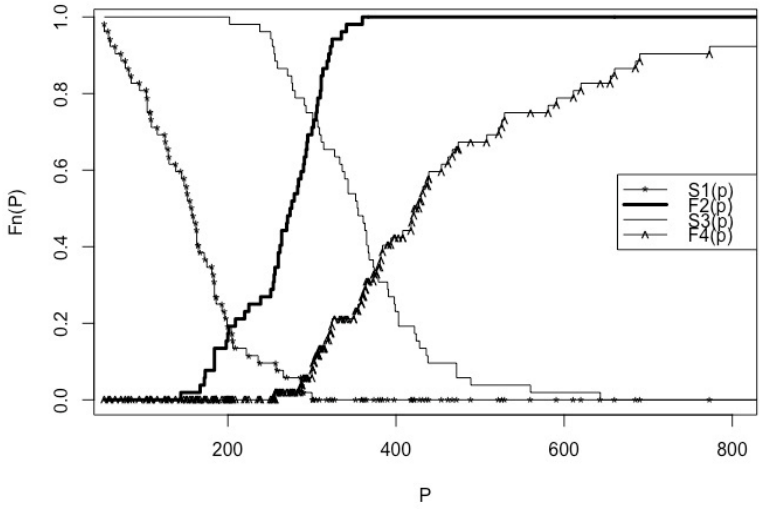


Рис. 1. Реализация метода PSM по реальным данным

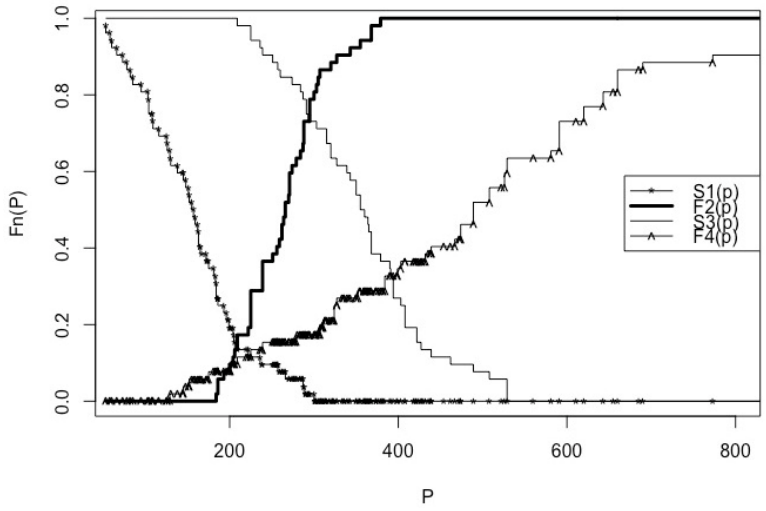


Рис. 2. Реализация модифицированного метода PSM по реальным данным

Применив данную модификацию для метода Штоцеля, получаем следующие результаты (см. рис. 3, 4): при классическом подходе получаем ценовые значения $P_1 = 157$, $P_2 = 422$ у.е./шт., при модифицированном подходе: $P_1 = 163.24$, $P_2 = 505.5$ у.е./шт.

Следовательно, при применении классического метода Штоцеля получаем ценовой диапазон [157;422] у.е./шт. товара, а при применении модифицированного метода Штоцеля [163.24;505.5] у.е./шт., который, как и при реализации метода PSM, оказался смещённым относительно ценового диапазона, полученного при реализации классического метода Штоцеля.

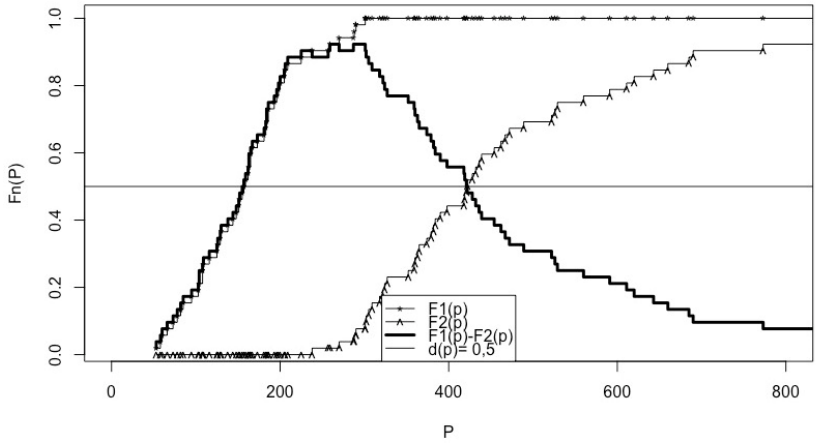


Рис. 3. Реализация метода Штоцеля по реальным данным

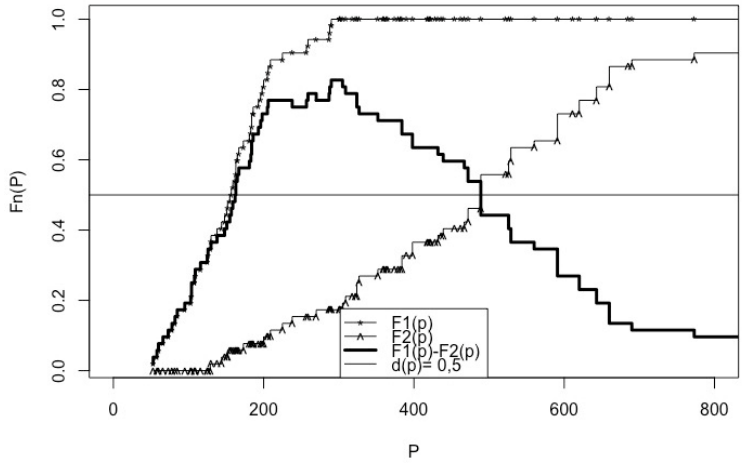


Рис. 4. Реализация модифицированного метода Штоцеля по реальным данным

Для метода Ван Вестендорпа (рис. 5, 6) при классическом подходе получаем цену $P_1 = 294$ у.е./шт., а при модифицированном подходе $P_1 = 209.7$ у.е./шт., которая, как и в предыдущих методах, смещена относительно цены, полученной при классическом подходе.

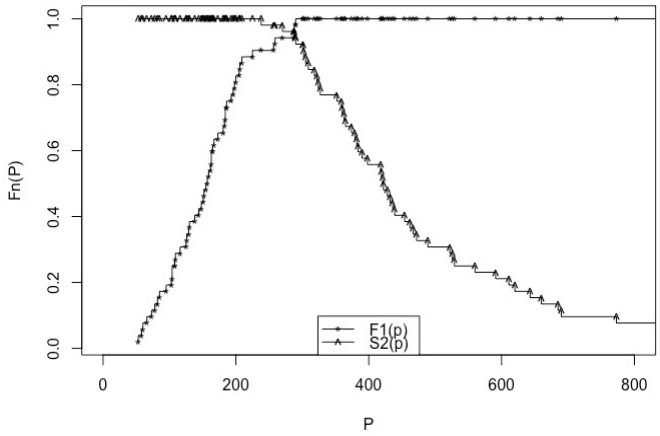


Рис. 5. Реализация метода Ван Вестендорпа на реальных данных

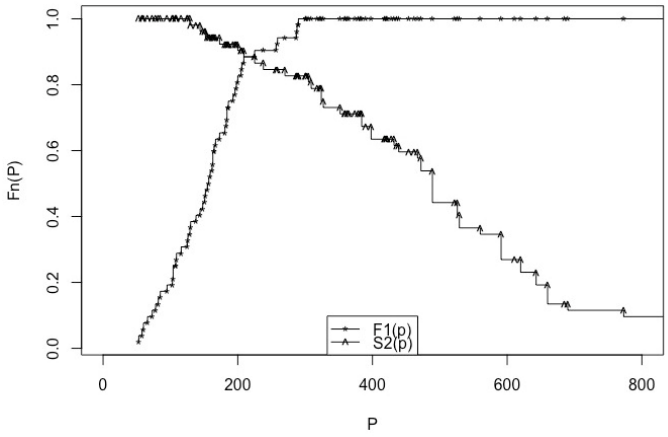


Рис. 6. Реализация модифицированного метода Ван Вестендорпа на реальных данных

Заключение

В данной работе предложена модификация методов ценообразования, ориентированных на срос, на случай линейной зависимости ценовых предпочтений потребителя. Для каждого из описанных методов (PSM, Штоцеля и Ван Вестендорпа) на реальных данных о ценовых предпочтений целевых потребителей нового программного продукта были рассмотрены модифицированный и классический подходы. Получены ценовые диапазоны, которые могут быть использованы производителем товара-новинки как базовые, в рамках которых производитель может назначать цену на свой продукт.

Данная модификация имеет право на существования, но требует дополнительных исследований. При модификации было выявлено искажение метода PSM, что может привести к недопустимым изменениям и нарушению работы метода. Это связано с тем, что при определенных обстоятельствах возникает дополнительное пересечение кривых ценовой чувствительности. Необходимо выявить условия и ограничения, при которых данный подход будет работать корректно и иметь преимущества перед классическим подходом.

ЛИТЕРАТУРА

1. *Kotler P., Armstrong G.* Principles of Marketingю – Prentice-Hall, 2016. – 398 p.
2. *Korzhov M.M., Kuzin D.A.* Comparative analysis and practical application of demand-driven pricing methods // Marketing News. – 2007. – № 4. – P. 9–17.
3. *Westendorp P.H.* NSS – Price Sensitivity Meter (PSM) – A New Approach to Consumer Perception of Prices // Venice Congress Main Sessions, European Marketing Research Society (ESOMAR). – Amsterdam, 1976. – P. 139–167.
4. *Stoetzel J.* Psychological/ Sociological Aspects of Price // Bernard Taylor and Gordon Wills (Eds.). Pricing Research, Princeton, NJ: Brandon Systems Press, 1970. – P. 70–74.
5. *Zenkova Zh., Krakovetckaia I.* Marketing of New Bio-Energy Drinks // Atiner conference paper series. – No: SME2013-0739. – 2013. – 14 p. – URL: http://www.atiner.gr/papers/SME2013-0739.pdf
6. *Боровков А.А.* Математическая статистика. – Новосибирск: Наука, 1997. – 772 с.
7. *Зенкова Ж.Н., Копнова Е.Е., Бараксанов Д.Н.* Применение метода PSM и алгоритма Тёрнбулла при определении цены нового программного продукта // Сборник статей XVII международной научно-практической конференции "Высокие технологии, фундаментальные исследования, инновации". С.-Пб.: Изд-во ТПУ, 2014. – 301 с.
8. *Зенкова Ж.Н., Краковецкая И.В.* Непараметрическая оценка Тёрнбулла для интервально-цензурированных данных в маркетинговом исследовании спроса на биоэнергетические напитки // Вестник томского государственного университета, Управление, вычислительная техника и информатика. – 2013. – № 3(24). – С. 64–69.
9. *Зенкова Ж.Н., Цыбульникова Н.Р.* Исследование цены спроса на продукцию ЗАО «Натуральное мыло» методом Штоцеля по цензурированным данным // Материалы IV Международной молодежной научной конференции «Математическое и программное обеспечение информационных, технических и экономических систем», Томск, 20-21 мая 2016 г. Издательский дом Томского государственного университета. – С. 62–65.
10. *Дмитриев Ю.Г. Зенкова Ж.Н., Зенков А.Г.* Статистические свойства оценок метода измерения ценовой чувствительности PSM // Вестник Томского государственного университета. Управление, вычислительная техника и информатика. – 2019. – № 47.

ПРОГНОЗИРОВАНИЕ НА ОСНОВЕ МОДЕЛЕЙ РАСТУЩИХ СИСТЕМ

Ю.Е. Пальянова

Томский государственный университет

Введение

В работе рассматривается применение математической модели растущих систем [1−3] к моделированию и прогнозированию цен различных финансовых котировок. Под растущей системой понимается любая совокупность объектов одной природы, которые вместе участвуют в процессах возникновения, создания и развития системы, подчиняющихся законам, одинаковым для всех объектов. Описание растущих систем возможно с помощью вероятностно-статистических моделей [1,2]. Системы подобного типа используются в геологии для изучения совокупности месторождений полезных ископае-

мых [4,5], для выявления закономерностей численности населения различных территорий [6], определения экстремального расхода воды в реках [7], расчета финансовых показателей и в других сферах практической деятельности.

В данной работе, исходя из предположения, что совокупность цен курсов валют или индексов за определенный период времени является растущей системой, проводится подбор математической модели, которая отражает зависимость максимальных цен котировок за разные периоды времени. Полученная модель приводит к специфической зависимости между ценами, позволяющей предсказывать тенденцию изменения их максимальных значений. Результаты применения модели иллюстрируются на модельных данных, а также курсах валют и индексов.

Населения различных территорий за определенный период времени также являются растущими системами. Для таких систем рассмотрена задача восстановления пропущенных данных о численности территорий на примере французских регионов за период с 1975 по 2018 год.

1. Описание модели для анализа цен акций

Пусть имеется N временных периодов с одинаковым количеством n дней. Наблюдается максимальное значение цены акции в каждый из n дней. Максимальные цены между днями будем считать независимыми и одинаково распределенными случайными величинами с функцией распределения $F(x)$. Обозначим максимальную цену акции за n дней в k-м периоде через $X_k^{(n)}$ и рассмотрим последовательность $X_1^{(n)},...,X_N^{(n)}$, отвечающую N периодам. При этом функция распределения $G^{(n)}(x) = P\left\{X_k^{(n)} < x\right\} = F^n(x)$.

Выберем в качестве $F(x)$ двойной экспоненциальный закон $F(x;\mu,\sigma) = \exp\{-\exp\{-(x-\mu)/\sigma\}\}$, $-\infty < x < \infty$, $-\infty < \mu < \infty$, $\sigma > 0$, предложенный в [7] для описания максимальных значений наблюдаемой величины. Упорядочим наблюдения по возрастанию: $X_{(1)}^{(n)} < X_{(2)}^{(n)} < ... < X_{(N)}^{(n)}$. Порядковые статистики $X_{(k)}^{(n)}$ можно рассматривать как оценки квантилей закона распределения $G^{(n)}(x)$ уровней $p_k = k/(N+1)$, $k = \overline{1,N}$. На основании этого, следуя [8], можно записать, $F_0^n\left(\left(X_{(k)}^{(n)} - \mu\right)/\sigma\right) = p_k + \varepsilon_{(k)}$, где $\varepsilon_{(k)} = u_{(k)} - p_k$, $u_{(k)}$ – k-я порядковая статистика выборки из равномерного в (0,1) закона. Отсюда

$$\frac{X_{(k)}^{(n)} - \mu}{\sigma} = F_0^{-1}\left(\left(p_k + \varepsilon_{(k)}\right)^{1/n}\right),\ k = \overline{1,N}, \tag{1}$$

$$\frac{X_{(k)}^{(n)} - \mu}{\sigma} \approx F_0^{-1}\left(p_k^{1/n}\right) + \varepsilon_{(k)},\ k = \overline{1,N}, \tag{2}$$

$$X_{(k)}^{(n)} \approx \mu + \sigma \cdot Y_{(k)}^{(n)} + \varepsilon_{(k)},\ k = \overline{1,N}, \tag{3}$$

где $Y_{(k)}^{(n)} = F_0^{-1}\left(p_k^{1/n}\right) = \left(-\ln\left(-n^{-1}\ln\left(p_k\right)\right)\right)$, $k = \overline{1,N}$, $F_0(x) = F(x;0,1)$ с $M\varepsilon_{(k)} = 0$ и матрицей ковариации $V = \operatorname{cov}\left(\varepsilon_{(k)}, \varepsilon_{(k')}\right) = \dfrac{k\left(N-k'+1\right)}{\left(N+1\right)^2\left(N+2\right)}$, $k,k' = \overline{1,N}$ для порядковых статистик из равномерного закона в (0,1). Модель (3) отражает линейную регрессионную зависимость значений максимальных цен акций от квантилей закона распределения $F_0(x)$. Параметры μ и σ находятся по обобщенному методу наименьших квадратов (ОМНК) [8].

Рассмотрим уравнение (1), разложив функцию F_0^{-1} по формуле Тейлора:

$$\frac{X_{(k)}^{(n)}-\mu}{\sigma}=F_0^{-1}\left(p_k^{1/n}\right)+\left[F_0^{-1}\left(p_k^{1/n}\right)\right]'\varepsilon_{(k)}+\frac{1}{2}\left[F_0^{-1}\left(p_k^{1/n}\right)\right]''\varepsilon_{(k)}^2+O\left(\varepsilon_{(k)}^3\right),$$

$$X_{(k)}^{(n)}=\mu+\sigma\left(F_0^{-1}\left(p_k^{1/n}\right)+\left[F_0^{-1}\left(p_k^{1/n}\right)\right]'\varepsilon_{(k)}+\frac{1}{2}\left[F_0^{-1}\left(p_k^{1/n}\right)\right]''\varepsilon_{(k)}^2+O\left(\varepsilon_{(k)}^3\right)\right),$$

$$EX_{(k)}^{(n)}=\mu+\sigma\left(F_0^{-1}\left(p_k^{1/n}\right)+\frac{1}{2}\left[F_0^{-1}\left(p_k^{1/n}\right)\right]''E\varepsilon_{(k)}^2+O\left(E\varepsilon_{(k)}^3\right)\right)+\sigma\left[F_0^{-1}\left(p_k^{1/n}\right)\right]'E\varepsilon_{(k)}.$$

Т.к. $\varepsilon_{(k)}$ – это порядковые статистики из равномерного распределения, то согласно [8] при $k=k'$ получим $E\varepsilon_{(k)}^2=\frac{k(N-k+1)}{(N+1)^2(N+2)}$.

$$EX_{(k)}^{(n)}=$$

$$=\mu+\sigma\left(F_0^{-1}\left(p_k^{1/n}\right)+\frac{1}{2}\left[F_0^{-1}\left(p_k^{1/n}\right)\right]''\frac{k(N-k+1)}{(N+1)^2(N+2)}+O\left(E\varepsilon_{(k)}^3\right)\right)+\sigma\left[F_0^{-1}\left(p_k^{1/n}\right)\right]'E\varepsilon_{(k)}.$$

Если пренебречь величиной $O\left(E\varepsilon_{(k)}^3\right)$, то

$$X_{(k)}^{(n)}=\mu+\sigma\left(F_0^{-1}\left(p_k^{1/n}\right)+\frac{1}{2}\left[F_0^{-1}\left(p_k^{1/n}\right)\right]''\frac{k(N-k+1)}{(N+1)^2(N+2)}\right)+\sigma\left[F_0^{-1}\left(p_k^{1/n}\right)\right]'E\varepsilon_{(k)},$$

$$EX_{(k)}^{(n)}=\mu+\sigma\left(F_0^{-1}\left(p_k^{1/n}\right)+\frac{1}{2}\left[F_0^{-1}\left(p_k^{1/n}\right)\right]''\frac{k(N-k+1)}{(N+1)^2(N+2)}\right). \quad (4)$$

Обозначим $\alpha_{(k)}^{(n)}=F_0^{-1}\left(p_k^{1/n}\right)+\frac{1}{2}\left[F_0^{-1}\left(p_k^{1/n}\right)\right]''\frac{k(N-k+1)}{(N+1)^2(N+2)}$.

Подставив $\alpha_{(k)}^{(n)}$ в (4), получим $EX_{(k)}^{(n)}=\mu+\sigma\alpha_{(k)}^{(n)}$. Таким образом, имеем уточненную, по сравнению с (3), модель:

$$X_{(k)}^{(n)}=\mu+\sigma\cdot\alpha_{(k)}^{(n)}+\varepsilon_{(k)},\ k=\overline{1,N}. \quad (5)$$

2. Алгоритм прогноза

Задача заключается в прогнозировании тенденции (увеличении или уменьшении) максимальной цены акции в $N+1$ периоде по сравнению с предыдущим. Прогнозирование тенденции максимальной цены финансовой котировки в $N+1$ периоде предлагается проводить следующим образом [1]. Пусть известны максимальные значения цен котировок за N периодов, и вычислены квантили $Y_{(1)}^{(n)}<...<Y_{(N+1)}^{(n)}$ за $N+1$ период. Поставим в соответствие N упорядоченным значениям цен $X_{(k)}^{(n)}$ N квантилей $Y_{(k)}^{(n)}$ из $N+1$ возможных. Наилучшим будем считать такое соответствие, при котором погрешность

$$STD=\sqrt{\sum_{k=1}^{N}\left(X_{(k)}^{(n)}-\hat{X}_{(k)}^{(n)}\right)^2}, \quad (6)$$

где $\hat{X}_{(k)}^{(n)}\approx\hat{\mu}+\hat{\sigma}\cdot Y_{(k)}^{(n)}+\varepsilon_{(k)}$, $k=\overline{1,N}$ (используется для модели (3)) или $\hat{X}_{(k)}^{(n)}=\hat{\mu}+\hat{\sigma}\cdot\alpha_{(k)}^{(n)}+\varepsilon_{(k)}$, $k=\overline{1,N}$ (используется для модели (5)) будет наименьшей по всем возможным переборам соответствий с учетом порядка, а также $\hat{\mu}$ и $\hat{\sigma}$, вычисляемых по ОМНК для каждого соответствия. Когда такое соответствие найдено, то становится известен «свободный» квантиль $Y_{(k)}^{(n)}$, которому соответствия не нашлось. Под-

ставляя этот квантиль в оцененное уравнение регрессии при наилучшем соответствии, находим прогнозное значение неизвестной цены акции на $N+1$ период и, соответственно, тенденцию цены.

3. Прогнозирование модельных данных

Возьмем в качестве $F(x)$ двойной экспоненциальный закон распределения $F(x;\mu,\sigma)=\exp\{-\exp\{-(x-\mu)/\sigma\}\}$. Сгенерируем 100 выборок для каждого набора значений параметров из этого закона с заданными μ, σ, N и n. Следовательно, можно увидеть растет или убывает значение $N+1$ периода. При этом качество прогноза будем отслеживать по среднему проценту верных прогнозов. Для сравнения будем использовать две модели, определенными формулами (3) и (5).

В табл. 1 приведены результаты по прогнозированию значения на $N+1$ период при фиксированном μ и разных N, n и σ. Также показано сколько раз тенденция этого значения была верно определена. По результатам прогноза обе модели из формул (3) и (5) улавливают тенденции одинаково.

Таблица 1

Результаты прогноза на $N+1$ период

Значения параметров	Средний процент верных прогнозов тенденции при n = 1, %	Средний процент верных прогнозов тенденции при n = 5, %
$\mu=2\ \ \sigma=5\ \ N=4$	85	65
$\mu=2\ \ \sigma=5\ \ N=9$	80	60
$\mu=2\ \ \sigma=0.1\ \ N=4$	90	70
$\mu=2\ \ \sigma=0.1\ \ N=9$	90	70

Табл. 1 показывает, что наибольшее влияние на точность прогноза тенденции оказывает масштабный параметр σ: чем больше σ при фиксированном μ, тем хуже модель оценивает все параметры и тем сложнее определить значение на $N+1$ период. При этом количество элементов N в выборке не так сильно влияет на качество прогноза.

В табл. 2 приведены результаты по прогнозированию значения на $N+1$ период при фиксированном σ и разных N, n и μ. Также показано сколько раз тенденция этого значения была верно определена.

Таблица 2

Результаты прогноза на $N+1$ период

Значения параметров	Средний процент верных прогнозов тенденции при n = 1 , %	Средний процент верных прогнозов тенденции при n = 5, %
$\mu=2\ \ \sigma=2\ \ N=4$	80	65
$\mu=2\ \ \sigma=2\ \ N=9$	76	61
$\mu=1000\ \ \sigma=2\ \ N=4$	90	75
$\mu=1000\ \ \sigma=2\ \ N=9$	85	70

По результатам из табл. 2 можно сделать вывод о том, что разные значения параметра положения μ несильно влияют на качество прогноза.

Таким образом, наиболее точный результат прогнозирования тенденции для значения в $N+1$ период достигается при $n=1$ и любом объеме выборки N. При этом обе модели из формул (3) и (5) отлавливают тенденции одинаково. Следовательно, данные результаты целесообразно использовать для дальнейшего прогноза реальных цен акций.

4. Прогноз тенденции по реальным данным

Рассмотрим реальные данные: курс евро по отношению к рублю и индекс ММВБ 10. Поскольку торги на бирже открыты с понедельника по пятницу, то за неделю возьмем пять дней, т.е. $N = 4$, $N + 1 = 5$, $n = 1$. Точность подгонки будем определять по наименьшему стандартному отклонению из формулы (6). Сравнивать полученные тенденции цен будем по моделям из формул (3) и (5). В табл. 3–4 знаком «+» показано, что данный способ отразил тенденцию цены на пятый день. Соответственно, знак «–» отражает тот факт, что выбранный способ не показывает реальную ситуацию. Приведем результаты прогноза для курса валют и индекса в таблицах 3–4.

Таблица 3

Результаты прогнозирования для курса валют евро-рубль для двойного экспоненциального закона

№ п/п	STD для модели (3)	Прогнозируемая цена курса для модели (3), руб.	STD для модели (5)	Прогнозируемая цена курса для модели (5), руб.
1	1.791	71.929 (+)	0.552	73.580 (+)
2	0.140	75.390 (−)	0.608	74.805 (−)
3	0.103	75.614 (+)	0.496	75.704 (+)
4	0.039	76.830 (+)	0.164	76.289 (+)
5	0.445	73.048 (+)	0.438	73.004 (+)
6	0.515	71.931 (−)	0.045	71.093 (−)
7	0.091	72.725 (+)	0.531	72.325 (+)
8	0.085	72.701 (−)	0.283	72.518 (−)
9	0.089	73.941 (+)	0.291	73.801 (+)
10	0.018	73.087 (+)	0.290	73.234 (+)
11	0.088	73.463 (+)	0.211	73.880 (+)
12	0.154	72.671 (+)	0.209	72.311 (+)
13	0.054	72.797 (−)	0.167	72.712 (−)
14	0.084	73.523 (+)	0.292	73.890 (+)
15	0.044	73.300 (−)	0.409	73.152 (−)
16	1.584	73.171 (+)	1.167	72.112 (+)
17	0.316	76.806 (+)	0.302	76.970 (+)
18	0.074	76.729 (+)	0.030	76.005 (+)
19	0.406	78.254 (+)	0.270	78.473 (+)
20	0.082	78.641 (−)	0.351	78.807 (−)
21	0.260	79.762 (−)	0.363	80.606 (−)
22	0.391	77.780 (+)	0.311	78.123 (+)
23	0.108	77.036 (+)	0.457	77.452 (+)
24	0.283	75.247 (−)	0.220	75.456 (−)
25	0.201	76.630 (+)	0.289	76.860 (+)
26	0.009	76.138 (−)	0.050	76.101 (−)

За представленный период реальная ситуация отражена в 17 случаях из 26, что соответствует 65%. При оценивании тенденции значения цены курса валюты на пятый день разница между моделями (3) и (5) приемлема с точки зрения практики.

Рассмотрим индекс ММВБ 10. Результаты прогноза тенденции цены представлены в табл. 4.

Таблица 4

Результаты прогнозирования для индекса ММВБ 10 для двойного экспоненциального закона распределения

№ п/п	STD для модели (3)	Прогнозируемая цена индекса для модели (3), руб.	STD для модели (5)	Прогнозируемая цена индекса для модели (5), руб.
1	1.279	4431 (–)	1.669	4403 (–)
2	2.612	4575 (+)	2.114	4546 (+)
3	34.006	4521 (+)	30.684	4509 (+)
4	7.632	4401 (–)	7.888	4396 (–)
5	7.998	4326 (+)	7.183	4377 (+)
6	5.959	4368 (+)	5.611	4398 (+)
7	3.804	4426 (+)	3.484	4407 (+)

№ п/п	STD для модели (3)	Прогнозируемая цена индекса для модели (3), руб.	STD для модели (5)	Прогнозируемая цена индекса для модели (5), руб.
8	3.786	4341 (–)	3.796	4312 (–)
9	2.876	4287 (–)	2.450	4254 (–)
10	9.236	4254 (+)	9.409	4239 (+)
11	14.281	4254 (–)	14.624	4245 (–)
12	12.974	4312 (+)	12.656	4314 (+)
13	10.841	4352 (+)	10.061	4348 (+)
14	8.961	4472 (+)	8.644	4454 (+)
15	6.245	4489 (+)	6.549	4478 (+)
16	6.245	4271 (+)	6.167	4310 (+)
17	4.449	4324 (–)	4.245	4326 (–)
18	6.052	4147 (+)	6.876	4158 (+)
19	14.019	4176 (–)	14.180	4179 (–)
20	4.909	4342 (+)	4.786	4383 (+)
21	22.698	4292 (+)	21.123	4305 (+)
22	4.844	4228 (+)	4.912	4224 (+)
23	26.599	4222 (–)	26.600	4239 (–)
24	4.770	4374 (–)	4.236	4376 (–)
25	2.810	4280 (+)	2.345	4289 (+)
26	1.664	4209 (+)	1.513	4216 (+)

За представленный период реальная ситуация отражена в 17 случаях из 26, что соответствует 65%. При оценивании значения тенденции цены индекса на пятый день разница между моделями (3) и (5) практически несущественна.

5. Восстановление численности населения

Для данной задачи воспользуемся методом наименьших квадратов. Пусть $X_1, X_2, ..., X_N$ – численности населения N территорий. Упорядочим их по возрастанию: $X_{(1)} < X_{(2)} < ... < X_{(N)}$. Пусть F– некоторая непрерывная функция распределения, для которой имеет место соотношение

$$X_{(k)} = F^{-1}(p_k) + \varepsilon_{(k)},\ k = \overline{1,N}, \tag{7}$$

где $p_k = k/(N+1)$, $k = \overline{1,N}$, ε_k – ошибки с $M\varepsilon_k = 0$ и $M\varepsilon_k^2 = \nu_k^2 < \infty$.

Далее будем рассматривать семейство сдвигово-масштабных функций распределения: $F(x;\mu,\sigma) = F_0\left((x-\mu)/\sigma\right)$, где $F_0(x) = F(x;0,1)$ – стандартная функция распределения. Использование таких семейств приводит к выражению

$$X_{(k)} = \mu + \sigma \cdot Y_{(k)} + \varepsilon_{(k)},\ k = \overline{1,N}, \tag{8}$$

где $Y_{(k)} = F_0^{-1}(p_k)$ – квантиль уровня p_k функции распределения F_0. Модель (8) отражает линейную регрессионную зависимость значений численности от квантилей стандартного закона.

Применим модель (8) для восстановления численности населения регионов Франции. Для этого в качестве $F(x)$ возьмем логнормальный закон распределения $\Phi\left(\left(\ln\left(X_{(k)}\right)-\mu\right)/\sigma\right)$, где $x > 0$, $\mu > 0$, $\sigma > 0$, при котором модель (8) примет вид:

$$\ln\left(X_{(k)}\right) = \mu + \sigma \cdot Y_{(k)} + \varepsilon_{(k)},\ k = \overline{1,N}, \tag{9}$$

где $Y_{(k)}$ – квантили стандартного нормального закона распределения.

Рассмотрим данные по численности французских регионов в период с 1975 по 2018 годы. Определим погрешность, которую дает данный метод:

$$A = \frac{1}{N}\sum_{k=1}^{N}\left|\frac{\hat{X}_{(k)} - X_{(k)}}{X_{(k)}}\right| \cdot 100\%, \tag{10}$$

где $\hat{X}_{(k)}$ – прогнозные значения $X_{(k)}$.

Рассмотрим 12 французских регионов, которые находятся в метрополии. Случайным образом исключим 5 территорий и с помощью алгоритма прогноза, описанного в 2, восстановим данные по численности. Пусть известна численность населения 7 регионов, а квантили вычислены для всех 12 регионов. Поставим в соответствие 7 упорядоченным значениям численности $X_{(k)}$, $k=\overline{1,7}$ 7 квантилей из 12 возможных. Наилучшим соответствием будем считать такое, при котором погрешность из формулы (10) будет наименьшей. При этом качество восстановленных данных будем оценивать по этой же формуле.

Таблица 5

Восстановление численности населения

Год	Соответствие и погрешность по формуле (10), %
1975	2.845
1976	7.393
1977	7.306
1978	4.075
1979	4.475
1980	3.638
1981	4.990
1982	6.379
1983	5.920
1984	3.313
1985	6.981
1986	4.975
1987	4.423
1988	4.471
1989	5.922
1990	5.296
1991	4.098
1992	6.859
1993	3.106
1994	4.372
1995	6.365
1996	3.111
1997	3.764
1998	7.227
1999	3.816
2000	2.064
2001	4.287
2002	4.674
2003	5.751
2004	4.271
2005	5.512
2006	4.360
2007	5.534
2008	4.374
2009	4.316
2010	4.410
2011	7.945
2012	5.034
2013	2.475
2014	5.082
2015	5.116
2016	4.883
2017	7.591
2018	2.316

Из полученных результатов можно сделать вывод о том, что погрешности являются приемлемыми с точки зрения практики, они не превышают 8%. Таким образом мо-

дель (9) целесообразно использовать для восстановления данных по численности населения.

Заключение

В работе рассмотрено применение математических моделей растущих систем для анализа максимальных цен курса валюты и индекса ММВБ 10, а также численности населения регионов Франции с помощью линейной регрессии, основанной на квантилях равноотстоящих уровней двойного экспоненциального закона распределения. На примерах показаны прогностические свойства моделей (3), (5), а также (9) как на модельных, так и на реальных данных. Выявлено, что разница в прогностических свойствах между моделями (3) и (5) несущественна с точки зрения практики. Модель (9) целесообразно использовать для восстановления данных по численностям населения.

Полученные результаты можно использовать для дальнейшего изучения прогнозных тенденций и реальных цен финансовых котировок, восстановления данных численности различных территорий.

ЛИТЕРАТУРА

1. *Дмитриев Ю.Г., Устинов Ю.К.* Математические модели растущих систем // Вычислительные технологии. – 2007. – Т. 12. – № 1. С. 68–75.

2. *Dmitriev Yu.G., Tarasenko F.P.* Towards a Modeling of Natural Systems // Cybernetics and Systems. – 2009. – № 40. – P. 236–248.

3. *Дмитриев Ю.Г., Тарасенко П.Ф.* Интерпретация геологических данных с использованием компьютерных технологий // Известия ТПУ. – 2008. – Т. 313. – № 5. – С. 26–32.

4. *Дмитриев Ю.Г., Тарасенко П.Ф.* Автоматизированная система «Октава» для геологического прогнозирования СО РАН // Вычислительные технологии. – 2003. – Вып. 8.

5. *Дмитриев Ю.Г., Тарасенко П.Ф.* Об одной природной закономерности // Материалы Всероссийской конференции с международным участием (ЗОНТ-09) 22-24 октября 2009. – Новосибирск. – Т. 1. – С. 4–13.

6. *Талейко А.В.* Закономерности формирования численности населения по территориям // Вестник Томского государственного университета. Управление, вычислительная техника и информатика. – 2009. – № 3. – С. 80–86.

7. *Смирнов Н.В., Дунин-Барковский И.В.* Курс теории вероятностей и математической статистики для технических приложений. – М.: Наука, 1969. – 512 с.

8. Введение в теорию порядковых статистик: пер. с англ. / под редакцией А. Я. Боярского – М.: Статистика, 1970. – 413 с.

РАСЧЕТ ЧИСТОЙ НЕТТО-ПРЕМИИ ПО ПОЛНЫМ И ЦЕНЗУРИРОВАННЫМ ДАННЫМ О СТРАХОВЫХ ВЫПЛАТАХ ПО КОМПЛЕКСНОМУ АВТОМОБИЛЬНОМУ СТРАХОВАНИЮ, КРОМЕ ОТВЕТСТВЕННОСТИ (КАСКО)

Ж.Н. Зенкова, К.А. Серебренникова
Томский государственный университет

Введение

Страховые услуги занимают одно из ведущих мест в рыночной инфраструктуре. Страхование дает возможность застрахованным не беспокоится об убытках от неблагоприятных обстоятельств, таких как финансовые риски, последствия аварий, различного рода стихийные бедствия и пр. [1,2]. В основе расчётов большинства страховых рисков в страховании ином, чем страхование жизни, лежат статистические модели, которые позволяют оценить стоимость страховки, базируясь на знании предыдущих выплат по конкретному виду страхования. Особое внимание уделяется оцениванию среднего уровня убытков, в том числе и по неполным выборкам.

В данной работе с помощью бутстреп-моделирования исследованы свойства оценки среднего уровня убытков по автомобильному страхованию КАСКО некоторой страховой компании по полной и цензурированной выборкам (прогрессивное цензурирование I типа), построены доверительные интервалы для средних.

1. Расчет чистой нетто-премии по полной и цензурированной выборкам

Чистая нетто-премия является основополагающей величиной при вычислении стоимости страхового полиса и рассчитывается с помощью формулы [2,3] $p = z \cdot \mathrm{E}(X)$, где $X \geq 0$ – размер выплаты – случайная величина с функцией распределения (ф.р.) $F(t) = P(X < t)$, $\mathrm{E}(X)$ – ее математическое ожидание, z – вероятность наступления страхового случая.

В действительности z и $\mathrm{E}(X)$ точно не известны, поэтому необходимо использовать статистические оценки, применяя их к накопленным наблюдениям. Таким образом, оценка чистой премии имеет вид $\hat{p} = \hat{z} \cdot \overline{X}$ где $\overline{X} = \frac{1}{N} \cdot \sum_{i=1}^{N} X_i$ – выборочное среднее, $\hat{z}$ – оценка вероятности наступления страхового случая, равная соотношению числа страховых случаев к числу заключенных за период договоров.

В практической деятельности при работе со статистической информацией нередко приходится сталкиваться с данными, которые представляются в виде цензурированных выборок [4], когда вполне допустима неопределенность касательно знаний результатов наблюдений и известно, что некоторые значения попали в определенные интервалы. В страховании такие ситуации встречаются довольно часто, т.к. не всегда возможно определить точный ущерб или последствия страхового случая, поэтому оценщик указывает приблизительную стоимость ущерба.

При наличии цензурирования выборка приобретает следующий вид: $(X, I) = \{(X_1, I_1), .., (X_N, I_N)\}$, где

$$I_i = \begin{cases} 0, & X_i \text{ – полное наблюдение;} \\ 1, & X_i \text{ – цензурированное наблюдение,} \end{cases}$$

и актуарию заведомо приходится работать с выборкой, цензурированной справа.

Работа с цензурированными данными значительно отличается от обработки полных данных и нуждается в особых подходах, при этом возникает значительная потеря информации, которая оказывает влияние на качество статистических процедур [3–11].

Рассмотрим самую простую ситуацию для анализа, когда цензура T_1 заранее известна, т.е. не случайна, количество неполных наблюдений в интервале $[T_1, +\infty)$ – с.в., численно равная доле $g \in (0,1)$ от числа полных наблюдений, попавших в интервал $[T_1, +\infty)$. Тогда оценка ф.р. определяется по формуле Каплана-Мейера, которая может быть представлена в виде [4]

$$F_N^C(t) = \left\{ \begin{array}{l} \dfrac{1}{N} \sum\limits_{i=1}^{N} \mathrm{I}_{[0,t)}(X_i) \overline{I}_i, \ 0 \leq t < T_1, \\ \dfrac{r}{N} + \dfrac{1}{(1-g)N} \sum\limits_{i=1}^{N} \mathrm{I}_{[T_1,t)}(X_i) \overline{I}_i, \ (N-r)(1-g) > 0, \\ r/N, \ (N-r)(1-g) = 0, \end{array} \right\}, \ T_1 \leq t \leq +\infty, \quad (1)$$

где $I_A(x) = \{0 : x \notin A, \ 1 : x \in A\}$, $\overline{I}_i = 1 - I_i$ для $i = \overline{1, N}$, r – число полных наблюдений в интервале $[0, T_1)$. Данная оценка является асимптотически несмещенной, ее дисперсия для $(N-r)(1-g) > 0$ определяется формулой [9,12]

$$\lim_{N\to\infty} NDF_N^C(t) = \begin{cases} F(t)(1-F(t)),\ 0 \le t < T_1, \\ F(t)(1-F(t)) + \dfrac{g(F(t)-p)(1-F(t))}{(1-p)(1-g)},\ T_1 \le t \le +\infty, \end{cases} \quad (2)$$

где $p = F(T_1)$. Очевидно, что $\lim_{N\to\infty} NDF_N^C(t) \ge NDF_N(t) = F(x)(1-F(x))$ – нормированной дисперсии эмпирической функции распределения (э.ф.р.) $F_N(t) = \dfrac{1}{N}\sum_{i=1}^{N} \mathrm{I}_{[0,t)}(X_i)$, что говорит о том, что точность оценивания при появлении цензурирования теряется.

Если в качестве оценки ф.р. использовать оценку $F_N^*(t)$, которая строится как обычная э.ф.р., при этом цензурированные данные рассматриваются как полные, то в итоге получается смещенная оценка, математическое ожидание и асимптотическая нормированная на N дисперсия которой определяются формулами [6]:

$$\lim_{N\to\infty} MF_N^*(t) = \begin{cases} F(t),\ 0 \le t < T_1, \\ g + F(t)(1-g),\ T_1 \le t \le +\infty, \end{cases}$$

$$\lim_{N\to\infty} NDF_N^*(t) = \begin{cases} F(t)(1-F(t)),\ 0 \le t < T_1, \\ F(t)(1-F(t)) + \dfrac{(1-g)(1-F(t))}{1-p}\big((1-p(1-g))F(t) - pg\big),\ T_1 \le t \le +\infty, \end{cases}$$

$$p = F(T_1).$$

На рис. 1 приведены графики математического ожидания оценки $F_N^*(t)$ и дисперсий обычной э.ф.р., оценки (2) и асимптотической нормированной на N среднеквадратической ошибки (MSE) оценки $F_N^*(t)$ для $N = 20$, $g = 0.2$, $p = 0.3$ для $F_N(t) = R_{(0,1)}(t)$.

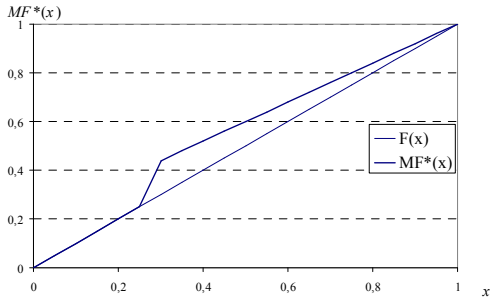


Рис. 1. Графики $MF_N^*(x)$ и $F_N(x)$

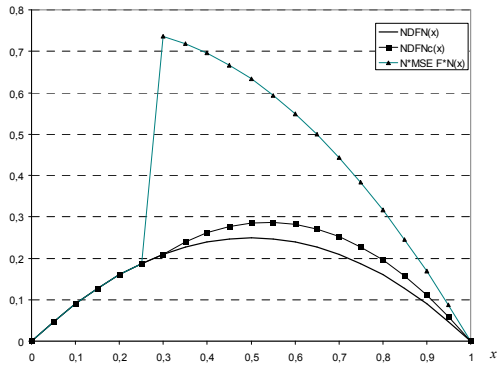


Рис. 2. Графики $\lim_{N\to\infty} NDF_N^C(x)$, $NDF_N(x)$ и $\lim_{N\to\infty} N \cdot MSE\left(F^*(x)\right)$ для $N = 20$, $g = 0.2$, $p = 0.3$ для $F_N(t) = R_{(0,1)}(t)$

Оценку среднего значения случайной величины X можно найти с помощью метода подстановки [13]

$$\overline{X}^{C} = \int_0^T t dF_N^C(t) = \frac{1}{N}\sum_{i=1}^{N} X_i \cdot \overline{I}_i \cdot \mathrm{I}_{[0,T_1)}(X_i) + \frac{1}{(1-g)N}\sum_{i=1}^{N} X_i \cdot \overline{I}_i \cdot \mathrm{I}_{[T_1,+\infty)}(X_i),$$

где $F_N^C(t)$ – оценка (1).

Предлагаемая методика применялась для расчета чистой нетто-премии в комплексном автомобильном страховании КАСКО для некоторой страховой компании. Рассматривались данные об ущербе для $N = 820$ страховых случаев, при этом было получено, что $\hat{z} = 0.183$.

При работе с полными выборками средний ущерб $\overline{X} = 79\ 230.55$ у.е./ед., чистая нетто-премия составила $\hat{p} = 14\ 499.19 =$ у.е./ед. Затем данные были искусственно процензурированы, при этом рассматривалась цензура $T_1 = 50\ 000$, доля $g = 0.2046$ (были процензурированы 71 из 347 значений, превышающих цензуру T_1). В результате расчетов получили, что $\overline{X}^{C} = 73\ 029.22$ у.е./ед., $\hat{p}^{C} = \hat{z} \cdot \overline{X}^{C} = 13\ 264.35$ Заметим, что наличие цензурирования привело к недооценке нетто-премии, а значит, и к снижению цены страхового полиса.

Для рассматриваемой выборки $\overline{X}^{*} = 64\ 164.72$ у.е./ед., что приводит к весьма существенной недооценке нетто-премии, т.к. $\hat{p}^{*} = \hat{z} \cdot \overline{X}^{*} = 11\ 742.14$.

2. Доверительные интервалы для чистых нетто-премий по полным и цензурированным данным

Главная идея метода заключается в замене теоретического исследования вычислительным экспериментом [14,15]. Предполагается, что существует множество данных, содержащих N наблюдений, которые образуют генеральную совокупность. Из генеральной совокупности случайным образом формируются выборки объема N с равными вероятностями ($1/N$) для каждого наблюдения. В итоге получают M выборок, по каждой из которых оценивается интересующий параметр исследуемой величины. После выполненных действий полученные оценки усредняют.

Значительным преимуществом бутстреп-процедуры является отсутствие требований о наличии информации о виде закона распределения исследуемой случайной величины, и в этом смысле она может рассматриваться как непараметрическая.

В данной работе с помощью бутстреп-моделирования производилось сравнение распределений итоговых оценок средних, строились доверительные интервалы для нетто-премии. Параметр моделирования $M = 50000$.

На рис. 3 представлены сглаженные бутстреп-гистограммы, построенные на основе разных выборок: $p(x)$ – по полной выборке, $p^{C}(x)$ – по цензурированной и $p^{*}(x)$ – по цензурированной как по полной, без учёта цензурирования. Очевидно, что игнорирование цензурирования привело к смещению плотности $p^{*}(x)$ влево, т.е., полученное с помощью бутстреп-метода среднее значение, равное 57 868.88 у.е./ед., существенно меньше, чем истинное 73 086.59 у.е./ед. (по полной выборке). Плотности $p(x)$ и $p^{C}(x)$ практически совпали, что говорит о высоком качестве оценки, построенной по цензурированной выборке, которая за счёт большого объема наблюдений ($N = 820$) и относительно небольшой доли цензурирования (20.46%) позволила получить практически истинное значение среднего, а именно, 73 008.26 у.е./ед.

Далее на основе бутстреп-выборок были построены доверительные интервалы для среднего ущерба с уровнем доверия 0.95:

1) по полной выборке: [60 941;88 652] у.е./ед.,

2) по цензурированной выборке: [60 904.56;88 536.62] у.е./ед.,

3) по цензурированной выборке без учёта цензурирования [48 413;71 335] у.е./ед.

Соответствующие доверительные интервалы с уровнем доверия 0.95 для чистых нетто-премий составили:

1) по полной выборке: [11 152.21;16 223.32],

2) по цензурированной выборке: [11 145.53;16 202.21],

3) по цензурированной выборке без учёта цензурирования [8 859.58;13 054.31].

Заметим, что доверительные интервалы для полной и цензурированной выборок практически совпали, а доверительный интервал, построенный без учёта цензурирования, сместился влево.

Таким образом, нельзя оценивать среднее значение по цензурированной выборке без учёта цензурирования, используя классическую формулу среднего арифметического, т.к. она дает недостоверный результат и приводит к неверному значению чистой нетто-премии.

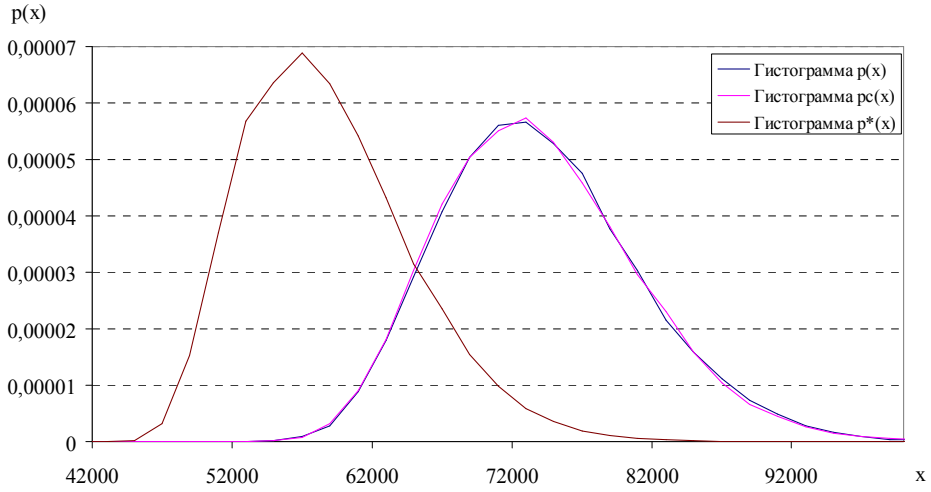


Рис. 3. Сглаженные бутстреп-гистограммы $p(x)$, $p^{C}(x)$ и $p^{*}(x)$

Заключение

В данной работе показано, что различные способы расчёта среднего значения по цензурированной выборке существенно влияют на точность оценивания чистой нетто-премии. На примере реальных данных о страховых выплатах по комплексному автомобильному страхованию, кроме ответственности (КАСКО) некоторой страховой компании показано, что игнорирование факта цензурирования существенно снижает значение чистой нетто-премии, несмотря на большой объем наблюдений и достаточно низкую долю цензурирования.

ЛИТЕРАТУРА

1. *Миронкина Ю.Н., Сорокин А.С.* Основы актуарных расчетов. – М.: Изд. Центр ЕАОИ, 2011. – 284 с.

2. *Фалин Г.И., Фалин А.И.* Теория риска для актуариев в задачах. – М.: Мир, Научный мир, 2004.

3. *Зенкова Ж.Н., Крайнова Е.А.* Оценка чистой нетто-премии с использованием дополнительной информации о квантиле функции распределения // Бизнес-информатика. – 2017. – № 4 (42). С. 55–63.

4. *Скрипник В.М., Назин А.Е., Благовещенский Ю.Г.* Анализ надежности технических систем по цензурированным выборкам. М.: Радио и связь, 1988. – 184 с.

5. *Дмитриев Ю.Г., Зенкова Ж.Н.* Об одной оценке симметричной функции распределения по цензурированной выборке // Материалы IV Всероссийской научно-практической конференция "Информационные технологии и математическое моделирование". Анжеро-Судженск. 2005. Ч.2. – Томск: Изд-во Том. ун-та, 2005. – С. 8–10.

6. *Дмитриев Ю.Г., Зенкова Ж.Н.* Об оценивании симметричной функции распределения по полным наработкам цензурированной выборки // Обозрение прикладной и промышленной математики. – 2005. – Т. 12. – Вып. 2. – С. 349–350.

7. *Зенкова Ж.Н., Муравлева М.А.* Расчет стоимости оборотных средств предприятия по цензурированным интервалом данным // Логистические системы в глобальной экономике [Электронный ресурс]: материалы Междунар. науч.-практ. конф. (3-4 марта 2014 г., Красноярск). Вып. 1. Электрон. сб. Сиб. гос. аэрокосмич. ун-т. Красноярск, 2014. – С. 103–107.

8. *Зенкова Ж.Н., Краковецкая И.В.* Непараметрическая оценка Тёрнбулла для интервально-цензурированных данных в маркетинговом исследовании спроса на биоэнергетические напитки // Вестник томского государственного университета, Управление, вычислительная техника и информатика. – 2013. – № 3(24). – С. 64–69.

9. *Зенкова Ж.Н., Макеева О.Б.* Применение методов обработки цензурированных данных при анализе оборачиваемости // Вестник науки Казахского агротехнического университета им. С. Сейфуллина, Астана, Казахстан. – 2014. –№3 (82). – С. 21–30.

10. *Зенкова Ж.Н., Муравлева М.А.* Асимптотическая несмещенность оценки функции распределения по однократно интервалом цензурированным данным с привлечением информации о симметрии // Новые информационные технологии в исследовании сложных структур: материалы X Российской конференции с международным участием. – Томск, Издательский Дом Томского государственного университета, 2014. – С. 114–115.

11. *Зенкова Ж.Н., Серебренникова К.А.* Расчёт чистой нетто-премии по однократно справа цензурированной выборке // Материалы VI Международной молодежной научной конференции Математическое и программное обеспечение информационных, технических и экономических систем. Томск, Изд-во ТГУ, 2018. – С. 112-116.

12. *Zenkova Z.N., Kolychev N.* Censored data analysis with using information about symmetry of distribution // Business inform. Ukraine. – 2011. – №7 (1). – P. 13–15.

13. *Боровков А.А.* Математическая статистика. – Новосибирск: Наука, 1997.

14. *Efron B.* Bootstrap Methods: Another Look at the Jackknife. –The Annals of Statistics, 1979.

15. *Efron B.* Nonparametric estimates of standard error: The jackknife, the bootstrap and other methods. – Biometrika, 1981.

НЕПАРАМЕТРИЧЕСКИЙ БКР-КРИТЕРИЙ НЕЗАВИСИМОСТИ

В.П. Шуленин, А.В. Пупков

Томский государственный университет

Введение

Критерий проверки гипотезы независимости предложили Блюм, Кифер и Розенблат в работе [1]. Для краткости будем его называть БКР-критерий независимости (см. также [4, с. 231]). При малых (конечных) объемах двумерной выборки $(X_1,Y_1),...,(X_n,Y_n)$, вычисление статистики этого критерия проводится с помощью наглядной графической процедуры, путём нанесения наблюдённых точек $(x_1,y_1),...,(x_n,y_n)$ на плоскость прямоугольной системы координат (см. рис.1) и формулу (2). Если объем n двумерной выборки достаточно большой (в асимптотике $n\to\infty$), то использование этой процедуры является утомительным и трудно реализуемым. Практическое применение этого критерия также ограничено тем фактом, что точное распределение статистики БКР-критерия при справедливости основной гипотезы независимости и при *конечных объёмах n* двумерной выборки $(X_1,Y_1),...,(X_n,Y_n)$ пока не найдено и отсутствуют таблицы критических значений статистики этого критерия. Однако указанные авторы получили асимптотическое распределение *нормированной статистики* БКР-критерия и привели её критические значения (см. табл. 1), а также таблицы А26 в [3, с. 431] и [2, с. 395]. Использование этих асимптотических результатов и возникающие при этом трудности вычисления *нормированной статистики* при достаточно больших объемах выборки, устраняются путем предложенной в данной работе программы, которая приводиться в приложении, что позволяет успешно и *корректно* применять БКР-критерий на практике.

Для описания БКР-критерия независимости уточним статистическую модель наблюдений, проверяемые гипотезы, формулу вычисления нормированной тестовой статистики, её критическую область и правило принятия решений.

1. БКР-критерий независимости

Статистическая модель наблюдений. Пусть (X,Y) – двумерная случайная величина с непрерывной (*неизвестной*) двумерной ф.р. $F_{XY}(x,y)$, $(x,y)\in R^2$, и $F_X(x)$, $F_Y(y)$ – одномерные функции распределения случайных величин X и Y. Имеется двумерная выборка $(X_1,Y_1),...,(X_n,Y_n)$ в виде *n независимых* пар случайных величин с одинаковой ф.р. $F_{XY}(x,y)$, $(x,y)\in R^2$ (т.е. $(X_1,Y_1),...,(X_n,Y_n)$ – н.о.р. двумерные случайные величины). Требуется проверить гипотезу независимости случайных величин X и Y.

Проверяемые гипотезы. Основная гипотеза независимости случайных величин X и Y записывается в виде

$$H_0 : F_{XY}(x,y) = F_X(x)F_Y(y), \tag{1}$$

она проверяется против альтернативы $H_1 : F_{XY}(x,y) \neq F_X(x)F_Y(y)$.

Статистика критерия. Для вычисления статистики БКР-критерия определим величину B_n в виде

$$B_n = n^{-5}\sum_{j=1}^{n}\left[N_1(j)N_4(j) - N_2(j)N_3(j)\right]^2, \tag{2}$$

где $N_i(j)$, $i=\overline{1,4}$ – число двумерных точек (x,y) в i-м, ($i=\overline{1,4}$) квадранте координатной плоскости системы с осями, параллельными осям исходной системы O_{XY} и с центром в точке $\left(X_j,Y_j\right)$, $j=\overline{1,n}$. Точнее, числа $N_i(j)$, $i=\overline{1,4}$ определяются следующим образом:

$N_1(j)$ – число выборочных пар $\left(X_1,Y_1\right),...,\left(X_n,Y_n\right)$, лежащих в области $R_1(j)=\left\{(x,y): x\le X_j, y\le Y_j\right\}$, $j=\overline{1,n}$;

$N_2(j)$ – число выборочных пар $\left(X_1,Y_1\right),...,\left(X_n,Y_n\right)$, лежащих в области $R_2(j)=\left\{(x,y): x> X_j, y\le Y_j\right\}$, $j=\overline{1,n}$;

$N_3(j)$ – число выборочных пар $\left(X_1,Y_1\right),...,\left(X_n,Y_n\right)$, лежащих в области $R_3(j)=\left\{(x,y): x\le X_j, y> Y_j\right\}$, $j=\overline{1,n}$;

$N_4(j)$ – число выборочных пар $\left(X_1,Y_1\right),...,\left(X_n,Y_n\right)$, лежащих в области $R_4(j)=\left\{(x,y): x> X_j, y> Y_j\right\}$, $j=\overline{1,n}$.

Для наглядности определения этих чисел см. рис. 1, на котором приведены реализации $\left(x_j,y_j\right)$ двумерных наблюдений и числа $N_i(j)$, $i=\overline{1,4}$, $j=\overline{1,20}$, при объеме выборки $n=20$.

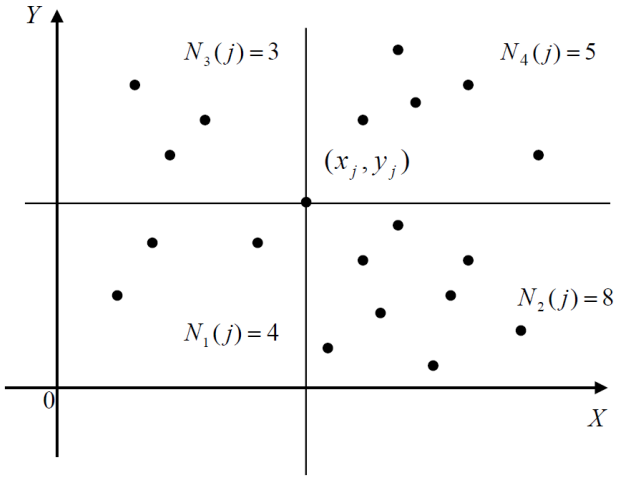


Рис. 1. Двумерные точки $\left(x_j, y_j\right)$, для определения чисел $N_i(j)$, $i=\overline{1,4}$, $j=\overline{1,20}$

Нормированный вариант статистики БКР-критерия, для которой найдено асимптотическое распределение (5), определяется в виде

$$T_n\left(\vec{X},\vec{Y}\right)=\frac{\pi^4}{2}nB_n=\frac{\pi^4}{2n^4}\sum_{j=1}^{n}\left[N_1(j)N_4(j)-N_2(j)N_3(j)\right]^2. \qquad (3)$$

Критическая область статистики. Асимптотическая критическая область уровня значимости α определяется в виде

$$\mathrm{X}_{1\alpha}=\left\{\left(\vec{x},\vec{y}\right): T_n\left(\vec{x},\vec{y}\right)\geq F^{-1}(1-\alpha)\right\}, \qquad (4)$$

где $F^{-1}(p)$, $0<p<1$ – квантильная функция для асимптотической функции распределения

$$F(z)=\lim_{n\to\infty}P\left\{\frac{\pi^4 n}{2}B_n\leq z \mid H_0\right\}, \qquad (5)$$

которая найдена (см. [2, с. 394]) с помощью обращения характеристической функции

$$M\{\exp(izB)\}=\prod_{j,k=1}^{\infty}\left[1-\frac{2iz}{\pi^4 j^2 k^2}\right]^{-1/2}$$

предельного распределения nB_n. Отметим, что квантили $F^{-1}(p)$, $0<p<1$ асимптотического распределения нормированной статистики (3) при справедливости гипотезы независимости (1) приведены в табл. 1, см. также [2, с.395].

Таблица 1

Асимптотические критические значения статистики (3) для БКР-критерия

$\alpha=0.10$	$\alpha=0.05$	$\alpha=0.01$
2.286	2.844	4.230

Критерий. Согласно (4), гипотеза независимости вида (1) отвергается при больших значениях статистики (3). Например, при $\alpha=0.05$ гипотезу независимости вида (1) следует отклонить при выполнении неравенства $T_n\left(\vec{x},\vec{y}\right)\geq 2.844$.

Замечание 1. Авторы БКР-критерия рекомендуют при малых объемах выборки (например, при $n<10$) использовать при изучении предельного распределения не nB_n, а $(n-1)B_n$. Следовательно, для малых объемов выборки, нормированный вариант статистики БКР-критерия записывается в виде

$$T_n^*\left(\vec{X},\vec{Y}\right)=\frac{\pi^4(n-1)}{2n^5}\sum_{j=1}^{n}\left[N_1(j)N_4(j)-N_2(j)N_3(j)\right]^2. \qquad (6)$$

Таким образом, при малых объемах выборки, БКР-критерий записывается в виде следующего правила:

отклонить H_0 на уровне значимости α, если $T_n^*\left(\vec{x},\vec{y}\right)\geq F^{-1}(1-\alpha)$,

принять H_0, если $T_n^*\left(\vec{x},\vec{y}\right)<F^{-1}(1-\alpha)$,

где критические значения $F^{-1}(1-\alpha)$ табулированы в [2, с. 395], см. также табл. 1.

Пример 1. Статистические данные взяты из работы [3, с. 245]. В эксперименте изучалось соотношение между запасами свободного пролина и содержанием коллагена в печени человека. Измерялся признак X – общее содержание коллагена (мг/г сухого вещества печени) и признак Y – свободный пролин (миллимоль/г сухого вещества печени). Объём выборки $n=7$.

Таким образом, в эксперименте изучается двумерная случайная величина (X,Y). Объём выборки $n=7$, и наблюдённые данные в виде реализация $\left(x_1,y_1\right),...,\left(x_7,y_7\right)$ выборки $\left(X_1,Y_1\right),...,\left(X_7,Y_7\right)$ приведены в табл. 2.

Таблица 2

Результаты опыта по изучению двумерной случайной величины (X,Y)

j	1	2	3	4	5	6	7
X	7.1	7.1	7.2	8.3	9.4	10.5	11.4
Y	2.8	2.9	2.8	2.6	3.5	4.6	5.0

Требуется, используя БКР-критерий, проверить гипотезу независимости случайных величин X и Y, которая записывается в виде $H_0: F_{XY}(x,y)=F_X(x)F_Y(y)$, она проверяется против альтернативы $H_1: F_{XY}(x,y)\neq F_X(x)F_Y(y)$.

Вычисление статистики БКР-критерия. В данном примере объем выборки $n=7$, поэтому, используя графическую процедуру, определяем числа $N_i(j)$, $i=\overline{1,4}$, $j=\overline{1,7}$ выборочных пар $\left(x_j,y_j\right)$ распределенных по областям $R_1(j)$, $R_2(j)$, $R_3(j)$, $R_4(j)$, $j=\overline{1,7}$.

При $j=1$, $\left(x_1,y_1\right)=(7{,}1,\ 2{,}8)$ получаем: $N_1(1)=1$, $N_2(1)=2$, $N_3(1)=1$, $N_4(1)=3$, $\left[N_1(1)N_4(1)-N_2(1)N_3(1)\right]^2=[1\cdot3-2\cdot1]^2=1^2=1$.

При $j=2$, $\left(x_2,y_2\right)=(7{,}1,\ 2{,}9)$ получаем: $N_1(2)=2$, $N_2(2)=2$, $N_3(2)=0$, $N_4(2)=3$, $\left[N_1(2)N_4(2)-N_2(2)N_3(2)\right]^2=[2\cdot3-2\cdot0]^2=6^2=36$.

…….

При $j=7$, $\left(x_7,y_7\right)=(11{,}4,\ 5{,}0)$ получаем: $N_1(7)=7$, $N_2(7)=0$, $N_3(7)=0$, $N_4(7)=0$, $\left[N_1(7)N_4(7)-N_2(7)N_3(7)\right]^2=[7\cdot0-0\cdot0]^2=0$.

Вычисленное по наблюденным данным табл. 2 значение $T_n^*\left(\vec{x},\vec{y}\right)$ статистики $T_n^*\left(\vec{X},\vec{Y}\right)$ вида (6) равно

$$T_n^*\left(\vec{x},\vec{y}\right)=\frac{\pi^4(n-1)}{2n^5}\sum_{j=1}^{n}\left[N_1(j)N_4(j)-N_2(j)N_3(j)\right]^2=$$

$$=\frac{\pi^4\cdot6}{2\cdot7^5}[1+36+25+9+100+36+0]=\frac{97{,}4091\cdot6}{2\cdot16807}[207]=3{,}60.$$

БКР-критерий. При уровне значимости $\alpha = 0.05$ критическое значение $F^{-1}(1-\alpha) = F^{-1}(0.95) = 2.844$ (см. табл. 1). Т.к. выполняется неравенство

$$\left\{T_n^*\left(\vec{x},\vec{y}\right) = 3.60\right\} > \left\{F^{-1}(0.95) = 2.844\right\},$$

гипотеза H_0 о независимости случайных величин X и Y отвергается на уровне значимости $\alpha = 0.05$. Таким образом, результаты эксперимента (см. табл. 2), указывают на наличие зависимости случайных величин X и Y.

Используя таблицу А26 в [3, с. 431], определяем Р-значение БКР-критерия, которое характеризует достигнутый (наименьший) уровень значимости отклонения гипотезы H_0 при наблюденном значении $T_n^*\left(\vec{x},\vec{y}\right) = 3.60$ статистики $T_n^*\left(\vec{X},\vec{Y}\right)$:

$$\text{Р-значение} = P\left(T_n^*\left(\vec{X},\vec{Y}\right) \geq T_n^*\left(\vec{x},\vec{y}\right) \mid H_0\right) = P\left(T_n^*\left(\vec{X},\vec{Y}\right) \geq 3.60\right) = 0.0205,$$

что также приводит к отклонению гипотезы независимости H_0.

Замечание 2. Отметим, что БКР- критерий асимптотически эквивалентен непараметрическому (свободному от распределения) критерию независимости Хёфдинга, который описан в [3, с. 242], и для которого табулирована вероятность верхнего хвоста распределения статистики при справедливости нулевой гипотезы для *конечных* объемов выборки $n = 5(1)9$, см. табл. А25 в [3, с. 429].

2. Описание программы вычисления статистики БКР-критерия

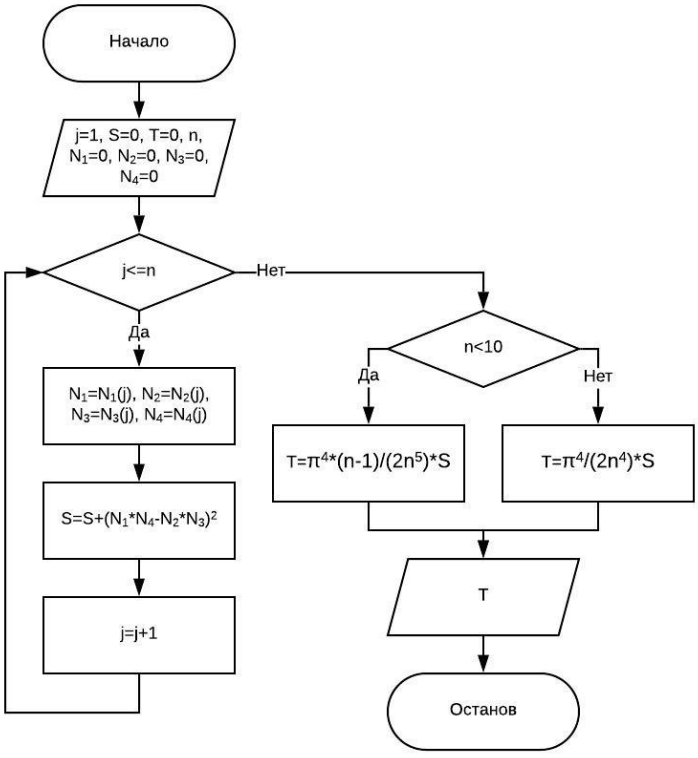


Рис. 1. Блок-схема программы вычисления статистики БКР-критерия

Заключение

В работе отмечаются трудности вычисления тестовой статистики БКР-критерия независимости при больших объемах двумерной выборки, которые устраняются путем

использования предложенной в работе программы для вычисления нормированной тестовой статистики, что позволяет использовать асимптотические результаты работы [1] и позволяет успешно и корректно применять БКР-критерий на практике.

ЛИТЕРАТУРА

1. *Blum J.R., Kiefer J., Rosenblatt M.* Distribution free tests of independence based on the sample distribution function // Ann. Math. Statist. – 1961. – N 21. P. 485–498.

2. *Оуэн Д.Б.* Сборник статистических таблиц. Изд. 2-е испр. – М.: ВЦ АН СССР, 1973. – 586 с.

3. *Холлендер М., Вулф Д.* Непараметрические методы статистики.– М.: Финансы и статистика, 1983. – 518 с.

4. *Шуленин В.П.* Дополнительные главы математической статистики (курс лекций). – Томск: Изд-во НТЛ, 2018. – 516 с.

АСИМПТОТИЧЕСКИ ЭФФЕКТИВНОЕ ОЦЕНИВАНИЕ В НЕПАРАМЕТРИЧЕСКОЙ ГЕТЕРОСКЕДАСТИЧНОЙ РЕГРЕССИИ*

С.С. Перелевский, Е.А. Пчелинцев

Томский государственный университет

Введение

Рассмотрим задачу асимптотически эффективного оценивания неизвестной функции $S(\cdot)$ гетероскедастичной регрессии, описываемую уравнением

$$y_j = S\left(x_j\right) + \sigma_j \xi_j,\ 1 \le j \le n, \tag{1}$$

где $x_j = j/n$, $\left(\xi_j\right)_{1\le j\le n}$ – последовательность независимых одинаково распределённых случайных величин с нулевым средним и дисперсией равной единице, $\left(\sigma_j\right)_{1\le j\le n}$ – коэффициенты волатильности, зависящие от x_j и функции S, n – число наблюдений.

Понятие асимптотической эффективности связано с оптимальной скоростью сходимости минимаксного риска, т.е. важным вопросом в результатах оптимальности является изучение точной асимптотики минимаксного риска, определенного ниже в (5). Проблема асимптотического непараметрического оценивания в модели гетероскедастичной регрессии изучались Эфроймовичем [1] и Пинскером [2]. Для доказательства асимптотической эффективности процедуры, необходимо показать, что ее асимптотический квадратичный риск совпадает с нижней границей, определяемой константой Пинскера [3]. В данной работе поставленная задача решается с использованием подхода, основанного на методах выбора модели и точных оракульных неравенств. Напомним, что метод выбора моделей появился в пионерских работах Акайке [4] и Мэллоуза [5], которые предложили ввести пенализационное слагаемое в критерий максимального правдоподобия. Далее, Барон, Бирже и Массар [6], Массар [7] и Кнайп [8] развили этот метод для получения неасимптотических оракульных неравенств в непараметрических регрессионных моделях с гауссовскими шумами в дискретном времени. К сожалению, этот метод не может быть применен в нашем случае для доказательства эффективности, поскольку коэффициент в основном слагаемом, в полученных оракульных неравенствах больше единицы. По этой причине в данной работе для эффективного оценивания функции в модели (1) применяется метод построения оракульных неравенств, развитый в [9]. В данной работе исследуется свойство асимптотической эффективности улучшенной процедуры выбора модели, предложенной в [10], для оценивания функции S в (1).

* Работа выполнена в рамках государственного задания Министерства науки и высшего образования РФ (проект № 2.3208.2017/4.6).

1. Оракульное неравенство для квадратического риска адаптивной процедуры выбора модели

Для оценивания неизвестной функции S в модели (1) предлагается использовать улучшенные взвешенные оценки МНК, определенные в [10],

$$S_\lambda^*(x) = \sum_{j=1}^{n} \lambda(j)\theta_{j,n}^* \phi_j(x), \tag{2}$$

где $\theta_{j,n}^* = \left(1 - \frac{c_n}{\|\tilde{\theta}_n\|}\right)\hat{\theta}_{j,n}$, $\|\tilde{\theta}_n\|^2 = \sum_{j=1}^{d} \hat{\theta}_{j,n}^2$, $\hat{\theta}_{j,n} = \theta_j + \frac{1}{n}\xi_{j,n}$, $\xi_{j,n} = \frac{1}{\sqrt{n}}\sum_{l=1}^{n} \sigma_l \xi_l \phi_j(x_l)$. Здесь коэффициент $d \approx n^\delta / \ln n$, $0 < \delta < 1$, $c_n \approx d/n$. Ясно, что $c_n \to 0$ при $n \to \infty$.

Метод выбора модели заключается в описании правила выбора процедуры из семейства оценок (2), т.е. нужно выбрать вектор весовых коэффициентов $\lambda \in \Lambda$ в (2). Для этого естественно минимизировать эмпирическую квадратическую ошибку вида

$$Err_n(\lambda) = \|S_\lambda^* - S\|_n^2 = \frac{1}{n}\sum_{l=1}^{n}\left(S_\lambda^*(x_l) - S(x_l)\right)^2 = \sum_{j=1}^{n}\lambda^2(j)\theta_{j,n}^{*2} - 2\sum_{j=1}^{n}\lambda(j)\theta_{j,n}^*\theta_j + \|S\|^2.$$

Однако, поскольку во втором слагаемом коэффициенты θ_j неизвестны, то минимизация этого выражения, не приводит к нахождению коэффициентов λ. Необходимо величины $\theta_{j,n}^*\theta_j$ заменить некоторыми оценками. В качестве таких оценок предлагаются $\bar{\theta}_{j,n} = \theta_{j,n}^*\hat{\theta}_{j,n} - \frac{1}{n}\hat{\varsigma}_n$, где $\hat{\varsigma}_n$ – некоторая оценка интегрированной дисперсии шума $\varsigma_n = \frac{1}{n}\sum_{j=1}^{n}\sigma_l^2$. При осуществлении этой замены в эмпирической ошибке нужно заплатить "штраф". Определим платежную функцию как

$$J_n(\lambda) = \sum_{j=1}^{n}\lambda^2(j)\theta_{j,n}^{*2} - 2\sum_{j=1}^{n}\lambda(j)\bar{\theta}_{j,n} + \rho\hat{P}_n(\lambda). \tag{3}$$

Здесь $0 < \rho < 1$ – некоторый положительный коэффициент, зависящий от n и такой, что $\rho(n) \to 0$ при $n \to \infty$ и пенализационное слагаемое $\hat{P}_n(\lambda) = \frac{|\lambda|^2 \hat{\varsigma}_n}{n}$, $|\lambda|^2 = \sum_{j=1}^{n}\lambda^2(j)$. Отметим, что в случае, когда последовательность $(\sigma_j)_{1\le j\le n}$, известна, то $P_n(\lambda) = \frac{|\lambda|^2 \varsigma_n}{n}$.

Полагая $\lambda^* = \arg\min_{\lambda\in\Lambda} J_n(\lambda)$, определим процедуру выбора модели равенством

$$S^* = S_{\lambda^*}^* \tag{4}$$

Заметим, что λ^* существует, поскольку множество Λ конечно. В случае, когда λ^* не единственное, берем любое из них.

Обозначим через P_n семейство распределений p в $\mathbb{R}^n$ векторов $(\xi_1,...,\xi_n)$ в модели (1) таких, что компоненты ξ_j совместно независимы, центрированы с единичной дисперсией и $\max_{1\le k\le n} \mathrm{E}(\xi_k^4) \le l_n^*$, где l_n^* – медленно возрастающая последовательность, т.е. $\lim_{n\to\infty} l_n = \infty$, $\lim_{n\to\infty} l_n / n^\delta = 0$ для любого $\delta > 0$. Легко видеть, что для любого $n \ge 1$ центрированное гауссовское распределение в $\mathbb{R}^n$ с единичной ковариационной матрицей принадлежит семейству P_n. Обозначим это распределение через q. Для оценки S^* определим следующий робастный среднеквадратический риск

$$\Re\left(S^*, S\right) = \sup_{p \in P_n} \mathbf{E}_{S,p} \left\| S^* - S \right\|_n^2, \tag{5}$$

где $\mathbf{E}_{S,p}$ – математическое ожидание относительно распределения $\mathbf{P}_{S,p}$ по наблюдениям $\left(y_1, ..., y_n\right)$ с фиксированной функцией S и фиксированным распределением $p \in P_n$ случайных величин $\left(\xi_j\right)_{1 \le j \le n}$ в модели (1).

В [10] получено следующее оракульное неравенство, определяющее неасимптотическую верхнюю границу для среднеквадратического риска (5) процедуры оценивания (4).

Теорема 1. Пусть наблюдения описываются уравнением (1). Тогда для любого $0 < \rho < 1/3$ среднеквадратический риск (5), предложенной процедуры выбора модели (4) для функции S удовлетворяет оракульному неравенству: $\Re\left(S^*, S\right) \le \dfrac{1-\rho}{1-2\rho} \min_{\lambda \in \Lambda} \Re\left(S_\lambda^*, S\right) + \dfrac{1}{n\rho} \Psi_n$, где $\lim_{n \to \infty} \Psi_n / n^\delta = 0$ для любого $\delta > 0$.

2. Асимптотическая эффективность

Пусть $C(\mathbb{R})$ - множество 1-периодических k раз дифференцируемых $\mathbb{R} \to \mathbb{R}$ функций. Предположим, что S принадлежит следующему множеству:

$$W_r^k = \left\{ f \in C_{per,1}^k(\mathbb{R}) : \sum_{j=0}^k \left\| f^{(j)} \right\|^2 \le r \right\},$$

где $\|\cdot\|$ обозначает норму в $L_2[0,1]$. Кроме того, предположим, что $r > 0$ и $k \ge 1$ – неизвестные параметры. Заметим, что множество W_r^k может быть представлено как эллипс в $L_2[0,1]$, т.е.

$$W_r^k = \left\{ f \in L_2[0,1] : \sum_{j=1}^\infty a_j \theta_j^2 \le r \right\},\ \theta_j = \left(f, \phi_j\right) = \int_0^1 f(t)\phi_j(t)dt,\ a_j = \sum_{l=0}^k \left\| \phi_j^l \right\|^2 = \sum_{i=0}^k \left(2\pi[j/2]\right)^{2i}.$$

Обозначим Σ_n – множество всех оценок $\hat{S}_n$ измеримых относительно наблюдений $\left(y_1, \ldots, y_n\right)$. Предположим, что коэффициенты волатильности $\left(\sigma_j(S)\right)_{j \ge 1}$ удовлетворяют следующим условиям.

C_1) $\sigma_j(S) = g\left(x_j, S\right)$ для некоторой неизвестной функции, $g : [0,\ 1] \times L_1[0,\ 1] \to \mathbb{R}_+$ которая квадратично интегрируема по x и такая, что

$$\lim_{n \to \infty} \sup_{W_r^k} \left| \frac{1}{n} \sum_{j=1}^n g^2\left(x_j, S\right) - \varsigma(S) = 0 \right|,$$

где $\varsigma(S) = \int_0^1 g^2\left(x_j, S\right)dx$. Кроме того, $g_* = \inf_{0 \le x \le 1} \inf_{W_r^k} g^2(x, S) > 0$, $\varsigma(S) < \infty$.

C_2) Для некоторого $x \in [0,\ 1]$, оператор $g^2(x, \cdot) : C[0,\ 1] \to \mathbb{R}$ дифференцируема в смысле Фреше при любой фиксированной функции f_0 из $C[0,\ 1]$, т.е. для некоторого f из некоторой окрестности f_0 в $C[0,\ 1]$,

$$g^2(x, f) = g^2\left(x, f_0\right) + \mathrm{L}_{x, f_0}\left(f - f_0\right) + \Upsilon\left(x, f_0, f\right),$$

где производная по Фреше $\mathrm{L}_{x, f_0}\left(f - f_0\right) : C[0,1] \to \mathbb{R}$ – ограниченный линейный оператор, а остаточный член $\Upsilon\left(x, f_0, f\right)$ для каждого $x \in [0,\ 1]$, удовлетворяет следующему свойству:

$$\lim_{\|f-f_0\|_\infty \to 0} \frac{\|\Upsilon(x, f_0, f)\|}{\|f - f_0\|_\infty} = 0,\ \|f\|_\infty = \sup_{0\le t\le 1} |f(t)|.$$

C_3) Существует некоторая положительная постоянная C^* такая, что для некоторой функции S из $C[0,1]$, оператор $L_{x,S}$, определенный в C_2), удовлетворяет следующему неравенству для любой функции f из $C[0,1]$:

$$|L_{x,S}(f)| \le C^* \left(|S(x) f(x)| + |f|_1 + \|S\| \|f\| \right),\ |f|_1 = \int_0^1 |f(t)| dt.$$

C_4) Функция $g_0(\cdot) = g(\cdot, S_0)$, соответствующая $S_0 \equiv 0$, является непрерывной на интервале $[0,1]$. Кроме того,

$$\lim_{\delta \to \infty} \sup_{0\le x\le 1} \sup_{\|S\|_\infty \le \delta} |g(x,S) - g(x,S_0)| = 0.$$

Справедливы следующие асимптотические результаты.

Теорема 2. Пусть для модели (1) выполнены условия C1) – C4). Тогда робастный риск (5) для всех $k \ge 1$ и $r > 0$ удовлетворяет следующему асимптотическому неравенству

$$\liminf_{n\to\infty} n^{\frac{2k}{2k+1}} \inf_{\hat{S}_n \in \Sigma_n} \sup_{S \in W_r^k} \frac{\Re(\hat{S}_n, S)}{\gamma_k(S)} \ge 1,$$

где $\gamma_k(S) = \Gamma_k^* \left(\varsigma(S) \right)^{2k/(2k+1)}$, $\Gamma_k^* = \left((2k+1) r \right)^{1/(2k+1)} \left(k/(\pi(k+1)) \right)^{2k/(2k+1)}$ – константа Пинскера.

Теорема 3. Пусть для модели (1) выполнены условия C1) – C4). Тогда робастный риск (5) процедуры выбора модели S^* для всех $k \ge 1$ и $r > 0$ удовлетворяет следующей асимптотической верхней границе

$$\limsup_{n\to\infty} n^{\frac{2k}{2k+1}} \sup_{S \in W_r^k} \frac{\Re(S^*_{\lambda^*}, S)}{\gamma_k(S)} \le 1.$$

Следствие 1. Из теорем 2 и 3 следует, что процедура выбора модели $S^*_{\lambda^*}$, определенная в (4), является асимптотически эффективной, т.е.

$$\lim_{n\to\infty} n^{\frac{2k}{2k+1}} \sup_{S \in W_r^k} \frac{\Re(S^*_{\lambda^*}, S)}{\gamma_k(S)} = 1.$$

3. Численное моделирование

В этом разделе проиллюстрируем теоретически установленные результаты с помощью численного моделирования в среде Matlab. В качестве функции S в модели (1) выберем

$$S_1(x) = x \sin(2\pi x) + x^2 (1-x) \cos(4\pi x),\ S_2(x) = \begin{cases} 0.5 - |x - 0.5|, & 0 \le x \le 1, \\ 0, & \text{иначе}, \end{cases}$$

$(\xi_j)_{1\le j\le n}$ – последовательность независимых $N(0,1)$ случайных величин, коэффициенты волатильности $\sigma_j^2 = 2 + S(x_j)$. Для вычисления весовых коэффициентов $(\lambda_\alpha)_{\alpha \in A}$, определенных в [10], положим

$$k^* = 100 + \sqrt{\ln n},\ \varepsilon = \frac{1}{\ln n},\ m = [\ln^2 n],\ \rho = \frac{1}{3 + \ln^2 n},\ \omega_\alpha = 100 + \left(A_\beta t n \right)^{\frac{1}{2\beta+1}}.$$

В табл. 1 приведены результаты поведения эмпирических среднеквадратических рисков по $N = 1000$ реализациям процедур выбора модели (4).

Таблица 1

Эмпирические квадратические асимптотические риски

n	101	501	1001	2001	10001
$n^{\frac{2k}{2k+1}}\frac{R(S^*,S_1)}{\gamma_k(S_1)}$	8.1458	5.7257	3.0856	1.0009	0.9125
$n^{\frac{2k}{2k+1}}\frac{R(S^*,S_2)}{\gamma_k(S_2)}$	11.1257	6.1254	3.1587	1.0015	0.9321

Из табл. 1 видно, что с ростом числа наблюдений n нормированные эмпирические среднеквадратические риски стремятся к единице, что численно подтверждает следствие 1.

Далее на рисунках представлены графики истинных функций S_1 и S_2 (сплошная линия) и их оценок (4) при $n = 501; 1001; 10001$.

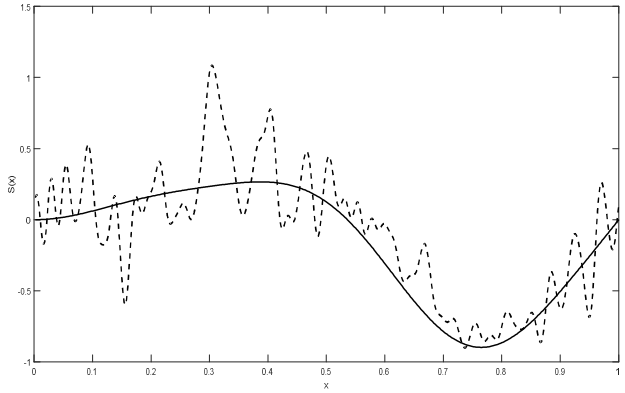


Рис. 1. Графики функции регрессии и ее оценок при n = 501

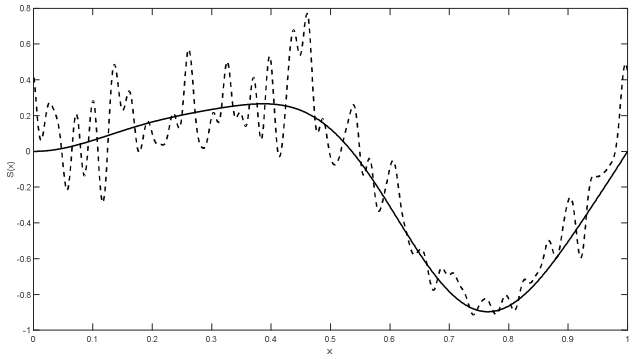


Рис. 2. Графики функции регрессии и ее оценок при n = 1001

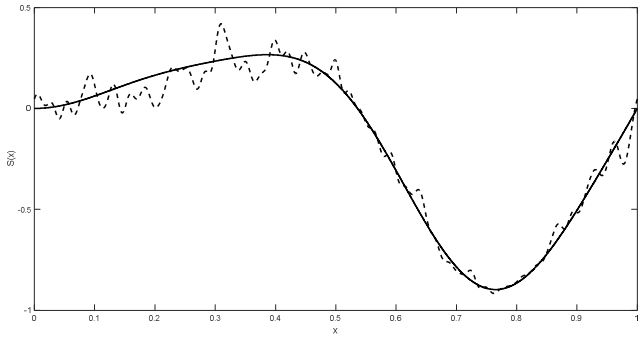


Рис. 3. Графики функции регрессии и ее оценок при $n = 10001$

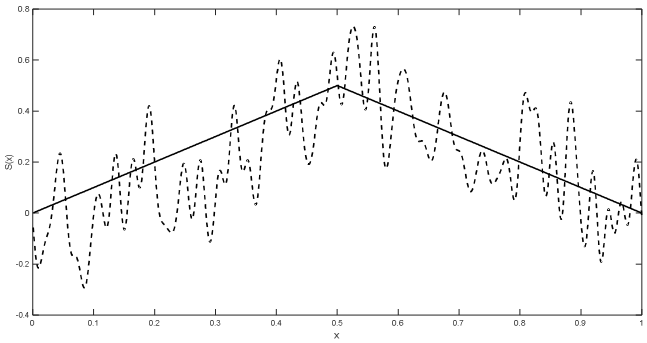


Рис. 4. Графики функции регрессии и ее оценок при $n = 501$

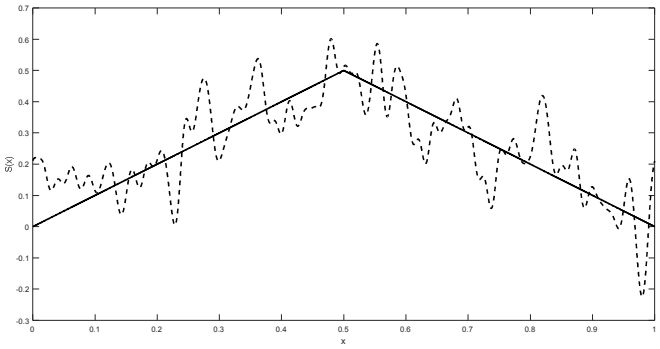


Рис. 5. Графики функции регрессии и ее оценок при $n = 1001$

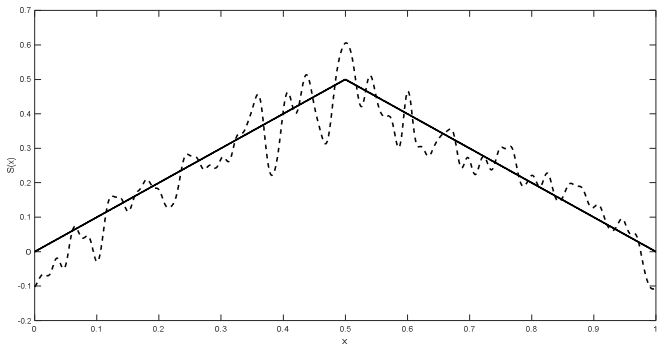


Рис. 6. Графики функции регрессии и ее оценок при $n = 10001$

ЛИТЕРАТУРА

1. *Efroimovich S.* Sequential design and estimation in heteroscedastic nonparametric regression // Sequential Analysis. – 2007. – N 26. – P. 3-25.

2. *Pinsker M.S.* Optimal filtration of square integrable signals in gaussian white noise // Problems Transimis. information. – 1981. – N 17. – P. 120-133.

3. *Galtchouk L., Pergamenshchikov S.* Asymptotic efficient sequential kernel estimates of the drift coefficient in ergodic diffusion processes // Statistical Inference for Stochastic Process. – 2006. – V. 9. – N 1. – P. 1–16.

4. *Akaike H.* A new look at the statistical model identification // IEEE Trans. on Automatic Control. – 1974. – V. 19. – N 7. – P. 716–723.

5. *Mallows C.* Some comments on Cp // Technometrics. – 1973. – V. 15. – P. 661–675.

6. *Barron A., Birge L., Massart P.* Risk bounds for model selection via penalization // Probab. Theory Relat. Fields. – 1999. – V. 113. – P. 301-415.

7. *Massart P.* A non-asymptotic theory for model selection // European Congress of Mathematics. Zurich: Eur. Math. Soc., 2005.

8. *Kneip A.* Ordered linear smoothers // Annals of Statistics. – 1994. – V. 22. – P. 835–866.

9. *Galtchouk L., Pergamenshchikov S.* Adaptive asymptotically efficient estimation in heteroscedastic nonparametric regression // Journal of the Korean Statistical Society. – 2009. – V. 38. – N 4. – P. 305–322.

10. *Пчелинцев Е.А., Перелевский С.С.* Адаптивное оценивание в гетероскедастичной непараметрической регрессии // Вестник ТГУ. Математика и механика. – 2019. – № 57. – С. 38–52.

ОБНАРУЖЕНИЕ РАЗЛАДКИ В ГАУССОВСКОМ ПРОЦЕССЕ AR(1) С НЕИЗВЕСТНЫМ ПАРАМЕТРОМ

А.В. Пупков

Томский государственный университет

Введение

В данной работе рассматривается подход к задаче обнаружения разладки в процессе AR(1) с неизвестным параметром после момента отклонения. Основная идея заключается в последовательной оценке параметра гауссовской авторегрессии по особой схеме. Используя эту оценку, можно получить статистику с неасимптотическим гауссовским распределением, дисперсию которой мы можем контролировать. В момент, когда происходит разладка процесса авторегрессии, меняется распределение полученной статистики. На основе этого изменения отслеживается разладка.

1. Постановка задачи

Рассмотрим процесс AR(1), с разладкой в момент времени ν, вида

$$x_k = \theta_0 x_{k-1} + \varepsilon_k,\ k = \overline{1, \nu - 1},$$

$$x_k = \theta x_{k-1} + \varepsilon_k,\ k \geq \nu,$$

где $x_0 = 0$, $\varepsilon_k \sim N(0,1)$, параметры принадлежат промежутку устойчивости процесса θ, $\theta_0 \in (-1,1)$.

Введем последовательность моментов остановки

$$\tau_i(h) = \inf\left\{ n > \tau_{i-1} : \sum_{k=\tau_{i-1}+1}^{n} x_{k-1}^2 \geq h \right\},\ \inf\{\varnothing\} = \infty, \tag{1}$$

где $\tau_0 = \tau_0(h) \equiv 0$. Далее, на каждом полученном интервале $\Delta\tau_i = (\tau_{i-1}, \tau_i]$, оценим значение параметра по следующей схеме

$$\hat{\theta}_i = \frac{1}{\tilde{h}_i} \sum_{k=\tau_{i-1}+1}^{\tau_i} \sqrt{\alpha_i(k)}\, x_{k-1} x_k, \tag{2}$$

где $\tilde{h}_i = \sum_{k=\tau_{i-1}+1}^{\tau_i} \sqrt{\alpha_i(k)}\, x_{k-1}^2$,

$$\alpha_i(k)=\begin{cases}1,\ \tau_{i-1}<k<\tau_i,\\ \beta,\ k=\tau_i,\end{cases}$$

причем β_i – корректирующий множитель, который определяется из выражения

$$\sum_{k=\tau_{i-1}+1}^{\tau_i-1} x_{k-1}^2+\beta_i x_{\tau_i-1}^2=h\,.$$

Теорема 1. Пусть $\left(\Delta\tau_i,\hat{\theta}_i\right)$, $\forall i>0$, $\tau_i<\nu$ – последовательная оценка параметра θ_0 исследуемого процесса на промежутке $\left(\tau_{i-1},\tau_i\right]$, определяемая формулами (1), (2) [1]. Тогда, если $\{\varepsilon_k\}_{k\ge1}$ – последовательность стандартных гауссовских случайных величин, то $\forall h>0$ и $\theta_0\in R$ выполняется $P_\theta\left\{\frac{\tilde{h}_i}{\sqrt{h}}\left(\hat{\theta}_i-\theta_0\right)\le z\right\}=\Phi(z)$, $\forall z\in R$, где $\Phi(z)=\frac{1}{\sqrt{2\pi}}\int_{-\infty}^{z}e^{-\frac{u^2}{2}}du$.

Рассмотрим статистики вида

$$t_i(h)=\frac{\tilde{h}_i}{h}\left(\hat{\theta}_i-\theta_0\right). \tag{3}$$

Поведение заданных статистик до и после разладки определяется выражением

$$t_i(h)=\begin{cases}Z_i,\ i<i_\nu,\\ Z_i+\Delta\theta+\zeta_i,\ i>i_\nu,\end{cases}$$

где

$$Z_i=\frac{1}{h}\sum_{k=\tau_{i-1}+1}^{\tau_i}\sqrt{\alpha_i(k)}x_{k-1}\varepsilon_k\sim N\left(0,\frac{1}{h}\right),\ \zeta_i=\left(\frac{\tilde{h}_i}{h}-1\right)(\theta-\theta_0)\to0,\ h\to\infty,\ \Delta\theta=(\theta-\theta_0),$$

i_ν – номер интервала $\Delta\tau_i$, в который произошла разладка, т.е. момент отклонения $\nu\in\Delta\tau_{i_\nu}$.

Сместим статистики вида (3). Поскольку малые изменения в значении параметра плохо уловимы, то можно установить минимальную разницу значений параметра до и после разладки, т.е. будем предполагать, что $|\theta-\theta_0|\ge L$.

Введем статистики вида

$$t_i^1(h)=t_i(h)-L,\ i=1,2,\ldots, \tag{4}$$

которые будем использовать в процедуре CUSUM, отслеживающей смещение параметра вверх, а статистики $t_i^2(h)=t_i(h)+L$, $i=1,2,\ldots$ – в процедуре, отслеживающей смещение вниз. Данное преобразование полезно, поскольку будет уменьшаться вероятность ложной тревоги, из-за смещения математического ожидания статистик, но будет увеличиваться среднее время запаздывания.

Введем следующие процедуры [2]:

$$N_L^1(h)=\inf\left\{n\ge1:\max_{1\le k\le n}\sum_{i=k}^{n}t_i^1(h)\ge c\right\} \tag{5}$$

и

$$N_L^2(h)=\inf\left\{n\ge1:\min_{1\le k\le n}\sum_{i=k}^{n}t_i^2(h)\le -c\right\} \tag{6}$$

для отслеживания скачка параметра вверх и вниз соответственно.

После введенных преобразований итоговая процедура будет иметь вид

$$\tilde{N}_1^2=\tilde{N}_1^2(h)=\min\left\{N_L^1(h),N_L^2(h)\right\}. \tag{7}$$

В следующем разделе рассмотрим теоретические выводы, связанные с предложенной моделью.

2. Среднее расстояние между ложными тревогами (нижняя граница)

Рассмотрим среднее расстояние между ложными тревогами в случае смещения параметра вверх ($\theta > \theta_0$). Необходимо получить нижнюю границу [3] для величины $E_0 N_L^1(h)$.

Стоит отметить, что статистики $t_i^1(h)$, $i = \overline{1, \nu - 1}$ обладают гауссовским распределением с математическим ожиданием ниже нуля и дисперсией, зависящей от параметра h, т.е. $t_i^1(h) \sim N\left(-L, \frac{1}{h}\right)$, $i = \overline{1, \nu - 1}$. В результате несложных преобразований имеем $E_0\left(e^{t_{i+1}^1(h)} \mid F_{\tau_i(h)}\right) = \exp\left\{-L + \frac{1}{2h}\right\}$, где $L > 0$ – минимальная разница между параметрами до и после разладки. Для того, чтобы $E_0\left(e^{t_{i+1}^1(h)} \mid F_{\tau_i(h)}\right) < 1$, необходимо, чтобы $\left(-L + 1/2h\right) < 0$, следовательно

$$h_* > \frac{1}{2L}. \tag{8}$$

Рассмотрим стохастическую последовательность [4] $X(h) = \left(X_n(h), F_{\tau_n(h)}\right)$,

где $X_n(h) = \exp\left(\sum_{k=1}^{n} t_k^1(h)\right)$. Последовательность $X(h)$ есть супермартингал при выполнении условия (8), поскольку для всех $n \geq 1$ выполняется

$$E_0 \mid X_n(h_*) \mid = \exp\left(-nL + \frac{n}{2h_*}\right) < \infty$$

и

$$E_0\left(X_{n+1}(h) \mid F_{\tau_n}\right) = E_0\left(e^{t_{n+1}^1(h)} X_n(h) \mid F_{\tau_n}\right) = X_n(h) E_0\left(e^{t_{n+1}^1(h)}\right) < X_n(h).$$

Сформулируем теорему.

Теорема 2. Пусть выполняется условие (8) для последовательности статистик (4). Тогда для процедуры (5) имеем соотношение [5] $E_0 N_L^1(h) \geq e^c$ для любого $c > 0$.

Рассмотрим случай смещения параметра вниз после разладки ($\theta_0 > \theta$) и изучим поведение величины $E_0 N_L^2(h)$. Поскольку $t_i^2(h) \sim N\left(L, \frac{1}{h}\right)$, $i = \overline{1, \nu - 1}$ и процедура (6) является зеркальным отражением процедуры (5), то данную ситуацию можно свести к рассмотрению процедуры (5) на основе последовательности статистик $\{-t_i^2(h)\}_{i \geq 1}$, которые имеют такое же распределение, что и величины $\left\{t_i^1(h)\right\}_{i \geq 1}$. Т.к. предложенное преобразование сводит ситуацию к рассмотренной выше, то можем заключить, что $E_0 N_L^2(h) \geq e^c$ для любого c при выполнении условия (8). Следовательно, для процедуры (7) имеем соотношение $E_0 \tilde{N}_1^2(h) = E_0 \min\left\{N_L^1(h), N_L^2(h)\right\} \geq e^c$.

Заключение

В настоящей работе предложена модификация процедуры CUSUM, которая не требует знания параметра процесса AR(1) после момента отклонения. Получена нижняя граница для среднего времени между ложными тревогами.

ЛИТЕРАТУРА

1. *Konev V.V.* On one property of martigales with conditionally Gaussian increments and its application in the theory of nonasymptotic inference // Doklady mathematics. – 2016. – Vol. 471. – № 5. – P. 523–527.

2. *Tartakovsky A., Nikiforov I., Basseville M.* Sequentional analysis: hypothesis testing and changpoint detection. – Boca Raton: A Chapman & Hall Book, 2015. – 584 p.

3. *Lai T.L.* Information bounds and quick detection of parameter changes in stochastic system // IEEE transaction on information theory. – 1998. –Vol. 44. – № 7. – P. 2917 – 2929.

4. *Ширяев А.Н.* Вероятность : Учеб. пособ. для вузов. – 2-е изд., перераб. и доп. – М: Наука. Гл. ред. физ.-мат. лит., 1989. – 640 с.

5. *Konev V., Vorobeychikov S.* Quickest detection of parameter changes in stochastic regression: nonparametric CUSUM // IEEE transaction on information theory. – 2017. – Vol. 63. – № 9. – P. 5588–5602.

ПОСТРОЕНИЕ ОДНОЭТАПНОЙ ПОСЛЕДОВАТЕЛЬНОЙ ПРОЦЕДУРЫ ОЦЕНИВАНИЯ ПАРАМЕТРОВ АВТОРЕГРЕССИОННОЙ МОДЕЛИ С НЕПРЕРЫВНЫМ ВРЕМЕНЕМ

А.О. Шерстобитова, Т.В. Емельянова

Томский государственный университет

Введение

В задачах прогнозирования и управления в динамических системах огромной актуальностью пользуются модели с непрерывным временем, описываемые стохастическими дифференциальными и стохастическими разностными уравнениями. Зачастую параметры таких уравнений неизвестны, поэтому перед использованием модели для решения основных задач фильтрации, прогнозирования, управления требуется идентифицировать параметры.

В настоящее время существует достаточно много исследований, посвященных задачам асимптотического оценивания [1–4]. Однако, для практических задач типична неасимптотическая проблема оценивания, когда требуется определить длину реализации, при которой оценки достигают заданной точности [5].

Рассматривается задача оценивания параметров устойчивой авторегрессионной модели с непрерывным временем. Для решения задачи оценивания параметров модели предложена одноэтапная последовательная процедура на основе метода наименьших квадратов, использующая специальное правило остановки. Процедура позволяет контролировать среднеквадратическую точность оценивания. Получено асимптотическое распределение предложенных оценок.

1. Постановка задачи

Рассматривается случайный p-мерный процесс $X_t = \left(X_1(t),...,X_p(t)\right)'$, описываемый системой линейных дифференциальных уравнений

$$dX_t = AX_t dt + BdW_t, \qquad (1)$$

где A и B – матрицы постоянных коэффициентов размера $p \times p$, W_t – стандартный p-мерный процесс броуновского движения.

$$X_t = \begin{pmatrix} x_t^1 \\ x_t^2 \\ \dots \\ x_t^p \\ \end{pmatrix},\ A = \begin{pmatrix} 0 & 1 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 & 0 \\ . & . & . & 1 & 0 \\ . & . & . & . & . \\ \theta_p & \theta_{p-1} & \dots & \dots & \theta_1 \end{pmatrix},\ B = \begin{pmatrix} 0 & \dots & 0 \\ 0 & \dots & 0 \\ 0 & \dots & \sigma \end{pmatrix},\ \sigma > 0\,.$$

Целью работы является исследование асимптотического распределения оценок неизвестных параметров, полученных с помощью одноэтапной последовательной процедуры, предложенной в работе [6].

Рассмотрим задачу оценивания матрицы неизвестных коэффициентов $A = \|a_{ij}\|$ по наблюдениям процесса X_t на промежутке $[0;T]$.

Для решения задачи оценивания вектора параметров $\theta = (\theta_1, \theta_2, ..., \theta_p)'$ широко применяется метод наименьших квадратов (МНК). Таким образом, оценка $\hat{\theta}_T$ по наблюдениям процесса на промежутке $[0;T]$ имеет вид

$$\hat{\theta}_T = M_T^{-1} \int_0^T X_s d \langle X_t \rangle_p\,, \tag{2}$$

здесь $\langle \alpha \rangle_i$ – i-я координата вектор-столбца $\alpha = (\alpha_1, ..., \alpha_p)'$ [6],

$$M_T = \int_0^T X_s X_s' ds \tag{3}$$

– выборочная информационная матрица Фишера. M_T^{-1} – обратная к матрице M_T, если она не вырождена, и $M_T^{-1} = 0$ в противном случае [6].

Оценка (2) содержит случайную матрицу M_T, вследствие чего при исследовании свойств этой оценки возникают сложности. Поэтому в работе [6] был предложен последовательный план (τ, θ^*), гарантирующий получение оценок с заданной среднеквадратической точностью.

2. Последовательная процедура. Вид асимптотического распределения последовательной процедуры оценивания

Предположим, что процесс авторегрессии (1) устойчив, т.е. все корни характеристического полинома $Q(z) = z^p - \theta_1 z^{p-1} - ... - \theta_p$ лежат в единичном круге. Пусть $H > 0$.

Для построения последовательной процедуры будем использовать лемму [6], результатом которой является оценка квадрата нормы уклонения оценки (2) от истинного значения параметра.

Лемма [6]. Пусть матрица M_T в (3) невырождена. Тогда квадрат нормы уклонения оценки (3) удовлетворяет неравенству $\left\|\hat{\theta}_T - \theta\right\|^2 \le \left\|M_T^{-2}\right\| \cdot \|m_T\|^2$, где $m_T = \int_0^T X_s dW_s$.

Заметим, что в силу леммы 1 [6] $\left\|M_T^{-2}\right\|^{\frac{1}{2}}$ монотонно убывает, поэтому определим длительность наблюдений процесса по правилу

$$\tau = \tau(H) = \inf\left\{ t > 0 : \left\|M_T^{-2}\right\|^{\frac{1}{2}} \le \frac{1}{H} \right\} \tag{4}$$

и последовательную оценку МНК неизвестных параметров по формуле

$$\theta^*(H) = M_{\tau(H)}^{-1} \int_0^{\tau(H)} X_s d\left\langle X_t \right\rangle_p \ [6]. \qquad (5)$$

Последовательный план (4), (5) позволяет контролировать среднеквадратическую точность получаемых оценок за счет выбора порога процедуры H. Длительность процедуры при этом пропорциональна порогу процедуры.

Асимптотическое распределение последовательных оценок устанавливает следующая теорема.

Теорема. Пусть задан процесс вида (1) $dX_t = AX_t dt + BdW_t$, где A и B – квадратные матрицы размера $p \times p$, W_t – стандартный p-мерный процесс броуновского движения,

$$X_t = \begin{pmatrix} x_t^1 \\ x_t^2 \\ \dots \\ x_t^p \\ \end{pmatrix}, \ A = \begin{pmatrix} 0 & 1 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 & 0 \\ . & . & . & 1 & 0 \\ . & . & . & . & . \\ \theta_p & \theta_{p-1} & \dots & \dots & \theta_1 \end{pmatrix}, \ B = \begin{pmatrix} 0 & \dots & 0 \\ 0 & \dots & 0 \\ 0 & \dots & \sigma \end{pmatrix}, \ \sigma > 0 .$$

Предположим, что для неизвестных параметров θ_i, $i = \overline{1, p}$ выполняется условие устойчивости, т.е. корни характеристического многочлена $Q(z) = z^p - \theta_1 z^{p-1} - \dots - \theta_p$ лежат в единичном круге.

Последовательный план $\left(\tau_H, \theta^*(H)\right)$ задается формулами

$$\tau = \tau(H) = \inf\left\{ t > 0 : \left\| M_T^{-2} \right\|^{\frac{1}{2}} \le \frac{1}{H} \right\}, \ \theta^*(H) = M_{\tau(H)}^{-1} \int_0^{\tau(H)} X_s d\left\langle X_t \right\rangle_p ,$$

где $H > 0$ – пороговое значение. Тогда вектор $\frac{1}{\sqrt{H}}\left(\theta^*(H) - \theta\right)$ имеет асимптотически нормальное распределение с параметрами $\left(0, F^{-1}\right)$.

3. Результаты имитационного моделирования

Применим полученные результаты к процессу AR(1) с неизвестным параметром θ

$$dx_t = \theta x_t dt + \sigma dw_t .$$

Проведено имитационное моделирование, в ходе которого вычислены моменты остановки τ_H, а также последовательные оценки параметра θ.

Рассмотрим результаты моделирования для процесса авторегрессии с непрерывным временем. Моделирование реализовывается при условиях: $\Delta t = 0.1$, $x_0 = 0$, объем выборки $N = 1000$. Для полученных оценок строится полигон частот.

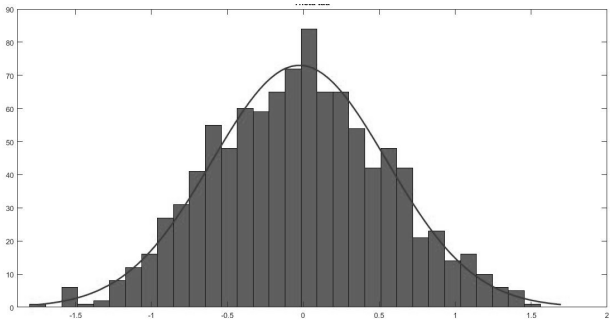


Рис. 1. Полигон частот для оценки параметра $\theta = 0.3$

С помощью критерия Харке-Бера была проведена проверка гипотезы о нормальном распределении последовательной оценки параметра θ с уровнем доверия $\gamma = 95\%$. Критерий подтвердил нормальность результатов оценивания.

Заключение

В работе представлена теорема, согласно которой последовательные оценки, построенные согласно процедуре (4), (5), имеют асимптотически нормальное распределение. Этот факт может быть использован для построения доверительных интервалов для параметров модели авторегрессии, а также для исследования оптимальности одноэтапной последовательной процедуры.

ЛИТЕРАТУРА

1. *Линьков Ю.Н.* Об оценках параметров процессов диффузионного типа. – В кн.: Теория случайных процессов. Вып. 9. Киев: Наук. думка, 1981. – С. 71–78.

2. *Кутоянц Ю.А.* Оценивание параметров случайных процессов. – Ереван, Изд-во АН АрмССР, 1980.

3. *Холево А.С.* Оценка параметра снова диффузионного процесса методом стохастической аппроксимации. – В кн.: Исследование по теории самонастраивающихся систем. М.: Наука, 1967, с. 179–200.

4. *Ибрагимов И.А., Хасьминский Р.З.* Асимптотическая теория оценивания. – М.: Наука, 1979.

5. *Конев В.В., Пергаменщиков С.М.* О последовательном оценивании параметров случайных процессов диффузионного типа // Проблемы передачи информации. – 1985. – Т. 21, Вып. 1. – С. 48–61.

6. *Емельянова Т.В., Конев В.В.* О последовательном оценивании параметров непрерывной авторегрессии // Вестник Томского гос. у-та: Математика и механика. – 2013. – № 5 (25). – С. 12–25.

III. МАТЕМАТИЧЕСКОЕ И ПРОГРАММНОЕ ОБЕСПЕЧЕНИЕ ВЫЧИСЛИТЕЛЬНЫХ МАШИН И КОМПЬЮТЕРНЫХ СЕТЕЙ

МАСКИРОВАНИЕ КОНСТАНТНЫХ НЕИСПРАВНОСТЕЙ ЛОГИЧЕСКИХ СХЕМ ВНЕШНИМИ ВЕНТИЛЬНЫМИ ПОДСХЕМАМИ

В.А. Провкин
Томский государственный университет

Введение

В процессе производства интегральных схем возможно изменение спецификации изготавливаемой схемы, возможно также обнаружение неисправностей на поздних стадиях производства схемы. В таких случаях приходится либо возвращать схемы на более ранние стадии производства, либо выбрасывать их как непригодные к использованию. Подобные ситуации снижают выход годных схем, что весьма нежелательно. Чтобы увеличить количество исправных схем, применяется ECO (Engineering Change Order) технологии. Неисправность в изготавливаемой схеме можно маскировать с помощью корректирующей схемы, подключаемой к входам корректируемой схемы и либо к выходам, либо к внутренним полюсам этой же схемы.

В [1] описан способ построения корректирующей схемы в условии доступности входов и выходов схемы. В данной работе рассматривается задача построения маскирующей схемы в условиях доступности входов корректируемой схемы и её внутреннего полюса. Задана некоторая комбинационная схема C. Известно, что один из внутренних полюсов этой схемы является неисправным. Требуется построить корректирующую (маскирующую) схему из вентилей, которая реализует частичную функцию этого полюса, и присоединить её к исходной схеме, чтобы замаскировать неисправность этого полюса. При этом желательно, чтобы корректирующая схема имела как можно меньший размер.

1. Постановка задачи

Рассмотрим комбинационную схему C (комбинационную составляющую последовательностной схемы), состоящую из вентилей. Каждому внутреннему полюсу v соответствует частично определённая булева функция. Частичная функция полюса определяется множествами единичных и нулевых наборов $M_1(f_v)$ и $M_0(f_v)$. Предполагается, что на каком-то полюсе имеет место константная неисправность. Это значит, что на всех наборах значений входных переменных этому полюсу схемы сопоставляется одно и то же значение 1(0), т.е. искажается частичная функция, сопоставляемая полюсу.

Рассмотрим частичную функцию f_1 и полностью определенную функцию f_2, представленные парами множеств единичных и нулевых наборов значений своих переменных: $M_1(f_1)$, $M_0(f_1)$; $M_1(f_2)$, $M_0(f_2)$. Будем говорить, что полностью определенная функция f_2 реализует частичную функцию f_1, если выполняются условие: пересечения множеств $M_1(f_1)$, $M_0(f_2)$ и множеств $M_0(f_1)$, $M_1(f_2)$ пусты. Это значит, что $M_1(f_2)$ содержит $M_1(f_1)$ и $M_0(f_2)$ содержит $M_0(f_1)$.

Каждому внутреннему полюсу в общем случае соответствует частичная булева функция, поскольку на некоторых наборах значений входных переменных (входных наборах) схемы соответствующее значение внутреннего полюса может не влиять на значения выходов схемы. Частичная функция, сопоставляемая внутреннему полюсу, определяется только на тех входных наборах, на которых смена соответствующего значения в рассматриваемом полюсе меняет значение на выходах схемы. Проявление не-

исправности внутреннего полюса схемы означает искажение частичной булевой функции этого полюса в области её определения. Соответственно, при построении маскирующей схемы («схемы-заплатки») необходимо использовать наборы из области определения частичной функции, чтобы найти, по возможности, более простую логическую схему, поведение которой представляется функцией f_2. В [2,3] для поиска функции-«заплатки» сводится к решению задачи QBF SAT.

2. Вероятностные оценки наблюдаемости внутреннего полюса логической схемы

Одной из характеристик внутреннего полюса является его наблюдаемость. Под наблюдаемостью внутреннего полюса схемы будем понимать возможность наблюдения смены единичного (нулевого) значения внутреннего полюса схемы на одном из ее выходов. Соответственно, чем ниже наблюдаемость полюса, тем менее определена частичная функция. Предполагается, что возникновение неисправности в схеме наиболее вероятно в полюсе с низкой наблюдаемостью.

Найдем вероятностную оценку наблюдаемости, используя ROBDD-графы. Вычисление оценки наблюдаемости полюса v комбинационной схемы C (комбинационной составляющей последовательностной схемы) относительно i-го выхода сводится к построению ROBDD-графа $R\left(C_v^i\right)$ подсхемы C_v^i, сопоставляемой этому выходу, в предположении, что полюс v является входом соответствующей одно выходной подсхемы. Граф реализует функцию $f_i\left(v,x_1,\ldots,x_n\right)$, причем, переменная v используется в качестве первой в разложении Шеннона при построении графа. Далее строится функция наблюдаемости относительно i-го выхода:

$$D_v f_i = f_i\left(0,x_1,\ldots,x_n\right) \oplus f_i\left(1,x_1,\ldots,x_n\right).$$

Эта функция зависит от входов схемы и принимает единичное значение на тех наборах, на которых изменения значений переменной v меняет i-й выход схемы.

Введём обозначения: $f_i^{v=0} = f_i\left(0,x_1,\ldots,x_n\right)$, $f_i^{v=1} = f_i\left(1,x_1,\ldots,x_n\right)$. В силу тождества

$$a \oplus b = a\bar{b} \vee \bar{a}b$$

функцию $D_v f_i$ можно записать в следующем виде:

$$D_v f_i = f_i^{v=0}\overline{f_i^{v=1}} \vee \overline{f_i^{v=0}} f_i^{v=1}$$

Построение этой функции сводится к операциям полиномиальной сложности над ROBDD-графами, представляющими функции $f_i^{v=0} = f_i\left(0,x_1,\ldots,x_n\right)$, $f_i^{v=1} = f_i\left(1,x_1,\ldots,x_n\right)$ и их инверсиями.

Для построения функции, представляемой наблюдаемость внутреннего полюса схемы относительно всех выходов, возьмём дизъюнкцию всех функций наблюдаемости относительно каждого выхода:

$$f^{obs} = \bigvee_{i=1}^{m}\left(D_v f_i\right) = \bigvee_{i=1}^{m}\left(f_i^{v=0}\overline{f_i^{v=1}} \vee \overline{f_i^{v=0}} f_i^{v=1}\right),$$

и представим ее в виде ROBDD-графа R^{obs}.

Считая единичные сигналы входных переменных равновероятными, вычисляем оценку наблюдаемости полюса v по графу R^{obs}.

Вероятность $p\left(\eta\right)$ единичного значения булевой функции η, сопоставляемой некоторой нетерминальной вершине μ ROBDD-графа, вычисляется через вероятности $p\left(\eta_\mu^{x_i=0}\right)$, $p\left(\eta_\mu^{x_i=1}\right)$ единичных значений функций $\eta_\mu^{x_i=0}$ и $\eta_\mu^{x_i=1}$, сопоставляемых дочерним вершинам нетерминальной вершины μ по следующему правилу (нетерминальная вершина μ отмечена переменной x_i)

$$p(\eta) = p(x_i)\, p\left(\eta_{\mu}^{x_i=1}\right) + p\left(\overline{x_i}\right) p\left(\eta_{\mu}^{x_i=0}\right).$$

3. Построение частичной функции и её реализации

Полюсу *v* сопоставляется полностью определенная функция $\varphi(x_1, x_2, \ldots, x_n)$, реализуемая подсхемой (она представляется ROBDD-графом $R(\phi)$), выходом которой является полюс *v*, а входами – переменные $x_1,\ldots,x_n$. Этому полюсу сопоставляется также частичная функция, реализуемая на полюсе *v* от тех же входных переменных, в условиях, когда эта подсхема является частью схемы *C*. Частичная функция определена на некоторых единичных $M_1(\phi)$ (нулевых $M_0(\phi)$) наборах полностью определенной функции $\varphi(x_1, x_2, \ldots, x_n)$. Только на этих наборах изменение значения функции $\varphi(x_1, x_2, \ldots, x_n)$ влияет на значение функции *f*, реализуемой схемой *C*.

Множества единичных и нулевых наборов частичной функции предложено представлять двумя ROBDD-графами $R_1(v)$ и $R_0(v)$. Построение частичной функции сводится к поиску всех тестовых наборов для константных неисправностей полюса *v*. Множество всех тестовых наборов для неисправности константа 0 есть множество $M_1(\phi)$, а множество всех тестовых наборов для неисправности константа 1 есть множество $M_0(\phi)$ частичной функции полюса. Если R^{obs} – граф, представляющий функцию наблюдаемости внутреннего полюса, то графы, представляющие множества области определения частичной функции, выражаются следующими равенствами:

$$R_1(v) = R(\phi) R^{obs} = R(\phi)\left[\bigvee_{i=1}^{m}\left(R\left(f_i^{v=0}\right)\overline{R\left(f_i^{v=1}\right)} \vee \overline{R\left(f_i^{v=0}\right)} R\left(f_i^{v=1}\right)\right)\right],$$

$$R_0(v) = \overline{R(\phi)} R^{obs} = \overline{R(\phi)}\left[\bigvee_{i=1}^{m}\left(R\left(f_i^{v=0}\right)\overline{R\left(f_i^{v=1}\right)} \vee \overline{R\left(f_i^{v=0}\right)} R\left(f_i^{v=1}\right)\right)\right].$$

Для извлечения всех наборов множеств $M_0(\phi)$ и $M_1(\phi)$ необходимо найти все пути, идущие из корневой вершины графа в терминальную вершину 1, и соответствующие им конъюнкции.

Построив частичную функцию выбранного полюса комбинационной схемы, необходимо синтезировать схему, реализующую эту функцию. С помощью этой схемы будем выполнять маскирование одиночной константой неисправности во внутреннем полюсе. Желательно, чтобы полученная маскирующая схема имела как можно меньший размер относительно маскируемой схемы.

Для синтеза маскирующей используем систему логического проектирования ABC. ABC – система синтеза и верификации логических схем, разработанная в Беркли, США. Интерфейс ABC представляет собой командную строку. Синтез схемы будем выполнять следующим образом:

1. С помощью команды read_pla считываем задание на синтез схемы в формате PLA – формате, описывающем область определения и таблицу истинности.
2. Командой strash выполняется непосредственный синтез схемы. В результате получается схема в виде And-Inverter Graph. Такая схема состоит из двухвходовых элементов «И» и элементов «НЕ».
3. Вывод полученной схемы в формат BLIF выполняется командой write_blif

В некоторых случаях может оказаться, что множество $M_0(\phi)$ или множество $M_1(\phi)$ является пустым. Система ABC доопределяет значения частичных функций нулями. Поэтому, если для исходной функции $M_1(\phi) = \varnothing$, $M_0(\phi) \neq \varnothing$, поступаем следующим образом: множество нулевых наборов объявляется множеством единичных

наборов. Далее синтезируем схему обычным образом. После этого к выходу полученной схемы добавляем инвертор. В результате получается схема, значение которой на выходе равно нулю на всех наборах множества $M_0(\phi)$. Такую схему будем использовать для маскирования константной неисправности в полюсе.

4. Маскирование неисправности внутреннего полюса внешней схемой

Маскирование константной неисправности на внутреннем полюсе схемы выполняется следующим образом. Пусть корректируемая схема C_i имеет n входов и m выходов. Обозначим v – внутренний полюс схемы, на котором в процессе тестирования была обнаружена константная неисправность (рис. 1).

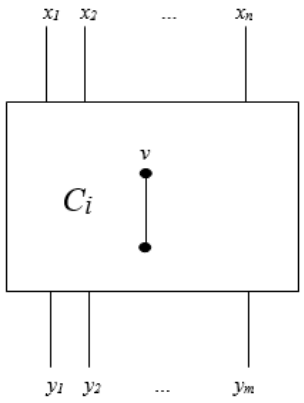


Рис. 1. Комбинационная схема с выделенным дефектным внутренним полюсом

Построив для этого полюса частичную функцию и её реализацию ранее описанным способом, мы получим некоторую вентильную схему, зависящую от входов схемы. С её помощью мы можем замаскировать неисправность полюса v при условии доступности входов схемы и внутреннего полюса. Для этого разрываются линии связи, идущие из полюса v, и выход маскирующей схемы C_p соединяется с входом того полюса, входом которого являлся полюс v. Входы маскирующей схемы подключаются к соответствующим входам исходной схемы C_i. Результат показан на рис. 2.

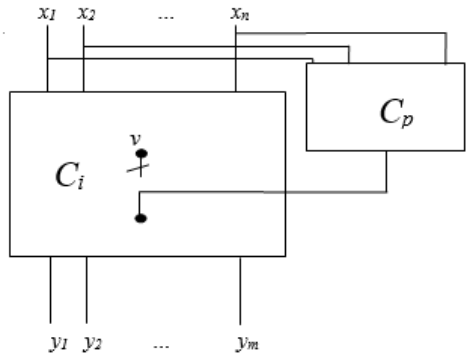


Рис. 2. Маскирование неисправности внутреннего полюса с помощью внешней схемы

Если неисправный полюс являлся точкой ветвления (рис. 3), то в этом случае разрываются все линии связи, исходящие из этого полюса, а выход маскирующей схемы разветвляется, и каждая ветвь направляется на соответствующий внутренний полюс маскируемой схемы (рис. 4).

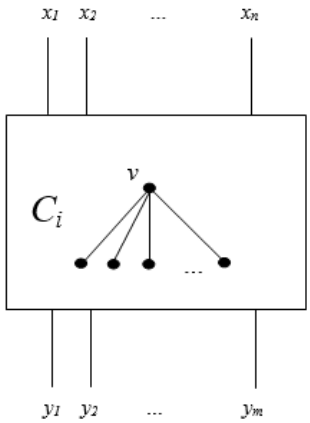


Рис. 3. Пример неисправного полюса, являющегося точкой ветвления

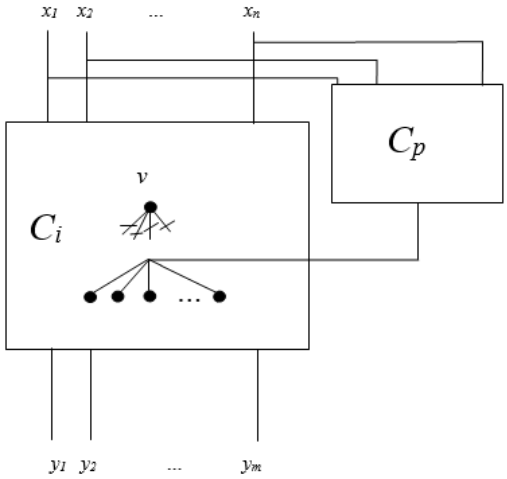


Рис. 4. Маскирование полюса, являющегося точкой ветвления

Заключение

Рассматриваются комбинационные логические схемы из вентилей. Для таких схем исследуется возможность маскирования константной неисправности на внутреннем полюсе схемы с помощью внешней вентильной подсхемы в условиях доступности входов схемы и внутреннего полюса схемы. Представлены алгоритм вычисления вероятностной оценки наблюдаемости и построения частичной функции внутреннего полюса схемы на основе операций с ROBDD-графами, а также алгоритм синтеза вентильной схемы, реализующей эту частичную функцию, ориентированный на использование системы логического проектирования ABC. Описан метод маскирования константной неисправности с помощью внешней подсхемы.

ЛИТЕРАТУРА

1. *Matrosova A., Chernyshov S., Goshin G., Kudin D.* Forming Patch Functions and Combinational Circuit Rectification // IEEE East-West Design & Test Symposium (EWDTS 2018), Kazan, 2018.

2. *Kai-Fu Tang, Po-Kai Huang, Chun-Nan Chou, and Chung-Yang (Ric) Huang* Multi-Patch Generation for Multi-Error Logic Rectification by Interpolation with Cofactor Reduction

3. *Shao-Lun Huang, Wei-Hsun Lin, Po-Kai Huang, and Chung-Yang (Ric) Huang* Match and Replace: A Functional ECO Engine for Multierror Circuit Rectification // IEEE Transactions on Computer-Aided Design of Integrated Circuits and Systems. March 2013. – V. 32. –N 3. – P. 467–478.

СИНТЕЗ ЧАСТИЧНО ПРОГРАММИРУЕМЫХ СХЕМ, ОРИЕНТИРОВАННЫЙ НА МАСКИРОВАНИЕ ТРУДНО ОБНАРУЖИМЫХ НЕИСПРАВНОСТЕЙ

В.А. Провкин
Томский государственный университет

Введение

В процессе производства интегральных схем возможно изменение спецификации изготавливаемой схемы, возможно также обнаружение неисправностей на поздних стадиях производства схемы. В таких случаях приходится либо возвращать схемы на более ранние стадии производства, либо выбрасывать их как непригодные к использованию. Подобные ситуации снижают выход годных схем, что весьма нежелательно. Кроме того, в схемы, изготовленных сторонними фирмами, могут быть внедрены вредоносные подсхемы (Trojan Circuits (TCs)) с целью разрушения системы или извлечения конфиденциальной информации. Срабатывание вредоносной подсхемы также может рассматриваться как проявление неисправности на линии связи, к которой присоединён выход вредоносной подсхемы.

Для снижения вероятности возникновения трудно обнаружимой неисправности или внедрения вредоносной подсхемы предлагается использовать частично-программируемые логические схемы [1,2]. Такие схемы представляют собой схемы, в которой подсхемы из традиционных вентилей соединены с программируемыми блоками памяти (LUT). Сначала проектируется схема из вентилей. Затем некоторые её подсхемы покрываются программируемыми блоками. Это даёт возможность либо полностью исключить возникновение неисправности (в случае, если покрываемая линия оказывается внутри программируемого блока), либо замаскировать эту неисправность в случае, если эта линия соединена со входом в блок, и имеется возможность сделать эту линию связи несущественной.

1. Постановка задачи

Рассмотрим комбинационную схему C (комбинационную составляющую последовательностной схемы), состоящую из вентилей с одним или двумя входами. Каждому внутреннему полюсу в общем случае соответствует частичная булева функция, которая определяется множествами единичных и нулевых наборов $M_1(f_v)$ и $M_0(f_v)$. Частичность функции возникает из-за того, что на некоторых наборах значений входных переменных (входных наборах) схемы соответствующее значение внутреннего полюса может не влиять на значения выходов схемы. Частичная функция, сопоставляемая внутреннему полюсу, определяется только на тех входных наборах, на которых смена соответствующего значения в рассматриваемом полюсе меняет значение на выходах схемы. Проявление неисправности внутреннего полюса схемы означает искажение частичной булевой функции этого полюса в области её определения. Предполагается, что включение вредоносной подсхемы наиболее вероятно для линий связи с низкой наблюдаемостью.

Рассмотрим частичную функцию f_1 и полностью определенную функцию f_2, представленные парами множеств единичных и нулевых наборов значений своих переменных: $M_1(f_1)$, $M_0(f_1)$; $M_1(f_2)$, $M_0(f_2)$. Будем говорить, что полностью определенная функция f_2 реализует частичную функцию f_1, если выполняются условие: пересечения множеств $M_1(f_1)$, $M_0(f_2)$ и множеств $M_0(f_1)$, $M_1(f_2)$ пусты. Это значит, что $M_1(f_2)$ содержит $M_1(f_1)$ и $M_0(f_2)$ содержит M0(f_1).

Пусть задана комбинационная схема и множество линий связи, в которые наиболее вероятно внедрение вредоносной подсхемы. Требуется покрыть схему программируемыми блоками таким образом, чтобы, по возможности, как можно больше линий из

этого множества оказалось внутри программируемого блока. При этом имеется ограничение на число входов в ПЛБ. Оставшиеся линии связи маскируются путём перепрограммирования программируемого блока, если у этого блока имеется свободный вход.

2. Вероятностные оценки наблюдаемости линий связи логической схемы

Под наблюдаемостью линии связи схемы будем понимать возможность наблюдения смены единичного (нулевого) значения линии связи схемы на одном из ее выходов. Соответственно, чем ниже наблюдаемость полюса, тем менее определена частичная функция. Предполагается, что включение вредоносной подсхемы наиболее вероятно к линии с низкой наблюдаемостью. В множество линий связи, которые требуется покрыть ПЛБ, включаются те линии, наблюдаемость которых ниже некоторого порогового значения.

Найдем вероятностную оценку наблюдаемости, используя ROBDD-графы. Вычисление оценки наблюдаемости линии *l* комбинационной схемы *C* (комбинационной составляющей последовательностной схемы) относительно *i*-го выхода сводится к построению ROBDD-графа $R\left(C_l^i\right)$ подсхемы C_v^i, сопоставляемой этому выходу, в предположении, что линия *l* является входом соответствующей одно выходной подсхемы. Граф реализует функцию $f_i\left(l,x_1,\ldots,x_n\right)$, причем, переменная *l* используется в качестве первой в разложении Шеннона при построении графа. Далее строится функция наблюдаемости относительно *i*-го выхода:

$$D_l f_i = f_i\left(0,x_1,\ldots,x_n\right) \oplus f_i\left(1,x_1,\ldots,x_n\right).$$

Эта функция зависит от входов схемы и принимает единичное значение на тех наборах, на которых изменения значений переменной *l* меняет *i*-й выход схемы.

Введём обозначения: $f_i^{l=0} = f_i\left(0,x_1,\ldots,x_n\right)$, $f_i^{l=1} = f_i\left(1,x_1,\ldots,x_n\right)$. В силу тождества $a \oplus b = a\bar{b} \vee \bar{a}b$ функцию $D_v f_i$ можно записать в следующем виде:

$$D_l f_i = f_i^{l=0}\overline{f_i^{l=1}} \vee \overline{f_i^{l=0}} f_i^{l=1}.$$

Построение этой функции сводится к операциям полиномиальной сложности над ROBDD-графами, представляющими функции $f_i^{l=0} = f_i\left(0,x_1,\ldots,x_n\right)$, $f_i^{l=1} = f_i\left(1,x_1,\ldots,x_n\right)$ и их инверсиями.

Для построения функции, представляющей наблюдаемость линии связи схемы относительно всех выходов, выполним операцию логического сложения для всех функций наблюдаемости относительно каждого выхода:

$$f^{obs} = \bigvee_{i=1}^{m}\left(D_l f_i\right) = \bigvee_{i=1}^{m}\left(f_i^{l=0}\overline{f_i^{l=1}} \vee \overline{f_i^{l=0}} f_i^{l=1}\right),$$

и представим ее в виде ROBDD графа R^{obs}.

Считая единичные сигналы входных переменных равновероятными, вычисляем оценку наблюдаемости линии *l* по графу R^{obs}.

Вероятность $p(\eta)$ единичного значения булевой функции η, сопоставляемой некоторой нетерминальной вершине μ ROBDD-графа, вычисляется через вероятности $p\left(\eta_\mu^{x_i=0}\right)$, $p\left(\eta_\mu^{x_i=1}\right)$ единичных значений функций $\eta_\mu^{x_i=0}$ и $\eta_\mu^{x_i=1}$, сопоставляемых дочерним вершинам нетерминальной вершины μ по следующему правилу (нетерминальная вершина μ отмечена переменной x_i)

$$p(\eta) = p(x_i)\,p\left(\eta_\mu^{x_i=1}\right) + p\left(\overline{x_i}\right)p\left(\eta_\mu^{x_i=0}\right).$$

3. Алгоритм покрытия заданного множества линий связи схемы программируемыми блоками

Сначала опишем вспомогательный алгоритм построения подсхемы с максимальным числом входов d, выход которой совпадает с заданной линией связи. Выполним следующие шаги алгоритма.

1. Формируем подсхему из заданной линии и двух соединенных с ней вентилей. Переходим к шагу 2.
2. Если число входов в текущую подсхему равно максимальному количеству входов d, то переходим к шагу 5. Иначе – на шаг 3.
3. Если есть линия с низкой наблюдаемостью, являющаяся входом в текущую подсхему, то определяется возможность включения её в эту подсхему в качестве внутренней линии. Пусть рассматриваемая подсхема имеет m входов, а элемент, выходом которого является текущая линия, имеет n входов. После подключения линии к подсхеме в ней будет $m+n-1$ входов. Следовательно, присоединение возможно только при условии $m+n-1 \le d$. Если данное ограничение выполнено, то эта линия вместе с вентилем, из которого она исходит, включается в эту подсхему и становится её частью. Далее переходим к шагу 2. Если же больше нет линий с низкой наблюдаемостью, которые являются входами в подсхему, то выполняем шаг 4.
4. Рассматривается линия, являющаяся выходом текущей подсхемы. Если её наблюдаемость выше порогового значения, то переходим на шаг 5. Иначе, если её наблюдаемость низка, то также пытаемся сделать эту линию частью подсхемы. Если это невозможно по причине того, что число входов в текущей подсхеме уже равно d, то переходим к шагу 5. Иначе, эта линя вместе с элементом, в который она входит, становится частью текущей подсхемы. Далее возвращаемся на шаг 2.
5. Работа алгоритма завершена, получена подсхема, которая будет покрыта программируемым блоком.

Теперь приведём алгоритм покрытия заданного множества линий связи схемы программируемыми блоками. Разделим элементы комбинационной схемы на уровни следующим образом: к 1-му уровню отнесём элементы, входы которых соединены непосредственно со входами схемы, к (i+1)-му уровню отнесём элементы, входы которых соединены с выходами элементов i-го и меньших уровней. Имеется список линий связи, значение оценки наблюдаемости которых ниже некоторой пороговой величины. Упорядочим этот список по неубыванию уровней вентилей, хотя бы один из входов которых соединен с концом линии связи из заданного списка, а при равных условиях упорядочиваем линии связи по неубыванию уровней вентилей выходы которых соединены с линиями связей из заданного списка. Пусть имеется упорядоченный список линии связи. Алгоритм покрытия:

1. Просматривается список линий связи. Ищется первая линия, которая не является входом или выходом программируемого блока. Если таких линий нет, то переходим к шагу 4. Иначе выбираем эту линию и переходим на шаг 2.
2. К выбранной линии применяется алгоритм построения подсхемы из этой линии с максимальным числом входов d. Из списка удаляются все линии, ставшие частью подсхемы. Переходим к шагу 3.
3. Рассматривается линия связи, выходящая из полученной подсхемы. Если она имеет низкую наблюдаемость, не является входом в ПЛБ, и другая линия связи, являющаяся выходом элемента, входом в который является рассматриваемая линия, также имеет наблюдаемость ниже порогового значения, то ко второй линии применяется алгоритм построения подсхемы из этой линии, но уже с максимальным числом входов $d-1$. Оставшийся вход будет использован для маскирования

неисправности первой линии в случае её обнаружения. После построения подсхемы шаг 3 повторяется, при этом рассматривается линия уже из новой подсхемы. Если линия, выходящая из подсхемы, имеет достаточно высокую наблюдаемость, то возвращаемся к шагу 1.

4. Работа алгоритма завершена, получена частично-программируемая комбинационная схема, состоящая из вентилей и программируемых блоков.

В случае, если некоторые линии заданного множества принадлежат точкам ветвления, то предлагается поступить следующим образом: элемент, выход которого соединён с несколькими линиями, дублируется. При этом рассматриваемая отделяемая линия связи соединяется с концом одного из элементов, а все оставшиеся линии, которые образовывали ветвление – с концом другого. Также дублируются линии, соединённые с входами вентиля, выход которого являлся точкой ветвления. Необходимо вычислить значения наблюдаемости для новых линий, а также заново вычислить наблюдаемости линий, которые были дублированы, т.к. значения наблюдаемости этих линий могли измениться. Во множество покрываемых линий могут добавиться новые линии связи. Этот процесс повторяется до тех пор, пока все линии с низкой наблюдаемостью не перестанут быть точками ветвления. Далее применяется описанный выше алгоритм.

4. Маскирование неисправности линии связи с использованием свободного входа

В [3] возможность маскирования неисправности линии сводится к задаче выполнимости квантифицированной конъюнктивной нормальной формы. Эта задача относится к классу PSPACE-полных задач (в отличие от задачи выполнимости пропозициональной конъюнктивной нормальной формы, которая относится к классу NP-полных задач). Отмечается, что проверка выполнимости квантифицированной конъюнктивной нормальной формы реализуется значительно медленнее, чем проверка выполнимости обычной конъюнктивной нормальной формы. В данной работе предлагается способ маскирования, не требующий решения задачи выполнимости. После покрытия схемы программируемыми блоками могут остаться линии, которые являются выходами ПЛБ, и одновременно входами в вентиль. Кроме того, могут остаться линии связи, которые не находятся внутри ПЛБ, но при этом они соединены с входом программируемого блока. В обоих случаях, если у ПЛБ имеется свободный вход, то возможно маскирование неисправности линии связи. В первом случае ПЛБ со свободным входом реализует функцию упомянутого выше вентиля. Маскирование выполняется следующим образом. Рассмотрим способ маскирования, представленный на рис. 1.

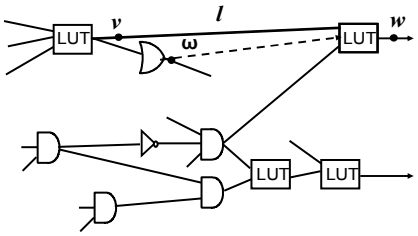


Рис. 1. Маскирование линии *l*, подключенной к входу LUT

Полюс *v* соединен линией *l* с входом u_i программируемого блока LUT, покрывающего подсхему C_{LUT} из вентилей. На рис. 1 линия *l* выделена жирным шрифтом. Входы подсхемы C_{LUT} являются либо входными, либо внутренними переменными схемы *C*. Число входных переменных LUT фиксировано, одна из переменных свободна, а функция от оставшихся переменных реализует подсхему из вентилей C_{LUT}, покрывающую соответствующий фрагмент из вентилей схемы *C*. Пусть в линию *l* включена вредоносная подсхема. Будем маскировать ее, используя свободный вход LUT. В дальнейшем этот LUT будем называть корректирующим. Обозначим его выход символом *w*. Изна-

чально идея маскирования заключалась в следующем: позиция пунктирной линии на рис. 2 заранее не определена, она может соединять свободный вход LUT с выходом другого элемента схемы, при выполнении определенных условий, а именно: функция этого элемента является реализацией частичной функции линии *l*. Однако предварительные исследования показали, что подходящая реализация частичной функции существует очень редко. Поэтому предложено использовать ту же самую реализацию путём добавления резервной линии связи (рис. 2).

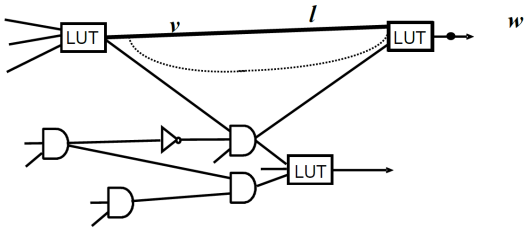


Рис. 2. Свободный вход LUT соединен с линией, дублирующей *l*

Для способа, предложенного на рис. 3, будем записывать в корректирующий программируемый блок единичные наборы полностью определенной функции, реализуемой подсхемой из вентилей, покрытой этим блоком и зависящей от его входных переменных, а именно от его $(m-1)$ переменных. Каждому единичному набору сопоставляется два набора в пространстве m переменных: один с единичным значением переменной u_m, другой – с нулевым значением этой переменной. Другими словами, если покрываемая блоком схема реализует функцию $f(u_1,\ldots,u_{m-1})$, то в программируемом блоке формируется функция, в которой переменная u_m является фиктивной (эта переменная сопоставляется пунктирной линии на рис. 1,2):

$$f_{LUT}(u_1,\ldots,u_{m-1},u_m=0)=f(u_1,\ldots,u_{m-1}),\ f_{LUT}(u_1,\ldots,u_{m-1},u_m=1)=f(u_1,\ldots,u_{m-1}).$$

В ходе работы алгоритма, описанного в разделе 3, могут также возникнуть линии связи с низкой наблюдаемостью, которые являются входами ПЛБ с более чем одним свободным входом. В этом случае возможно маскирование нескольких линии связи (а именно – стольких линий, сколько имеется свободных входов). Тогда при программировании этого блока в нём формируется функция с несколькими фиктивными переменными.

В случае, если на линии связи *l* имеет место неисправность или в неё включена вредоносная подсхема, то этот LUT необходимо перепрограммировать таким образом, чтобы переменная u_i стала фиктивной, а переменная u_m – существенной. При перепрограммировании LUT формируем функцию

$$f_{LUT}(u_i=1)\wedge u_m \vee f_{LUT}(u_i=0)\wedge \overline{u_m}$$

из единичных наборов функции корректирующего LUT. Эти наборы получены непосредственно по структуре подсхемы, покрытой корректирующим LUT, и теперь представляют функцию этого LUT в пространстве $u_1,\ldots,u_{i-1},u_{i+1},\ldots,u_m$ его входных переменных. Она отличается от исходной функции LUT тем, что в ней переменная u_i заменена переменной u_m, и переменная u_m, (как прежде переменная u_i) теперь является существенной. С целью маскирования ТС делаем переменную u_i несущественной. Это значит, что каждому единичному набору функции

$$f_{LUT}(u_i=1)\wedge u_m \vee f_{LUT}(u_i=0)\wedge \overline{u_m}$$

необходимо сопоставить два набора в пространстве m переменных: один с единичным значением переменной u_i, а другой – с нулевым значением этой переменной. Тогда ТС, подключенный к входу u_i, не может изменить корректного поведения схемы *C*.

5. Экспериментальные результаты

Алгоритм покрытия линий заданного подмножества программируемыми блоками без свободного входа реализован программно. Выполнены эксперименты на контрольных примерах схем из набора MCNC. При этом число входов в ПЛБ ограничено 8. Для этих схем выбрано пороговое значение, линии связи с оценками наблюдаемости ниже этого значения рассматривались как наиболее вероятные для включения вредоносной подсхемы. К этим схемам применён алгоритм покрытия множества линий с наблюдаемостью ниже заданного значения. На вход программы подаётся схема в формате BLIF, на выходе формируется схема, в которой линии с низкой наблюдаемостью покрыты ПЛБ, также в формате BLIF. Корректность работы программы и эквивалентность исходной и полученной схем проверена с помощью системы логического проектирования ABC применением команды cec, предназначенной для верификации комбинационных схем. В табл. 1,2 приведены данные о количестве вентилей и линий связи в исходной схеме, в схеме, полученной путём разделения точек ветвления, и в схеме, к которой применён алгоритм покрытия программируемыми блоками. В табл. 3 показано, сколько ПЛБ и с каким количеством входов использовано для покрытия линий с низкой наблюдаемостью.

Таблица 1

Сравнение количества элементов и линий связи до и после разделения ветвящихся линий

Название схемы	Пороговое значение	В исходной схеме			После расщепления		
		Вентилей	Линий связи	Линий связи с низкой наблюдаемостью	Вентилей	Линий связи	Линий связи с низкой наблюдаемостью
9symml_rs	0.01	309	503	38	363	586	73
apex6_rs	0.01	856	1573	103	934	1709	170
cordic_rs	0.01	78	135	40	96	165	54
x3	0.01	1015	1752	58	1035	1790	68
vda_rs	0.01	801	1497	41	839	1560	61
term1_rs	0.01	252	433	132	404	704	230
i4	0.01	294	542	36	294	542	36
k2_rs	0.01	1680	3078	191	1900	3446	301
too_large_rs	0.001	1220	2047	341	1462	2463	519
t481_rs	0.001	2392	4267	180	2734	4852	375
frg1_rs	0.01	174	303	100	193	339	108

Таблица 2

Сравнение количества элементов и линий связи до и после покрытия линий с наблюдаемостью ниже порога

Название схемы	Пороговое значение	После расщепления			После покрытия			
		Вентилей	Линий связи	Линий связи с низкой наблюдаемостью	Вентилей	Линий связи	Линий связи с низкой наблюдаемостью	Линий связи, для которых возможно маскирование
9symml_rs	0.01	363	586	73	259	513	0	0
apex6_rs	0.01	934	1709	170	749	1556	17	7
cordic_rs	0.01	96	165	54	41	116	5	4
x3	0.01	1035	1790	68	946	1725	3	0
vda_rs	0.01	839	1560	61	751	1499	0	0
term1_rs	0.01	404	704	230	147	594	20	8
i4	0.01	294	542	36	234	506	0	0
k2_rs	0.01	1900	3446	301	1491	3145	0	0
too_large_rs	0.001	1462	2463	519	906	1997	53	30
t481_rs	0.001	2734	4852	375	2194	4477	0	0
frg1_rs	0.01	193	339	108	75	240	9	6

Таблица 3

Сравнение количества элементов и линий связи до и после разделения ветвящихся линий

Название схемы	Входов	Выходов	Программируемых блоков							
			Всего	2-входовых	3-входовых	4-входоых	5-входовых	6-входовых	7-входовых	8-входовых
9symml_rs	9	1	31	6	15	9	1	0	0	0
apex6_rs	135	99	36	3	8	7	3	3	4	8
cordic_rs	23	2	7	0	0	2	0	0	2	3
x3	135	99	24	2	10	6	2	2	0	2
vda_rs	17	39	27	5	10	10	1	0	1	0
term1_rs	34	10	47	3	15	3	6	4	3	13
i4_rs	192	6	24	0	16	4	4	0	0	0
k2_rs	45	45	109	7	50	29	7	11	3	1
too_large_rs	38	3	90	15	13	11	5	5	18	23
t481_rs	16	1	165	2	82	65	16	0	0	0
frg1_rs	28	3	19	2	3	2	1	1	1	9

Из табл. 3 видно, что предложенный алгоритм способен значительно снизить число линий с низкой наблюдаемостью в схеме, и, таким образом, снизить вероятность внедрения вредоносной схемы или возникновения трудно обнаружимой неисправности. В некоторых случаях возможно полное покрытие линий программируемыми блоками. В остальных же случаях неисправности на некоторых линиях связи могут быть замаскированы путём перепрограммирования соответствующего блока. Также можно заметить, что число вентилей и линий связи в схеме уменьшается после покрытия, несмотря на то что после разделения всех точек ветвления сложность схемы возрастает.

Заключение

Рассматриваются комбинационные схемы из вентилей. Для таких схем предложен алгоритм покрытия линий программируемыми блоками с ограниченным числом входов с целью снижения возможностей включения в эти линии вредоносных подсхем. Разработана программная реализация описанного алгоритма. Проведены эксперименты на контрольных примерах. Показано, что во многих случаях возможно значительное снижение риска внедрения вредоносной подсхемы (ТС), а также нейтрализация её действия.

ЛИТЕРАТУРА

1. *Yamashita S., Yoshida H., Fujita M.* Increasing yield using partially-programmable circuits // Workshop on Synthesis and System Integration of Mixed Information technologies (SASIMI). – 2010. – P. 237–242.

2. *Jo S., Matsumoto T., Fujita M.* SAT-based automatic rectification and debugging of combinational circuits with LUT insertions // Proc. Of IEEE Asian Test Symposium. – 2012. P. 19–24.

3. *Matrosova A., Ostanin S.* Trojan Circuits Masking and Debugging of Combinational Circuits with LUT Insertion // 2018 IEEE International Conference on Automation, Quality and Testing, Robotics. AQTR 2018 (THETA 21), 24-26 may 2018, Cluj-Napoca, Romania. [Cluj-Napoca], 2018. – P. 462–467.

ПОСТРОЕНИЕ ЧАСТИЧНЫХ ФУНКЦИЙ ДЛЯ ПОДСХЕМ

А.С. Ложкин, А.Ю. Матросова, Э.С. Фатеева

Томский государственный университет

Введение

Производство логических схем высокого уровня интеграции и высокого быстродействия связано с большими трудозатратами. К сожалению, даже на последних этапах проектирования возможно обнаружение неисправности в схеме. Возвращение к ранним этапам проектирования является дорогостоящим, а выбрасывание неисправной схемы приводит к снижению выхода годных схем, что отражается на себестоимости производимых схем. Конечно, маскировать неисправность можно, используя дублирование в

предположении, что неисправность возможна только в одной из двух схем. Тогда достаточно использовать вместо неисправной схемы исправную схему. Возможно также применение устойчивых к неисправностям схем с использованием дублирования и детекторов кодов [1–3]. Однако этот подход связан с существенными аппаратурными затратами, которые часто оказываются неприемлемыми на практике. В [4] предложен подход к синтезу частично программируемых схем, в которых имеет место аппаратурная избыточность в пределах 10% и обеспечивается маскирование более 80% неисправностей. Подход основан на возможности представления поведения функций, сопоставляемых подсхемам, реализациями соответствующих частичных функций, конкретный вид которых определяется способом включения подсхемы в схему. Причем, для подсхемы подмножества единичных и нулевых наборов значений переменных частичной функции компактно представляются Reduced Ordered Binary Decision Diagram (ROBDD-графами) и, следовательно, обеспечивается возможность поиска лучших ее реализаций. Речь идет о частичных функциях, зависящих от входных переменных схемы, которой рассматриваемая подсхема принадлежит.

В данной работе предлагается алгоритм извлечения из частичной функции на множестве входных переменных схемы частичной функции от заданного множества внутренних переменных схемы, основанный на использовании троичного моделирования. Применяется специальное троичное моделирование, позволяющее определять допустимость интервала для функции или ее инверсии процедурой, сложность которой является линейной функцией от числа вершин ROBDD-графа, представляющего эту функцию. Воспользовавшись таким троичным моделированием, получаем интервальное описание областей, содержащих искомые множества единичных и нулевых наборов частичной функции в пространстве выделенного подмножества внутренних переменных схемы. Далее из этих областей двоичным моделированием выделяются множества единичных и нулевых наборов частичной функции. Эти множества могут затем использоваться для синтеза схемы, маскирующей неисправность выходного полюса подсхемы. Входы маскирующей подсхемы подключаются к внутренним переменным заданного подмножества и к выходу подсхемы. Такое подключение используется в рамках Engineering Change Order (ECO) технологий. Будем иметь в виду, что маскирующая подсхема корректирует не только неисправность выходного полюса рассматриваемой подсхемы, но и любую неисправность фрагмента этой подсхемы от ее выхода до ближайших точек ветвления подсхемы.

1. Постановка задачи

Пусть в некоторой схеме выделена подсхема с неисправным выходным полюсом *v* (рис. 1 а). Этой подсхеме соответствует частичная функция *f*, представленная множеством $M_1^f(X)$ единичных и множеством $M_0^f(X)$ нулевых наборов входных переменных $X = (x_1, ..., x_n)$ схемы.

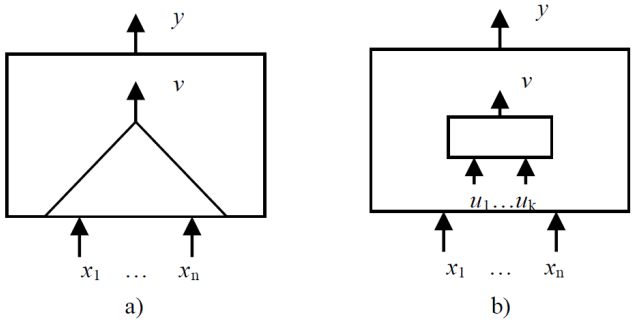


Рис. 1. Схема с неисправной подсхемой

Необходимо получить представление этой функции множествами единичных – $M_1^f(U)$ и нулевых $M_0^f(U)$ наборов внутренних переменных множества U, $U=(u_1,...,u_k)$ (рис. 1 b). Полностью определенные функции, сопоставляемые внутренним переменным, представляются ROBDD-графами R_i. Получать множества $M_1^f(U)$ и $M_0^f(U)$ предлагается, используя специальное троичное моделирование [5] с последующим применением к его результатам двоичного моделирования.

2. Построение частичных функций для подсхем

Специальное троичное моделирование логических схем [5] может быть использовано при решении различных проблем диагностирования, например для построения тестов, верификации и др.

Обозначим символом × неопределенное значение переменной. Вектор $\alpha=(\alpha_1,...,\alpha_n)$, компоненты которого принимают значения из множества $\{0,1,\times\}$, будем называть троичным вектором. Троичный вектор $\alpha=(\alpha_1,...,\alpha_n)$ задает интервал $N(\alpha)$ в булевом пространстве размерности n. В булевом пространстве размерности n выделим область единичных M_1^φ и область нулевых M_0^φ наборов полностью определенной булевой функции φ, сопоставляемой внутреннему полюсу из множества U. Если интервал $N(\alpha)\subseteq M_1^\varphi$ (M_0^φ), т.е. $N(\alpha)$ полностью содержится в области единичных (нулевых) значений функции φ, то эта функция на троичном векторе α принимает значение 1 (0). Если $N(\alpha)\cap M_1^\varphi\neq\varnothing$ и $N(\alpha)\cap M_0^\varphi\neq\varnothing$, т.е. интервал $N(\alpha)$ пересекается с обеими областями определения частичной функции, то функция φ на троичном векторе α принимает значение ×. Проиллюстрируем сказанное на диаграммах Вена (рис. 2):

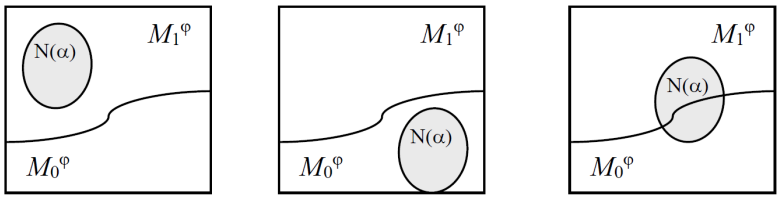


Рис. 2. Диаграммы Вена: a) $\varphi(\alpha)=1$; b) $\varphi(\alpha)=0$, c) $\varphi(\alpha)=\times$

На рис. 2 прямоугольником представлено булево пространство размерности n. Оно разделено на области M_1^φ и M_0^φ. На рис. 2a (2b) интервал $N(\alpha)$ полностью содержится в области единичных (нулевых) наборов булевой функции φ, и в этом случае точным значением функции φ, на троичном векторе α является значение 1 (0). На рисунке 2c интервал $N(\alpha)$ пересекается с обеими областями, в этом случае функция φ принимает значение × на троичном векторе α.

В процедуре вычисления троичного значения функции, зависящей от входных переменных и представленной графом ROBDD, используется вероятностный подход к специальному троичному моделированию булевых функций. Он заключается в следующем. Пусть необходимо вычислить троичное значение β булевой функции $\varphi(X)$ на троичном векторе α. Троичному вектору $\alpha=(\alpha_1,...,\alpha_n)$ сопоставляется распределение вероятностей единичных значений входных переменных $\rho(X)=(p_1,...,p_n)$. Если $\alpha_i=1$ (0), то $p_i=1$ (0). Если $\alpha_i=\times$, то $p_i=\rho_i$, где $0<\rho_i<1$, например, $1/2$. Это зна-

чит, что, если $\alpha=(0,\times,1,\times)$, то $\rho(X)=\left(0,1/2,1,1/2\right)$. Далее вычисляется вероятность $\rho(\varphi)$ единичного значения функции $\varphi(X)$ при заданном $\rho(X)$. Троичное значение β (результат специального троичного моделирования) получаем по следующему правилу: если $\rho(\varphi)=1\ (0)$, то $\beta=1\ (0)$, если $0<\rho(\varphi)<1$, то $\beta=\times$.

Для вычисления вероятности $\rho(\varphi)$ единичного значения функции $\varphi(X)$ – при заданном распределении вероятностей $\rho(X)$ – используется следующий алгоритм:

1. Пронумеруем все нетерминальные вершины ROBDD-графа, представляющего функцию $\varphi(X)$. Для этого достаточно разбить его на уровни таким образом, чтобы вершины, помеченные одинаковыми переменными, находились на одном уровне. К первому уровню отнесем вершины, помеченные переменной, которая при построении ROBDD-графа выбирается последней; ко второму уровню – предпоследней и т.д. Затем, пронумеруем произвольным образом вершины внутри уровня. В результате вершины любого уровня имеют большие номера, чем вершины предшествующих уровней.

2. В ROBDD-графе 1-концевой вершине поставим в соответствие вероятность $\rho(\varphi_w)=1$, 0-концевой вершине – вероятность $\rho(\varphi_w)=0$. Начиная с первого уровня для каждой вершины w ROBDD-графа (рис. 3), помеченной переменной x_i, вычисляем вероятность $\rho(\varphi_w)$ по формуле $\rho(\varphi_w)=\rho(x_i)\rho(\varphi_w(x_i=1))+\rho(\overline{x}_i)\rho(\varphi_w(x_i=0))$.

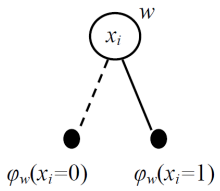


Рис. 3. Представление вершин в ROBDD-графе

Вероятности $\rho(\varphi)$ единичного значения функции φ, сопоставляемой внутреннему полюсу из множества U, соответствует вероятность, вычисленная для корневой вершины ROBDD-графа, представляющего функцию φ. Речь идет о вычислении вероятности единичного значения функции в условиях распределения вероятностей единичных значений входных переменных, соответствующих заданному троичному вектору.

Пример. Пусть имеется множество единичных M_1^f и множество нулевых M_0^f наборов значений частичной функции f, сопоставляемой полюсу v и зависящей от входных переменных $X=(x_1,x_2,x_3,x_4,x_5)$:

$$M_1^f=\begin{matrix}x_1\,x_2\,x_3\,x_4\,x_5\\1\times0\times\times\\0\,0\times0\times\\\times\,1\,1\,1\,0\\0\,0\,1\times1\end{matrix}\qquad M_0^f=\begin{matrix}x_1\,x_2\,x_3\,x_4\,x_5\\0\,1\,0\times\times\\1\,0\,1\times\times\\\times\,1\,1\,0\,0\\0\,0\,0\,1\times\end{matrix}$$

Построим соответствующее множество единичных M_1^{f*} и множество нулевых M_0^{f*} наборов значений переменных функции f^*, зависящей от внутренних переменных $U=(u_1,u_2,u_3)$.

Будем иметь в виду, что зависимость внутренних переменных множества U от входных переменных $X=(x_1,x_2,x_3,x_4,x_5)$ представляется ROBDD-графами (рис. 4):

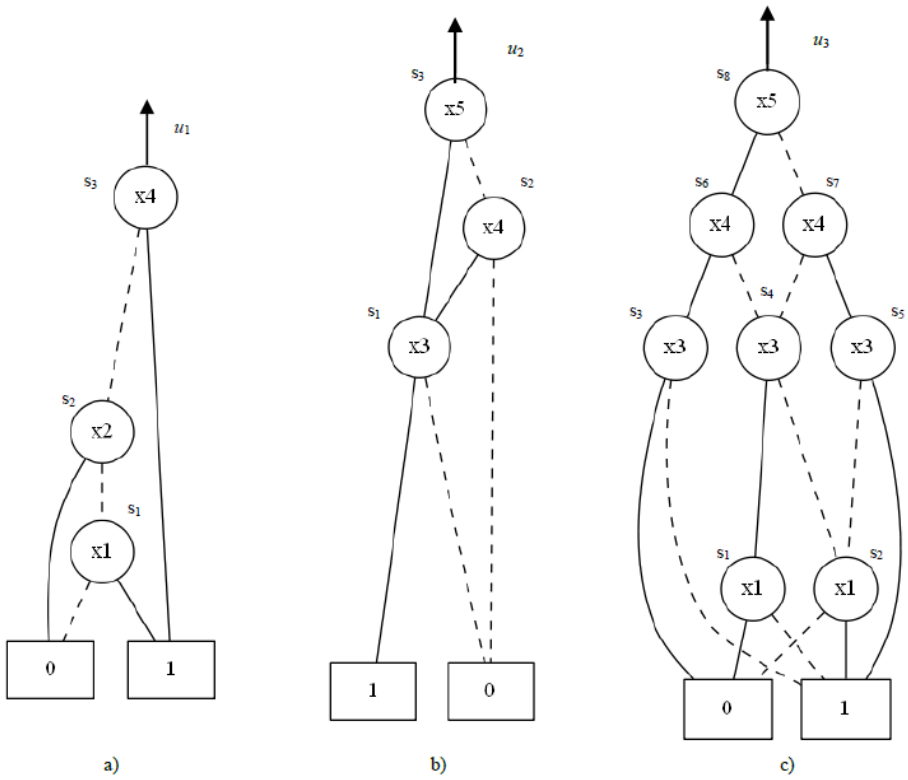


Рис. 4. ROBDD-графы, представляющие функции внутренних переменных U: a) u_1; b) u_2; c) u_3

Вычислим вероятности $\rho(u_i)$ единичного значения функций u_i, например, на троичном векторе $\alpha = (1,\times,0,\times,\times) \subseteq M_1^f$, которому соответствует распределение вероятностей $\rho(X) = \left(1, \frac{1}{2}, 0, \frac{1}{2}, \frac{1}{2}\right)$, и найдем троичный вектор $\beta = (u_1,u_2,u_3) \subseteq U_1^f$:

1. Пронумеруем вершины в графе 4a как показано на рис. 4a и посчитаем $\rho(u_1)$:

$$\rho(s_1) = \rho(x_1)\cdot 1 + \rho(\overline{x}_1)\cdot 0 = 1\cdot 1 + 0\cdot 0 = 1,$$

$$\rho(s_2) = \frac{1}{2}\cdot 0 + \frac{1}{2}\cdot \rho(s_1) = \frac{1}{2}\cdot 0 + \frac{1}{2}\cdot 1 = \frac{1}{2},$$

$$\rho(u_1) = \rho(s_3) = \frac{1}{2}\cdot 1 + \frac{1}{2}\cdot \rho(s_2) = \frac{1}{2}\cdot 1 + \frac{1}{2}\cdot \frac{1}{2} = \frac{3}{4}.$$

Т.к. $0 < \rho(u_1) < 1$, то $u_1 = \times$.

2. Аналогичным образом, для графов 4b и 4c получим $\rho(u_2) = 0$ и $\rho(u_3) = 1$. Следовательно, $u_2 = 0$ и $u_3 = 1$.

Таким образом, вектор β от переменных u_1, u_2, u_3 представляется в виде $(\times,0,1)$. Вычисляем аналогичным образом троичные векторы β на оставшихся векторах α из M_1^f, получаем соответствующее множество троичных векторов в пространстве переменных множества U:

$$U_1^f = \begin{matrix} u_1\ u_2\ u_3 \\ \times\ 0\ 1 \\ 0\ 1\ \times \\ 1\ 1\ 0 \\ \times\ 1\ 1 \end{matrix}$$

Для нахождения троичных векторов на множестве нулевых значений $\alpha \subseteq M_0^f$ функции f, зависящей от входных переменных $X = (x_1, x_2, x_3, x_4, x_5)$, необходимо вычислить вероятности $\rho(u_i)$ нулевого значения функций u_i, переименовав терминальные узлы в ROBDD-графах на рис. 4. Результатом является множество троичных векторов, сопоставляемых множеству нулевых значений частичной функции f, зависящих от внутренних переменных множества U:

$$U_0^f = \begin{matrix} u_1\ u_2\ u_3 \\ 0\ 1\ \times \\ 0\ \times\ 1 \\ 1\ 0\ \times \\ 0\ 1\ 0 \end{matrix}$$

Итак, с помощью проделанных вычислений от частичных функций полюса, зависящих от входных переменных, мы перешли к троичным векторам, зависящим от внутренних переменных схемы, порожденных этой частичной функцией (рис. 5):

$$M_1^f = \begin{matrix} x_1\ x_2\ x_3\ x_4\ x_5 \\ 1\ \times\ 0\ \times\ \times \\ 0\ 0\ \times\ 0\ \times \\ \times\ 1\ 1\ 1\ 0 \\ 0\ 0\ 1\ \times\ 1 \end{matrix} \longrightarrow U_1^f = \begin{matrix} u_1\ u_2\ u_3 \\ \times\ 0\ 1 \\ 0\ 1\ \times \\ 1\ 1\ 0 \\ \times\ 1\ 1 \end{matrix}$$

$$M_0^f = \begin{matrix} x_1\ x_2\ x_3\ x_4\ x_5 \\ 0\ 1\ 0\ \times\ \times \\ 1\ 0\ 1\ \times\ \times \\ \times\ 1\ 1\ 0\ 0 \\ 0\ 0\ 0\ 1\ \times \end{matrix} \longrightarrow U_0^f = \begin{matrix} u_1\ u_2\ u_3 \\ 0\ 1\ \times \\ 0\ \times\ 1 \\ 1\ 0\ \times \\ 0\ 1\ 0 \end{matrix}$$

Рис. 5. Множество векторов, полученное на этапе троичного моделирования

Будем иметь в виду, что троичные векторы в пространстве переменных множества U содержат множества нулевых и единичных наборов частичной функции f^* в пространстве переменных множества U. Вычислим множества единичных и нулевых наборов этой частичной функции, используя двоичное моделирование.

В качестве решения предлагается доопределить данные интервалы, с целью получения всевозможных двоичных векторов, соответствующих этим интервалам и выполнить для них двоичное моделирование. Будем иметь в виду, что если булев вектор порожден троичным вектором частичной функции f из множества единичных (нулевых) наборов, то во множества частичной функции f^* следует включать только те булевы векторы в пространстве переменных множества U, на которых полностью определенная функция f^{**} принимает единичное (нулевое) значение. Двоичное моделирование функции f^{**} выполняем по структурному описанию подсхемы с выходным полюсом v и входными полюсами множества U.

В результате двоичного моделирования получаем множества единичных $M_1^{f^*}$ и множество нулевых $M_0^{f^*}$ наборов значений переменных функции f^*.

Пусть в нашем распоряжении имеется ROBDD-граф, представляющий полностью определенную функцию *f***. Построение множества единичных и нулевых наборов функции *f** выполним с использованием этого графа.

На этапе троичного моделирования получено отображение (рис. 5). Доопределяя интервалы U_1^f и U_0^f, получим всевозможные двоичные векторы, им соответствующие. Так, например, первому интервалу из U_1^f соответствует пара векторов

$u1\ u2\ u3$
0 0 1
1 0 1

а первому интервалу из U_0^f соответствует пара векторов

$u1\ u2\ u3$
0 1 0
0 1 1

Пусть функция *f*** представляется следующим ROBDD-графом от внутренних переменных u_1, u_2, u_3 (рис. 6). Выполним двоичное моделирование, например, на первом интервале из множества U_1^f, т.е. на векторах

$u1\ u2\ u3$
0 0 1
1 0 1

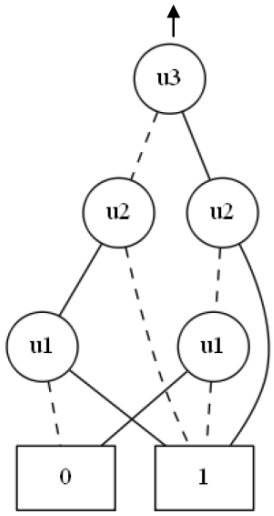


Рис. 6. ROBDD-граф

Единичному (нулевому) значению переменной u_i соответствует сплошная (пунктирная) дуга, исходящая из этой вершины. Если двоичный вектор задает путь, ведущий в 1-концевую (0-концевую) вершину, то этот вектор принадлежит множеству единичных (нулевых) наборов значений переменных значений.

Таким образом, вектор (0,0,1) принадлежит множеству M_1^{f**}, а вектор (1,0,1) – множеству M_0^{f**}. Следовательно, во множество M_1^{f*} включаем вектор (001).

Выполнив двоичное моделирование на остальных векторах, получим следующие наборы:

$$M_1 f^* = \begin{matrix} u_1 u_2 u_3 \\ 1\ 1\ 0 \\ 1\ 1\ 1 \\ 0\ 1\ 1 \\ 0\ 0\ 0 \\ 0\ 0\ 1 \end{matrix} \qquad M_0 f^* = \begin{matrix} u_1 u_2 u_3 \\ 0\ 1\ 0 \\ 1\ 0\ 1 \end{matrix}$$

Заключение

Рассматривается алгоритм построения частичных функций для подсхем, который позволяет получить множества единичных и нулевых наборов функции от внутренних переменных. По этим наборам может быть построена компактная схема, позволяющая маскировать неисправность в исходной схеме с низкой аппаратурной избыточностью.

ЛИТЕРАТУРА

1. *Matrosova A., Andreeva V., Sedov Yu.* Survivable Discrete Circuit Design // Proceedings of the 8-th IEEE International On-Line Testing Workshop. Isle of Bendor. – 2002. – P. 44–48.

2. *Efanov D.V., Sapozhnikov V.V., Sapozhnikov VI.V.* On Summation Code Properties In Functional Control Circuits // Automation and Remote Control. – 2010. – Vol. 71. — Is. 6. – P. 1117–1123.

3. *Blyudov A.A., Efanov D.V., Sapozhnikov V.V., Sapozhnikov VI.V.* Summation Codes for Organization of Control of Combinational Circuits. Automation and Remote Control // Automation and Remote Control. – 2013. – Vol. 74. – Is. 6. – P. 1020–1028.

4. *Матросова А.Ю., Кириенко И.Е., Томков В.В., Мирютов А.А.* Обеспечение надежности физических систем: синтез частично программируемых логических схем // Известия вузов. Физика. – 2014. – Т. 57. – № 6. – С. 127–132.

5. *Matrosova A., Golubeva O., Oshlakova T.* On Correction of the Results of a Ternary Simulation and a Preliminary Estimation of the Correction Results // Proceedings of the 6-th Biennial Conference on Electronics and Microsystems Technology «Baltic Electronics Conference», Tallinn. – 1998. – P. 183–186.

О СВОЙСТВАХ ROBDD-ГРАФОВ, ПРЕДСТАВЛЯЮЩИХ ТЕСТОВЫЕ ПАРЫ ДЛЯ РОБАСТНО ТЕСТИРУЕМЫХ НЕИСПРАВНОСТЕЙ ЗАДЕРЖЕК ПУТЕЙ

В.З. Тычинский, В.В. Андреева

Томский государственный университет

Введение

С увеличением уровня интеграции и быстродействия современных схем возможно возникновение различных непредусмотренных емкостей, индуктивностей, сопротивлений и т.д. Это приводит к дополнительным задержкам некоторых путей, которые невозможно рассчитать традиционными методами. Такие задержки рассматриваются как неисправности схемы. Одной из основных моделей таких неисправностей являются неисправности задержек путей (Path Delay Faults (PDFs)).

Различают робастно тестируемые и не робастно тестируемые неисправности задержек путей. Обнаружение робастно тестируемой неисправности позволяет однозначно определить путь, на котором она возникла. Обнаружение не робастно тестируемой неисправности такой возможности не дает. Если в логической схеме выявлены робастно тестируемые неисправности, то создаваемая схема отправляется разработчикам. В свою очередь они могут ее корректировать, чтобы сохранить быстродействие, определенное для схемы ранее, или понизить быстродействие с учетом выявленных задержек путей.

Обычно тестируют не все пути, а только те, задержки которых близки к максимальной. Для тестирования задержки пути необходимы два вектора – тестовая пара (v_1, v_2). Векторы пары отличаются инверсными значениями по переменной, отмечающей начало пути, а, возможно, и значениями по другим переменным. Задержки проти-

воположных перепадов значений сигналов (rising transitions и falling transitions) могут различаться, и поэтому для каждого пути в общем случае строится две пары.

В настоящее время разработаны различные методы получения тестовых пар для обнаружения неисправностей задержек путей. Среди них особый интерес представляет алгоритм, в котором поиск тестовых пар сводится к операциям над ROBDD-графами. Будем иметь в виду, что большинство операций, связанных с ROBDD-графами, обладают полиномиальной сложностью.

При тестировании неисправностей, в частности, в процессе изготовления интегральной схемы, очень важно использовать тестовые последовательности, ориентированные на потребление, по возможности, меньшей мощности. Потребляемая мощность тем больше, чем больше смен значений компонент при переходе от одного вектора последовательности к другому вектору. Поэтому важно минимизировать число таких смен значений насколько это возможно. Дело в том, что тестовые последовательности, как правило, потребляют большую мощность, чем последовательность примерно той же длины в рабочем режиме функционирования проектируемой схемы. Не учитывая этот факт, можно разрушить схему в процессе тестирования при ее изготовлении, в то время как схема могла бы корректно функционировать в рабочем режиме.

1. Постановка задачи

Основной задачей работы является анализ алгоритма построения тестовых пар для обнаружения неисправностей задержек путей с использованием бенчмарков ISCAS'89. Анализ включает в себя как общую оценку затрат ресурсов на поиск тестовых пар, так и оценку возможностей алгоритма по снижению количества перепадов в построенной тестовой последовательности.

2. Обнаружение робастно тестируемых неисправностей задержек путей

Как известно, тестирование неисправностей задержек путей на практике выполняется методами сканирования [1], при которых вектор v_2 получается из вектора v_1 либо сдвигом v_1, либо v_2 порождается реакцией комбинационной составляющей схемы с памятью на вектор v_1. Ясно, что при таком подходе тестовые пары для робастно тестируемых неисправностей далеко не всегда удается сформировать. Поэтому, наряду с традиционными методами тестирования развиваются альтернативные методы тестирования, например, Random Access Scan (RAS) методы, которые позволяют существенно сократить время тестирования, потребление мощности и множество тестовых наборов [2,3]. В рамках этих методов могут использоваться тестовые пары, гарантированно обнаруживающие неисправности задержек путей, в частности, робастно тестируемые неисправности задержек путей. Последовательности тестовых пар, ориентированные на сокращение длины теста в целом и на снижение потребляемой при тестировании мощности, могут быть получены с использованием подмножеств тестовых пар робастно тестируемых неисправностей задержек путей, построенных для каждого из выделенных путей.

В рассмотренном в данной работе алгоритме тестовые пары получают с использованием булевой разности для рассматриваемого пути [4].

3. Вычисление булевой разности для пути α

Путём в комбинационной схеме называют последовательность элементов, в которой выход предыдущего элемента является входом последующего элемента. Один из входов первого элемента последовательности является входом схемы. Выход последнего элемента последовательности является выходом схемы.

Представим путь α одновыходной комбинационной схемы C, входы которой обозначены переменными $x_1,\ldots,x_n$, последовательностью следующих символов: x_i, $u_1,u_2,\ldots,u_r$. Здесь r – длина пути α, x_i – переменная, отмечающая начало пути (вход схе-

мы C), переменные $u_1,u_2,\ldots,u_r$ отмечают выходы элементов пути. Переменная u_r отмечает выход схемы C. Заметим, что схема C может быть комбинационной частью последовательностной схемы.

Пусть переменные u_i, u_{i-1} отмечают выходы соседних элементов пути α. Рассмотрим подсхему C_{u_i} схемы C. Выход этой подсхемы отмечается переменной u_i, а входы – переменными $x_1,\ldots,x_n,u_{i-1}$. Здесь переменная u_{i-1} является входной переменной подсхемы C_{u_i} наряду с переменными $x_1,\ldots,x_n$ и одновременно входной переменной элемента с выходом, отмеченным переменной u_i.

Обозначим символом $D_{u_i} / D_{u_{i-1}}$ булеву разность, вычисляемую для функции, реализуемой подсхемой C_{u_i}, по переменной u_{i-1}. Тогда булева разность для пути α принимает следующий вид:

$$D_\alpha = \left(D_{u_r} / D_{u_{r-1}}\right) \wedge \left(D_{u_{r-1}} / D_{u_{r-2}}\right) \wedge \ldots \wedge \left(D_{u_2} / D_{u_1}\right) \wedge \left(D_{u_1} / D_{x_i}\right).$$

4. Получение тестовых наборов для робастно тестируемых неисправностей с использованием ROBDD-графов

Для вычисления булевой разности $D_{u_i} / D_{u_{i-1}}$ необходимо выполнить следующие шаги:

1. Получить ROBDD $R\left(f_{u_i}\right)$ – граф, представляющий функцию подсхемы C_{u_i}. ROBDD-граф $R\left(f_{u_i}\right)$ зависит от переменных $x_1,\ldots,x_n,u_{i-1}$. Этот граф получается при условии, что переменная u_{i-1} используется в качестве первой переменной при выполнении декомпозиции Шеннона.

2. Получить из $R\left(f_{u_i}\right)$ два ROBDD-графа: $R\left(f_{u_i}^{u_{i-1}=0}\right)$, $R\left(f_{u_i}^{u_{i-1}=1}\right)$. Их корни являются дочерними вершинами корня ROBDD-графа $R\left(f_{u_i}\right)$. ROBDD-графы $R\left(f_{u_i}^{u_{i-1}=0}\right)$, $R\left(f_{u_i}^{u_{i-1}=1}\right)$ реализуют функции, которые получаются из функции f_{u_i} заменой переменной u_{i-1} константами 0 и 1 соответственно.

3. Выполнить перемножения функций $f_{u_i}^{u_{i-1}=0}$, $\overline{f}_{u_i}^{u_{i-1}=1}$ и функций $\overline{f}_{u_i}^{u_{i-1}=0}$, $f_{u_i}^{u_{i-1}=1}$, используя соответствующие ROBDD-графы. Объединив полученные при перемножении функции, получим ROBDD $R\left(D_{u_i} / D_{u_{i-1}}\right)$:

$$R\left(D_{u_i} / D_{u_{i-1}}\right) = \left(R\left(f_{u_i}^{u_{i-1}=0}\right) \wedge R\left(\overline{f}_{u_i}^{u_{i-1}=1}\right)\right) \vee \left(R\left(\overline{f}_{u_i}^{u_{i-1}=0}\right) \wedge R\left(f_{u_i}^{u_{i-1}=1}\right)\right).$$

Отметим, что ROBDD-графы $R\left(\overline{f}_{u_i}^{u_{i-1}=0}\right)$, $R\left(\overline{f}_{u_i}^{u_{i-1}=1}\right)$ получаются из ROBDD-графов $R\left(f_{u_i}^{u_{i-1}=0}\right)$, $R\left(f_{u_i}^{u_{i-1}=1}\right)$ в результате перестановки их терминальных вершин.

Выполнив перемножение графов, получаем граф $R\left(D_\alpha\right)$, представляющий булеву разность для пути α:

$$R\left(D_\alpha\right) = R\left(D_{u_r} / D_{u_{r-1}}\right) \wedge R\left(D_{u_{r-1}} / D_{u_{r-2}}\right) \wedge \ldots \wedge R\left(D_{u_2} / D_{u_1}\right) \wedge R\left(D_{u_1} / D_{x_i}\right).$$

Следует иметь в виду, что всякая простая цепь, соединяющая корень ROBDD-графа $R\left(D_\alpha\right)$ с его единичной терминальной вершиной, представляет конъюнкцию в

пространстве переменных $x_1,\ldots,x_n$ (в конъюнкции могут присутствовать не все переменные). Булев вектор μ на множестве переменных $x_1,\ldots,x_n$, обращающий эту конъюнкцию в единицу, есть вектор v_2 тестовой пары (v_1, v_2) либо для rising transition (если вектор μ обращает схему C в 1), либо для falling transition (если вектор μ обращает схему C в 0) пути α.

Следует отметить, что векторы пары (v_1, v_2) представляют полные состояния последовательностной схемы. Входные составляющие полных состояний этих векторов являются предпоследним и последним входными векторами последовательности, доставляющей эту пару из начального состояния q_0.

Итак, ROBDD R_α представляет все тестовые наборы v_2 для rising и falling transitions, которые не отделены друг от друга. Необходимо их разделить. ROBDD для rising transition обозначим в виде $R_{rise} = R(D_\alpha) \wedge x_i$. ROBDD для falling transition обозначим через $R_{fall} = R(D_\alpha) \wedge \overline{x}_i$.

Поскольку литера ЭНФ, сопоставляемая рассматриваемому пути, может иметь различные знаки инверсии, знак инверсии над переменной x_i также может изменяться.

Обозначим символом R'_{rise} ROBDD-граф, полученный из графа R_{rise} удалением переменной x_i, а символом R'_{fall} – ROBDD-граф, полученный из графа R_{fall} удалением переменной x_i. В обоих случаях наличие инверсии над переменной x_i не имеет значения.

Обозначим символом R_{rob} ROBDD-граф, представляющий тестовые пары соседних по переменной x_i наборов для робастно тестируемых неисправностей задержек пути α: $R_{rob} = R'_{rise} \wedge R'_{fall}$.

Заметим, что граф R_{rob} не содержит переменной xi. Путь от корня графа R_{rob} до его единичной концевой вершины представляет конъюнкцию, такую, что булев вектор в пространстве переменных $x_1,\ldots,x_{i-1},x_{i+1},\ldots,x_n$ задает тестовую пару в пространстве n переменных. Один из векторов пары получается приписыванием переменной x_i значения 0, а другой – значения 1. Из тестовой пары, как указано выше, формируются тройки векторов, обнаруживающие задержки инверсных перепадов сигналов рассматриваемого пути.

Проиллюстрируем работу алгоритма на примере. Рассмотрим комбинационную схему C (рис. 1):

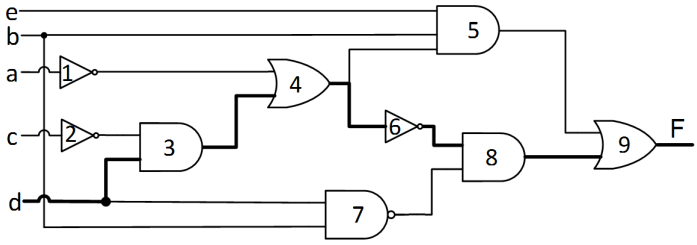


Рис. 1. Комбинационная схема C

Пусть путь α включает элементы с номерами 3, 4, 6, 8, 9. Начало пути отмечается переменной d. Таким образом, $\alpha = d,u_3,u_4,u_6,u_8,u_9$. Для наглядности и компактности соответствующие ROBDD-графы будем представлять в виде ДНФ:

$$D_{u_9} / D_{u_8} = (u_5 \vee (u_8 = 0)) \oplus (u_5 \vee (u_8 = 1)) = \overline{u}_5 = \overline{b} \vee \overline{e} \vee ac \vee a\overline{d}$$

$$D_{u_8} / D_{u_6} = ((u_6 = 0) \wedge u_7) \oplus ((u_6 = 1) \wedge u_7) = u_7 = \overline{b} \vee \overline{d}$$

$$D_{u_6} / D_{u_4} = (u_4 = 0) \oplus (u_4 = 1) = 1$$

$$D_{u_4} / D_{u_3} = (u_1 \vee (u_3 = 0)) \oplus (u_1 \vee (u_3 = 1)) = \overline{u}_1 = a$$

$$D_{u_3} / d = (u_2 \wedge (d = 0)) \oplus (u_2 \wedge (d = 1)) = u_2 = \overline{c}$$

$$D_{\alpha} = (\overline{b} \vee \overline{e} \vee ac \vee a\overline{d}) \wedge (\overline{b} \vee \overline{d}) \wedge 1 \wedge a \wedge \overline{c} = a\overline{c}\overline{b} \vee a\overline{c}\overline{d}$$

$$D_{rise} = (a\overline{c}\overline{b} \vee a\overline{c}\overline{d}) \wedge d = a\overline{c}\overline{b};\ D'_{rise} = a\overline{c}\overline{b}$$

$$D_{fall} = (a\overline{c}\overline{b} \vee a\overline{c}\overline{d}) \wedge \overline{d} = a\overline{c}\overline{d};\ D'_{fall} = a\overline{c}$$

$$D_{rob} = a\overline{c}\overline{b} \wedge a\overline{c} = a\overline{c}\overline{b}$$

Заметим, что ДНФ D_{rob} не содержит переменной xi. Конъюнкции ДНФ D_{rob} таковы, что соответствующие им булевы векторы в пространстве переменных $x_1,\ldots,x_{i-1},x_{i+1},\ldots,x_n$ задают тестовые пары в пространстве *n* переменных. Один из векторов пары получается приписыванием переменной x_i значения 0, а другой – значения 1. Из тестовой пары формируются тройки векторов, обнаруживающие задержки инверсных перепадов сигналов рассматриваемого пути.

5. Используемые технологии

5.1. Язык C++ и фреймворк Qt

Язык программирования C++ – компилируемый, статически типизированный язык программирования общего назначения. Он поддерживает такие парадигмы программирования, как процедурное программирование, объектно-ориентированное программирование, обобщённое программирование. Язык имеет богатую стандартную библиотеку, которая включает в себя распространённые контейнеры и алгоритмы, ввод-вывод, регулярные выражения, поддержку многопоточности и другие возможности. C++ сочетает свойства как высокоуровневых, так и низкоуровневых языков.

Использование языка C++, ввиду наличия объектно-ориентированной парадигмы, упрощает и ускоряет разработку программного обеспечения в сравнении, например, с Си. Также стоит отметить, что для данной задачи достаточно остро стоит вопрос производительности, поскольку рассмотренные в работе алгоритмы требуют большого количества вычислений. С этой точки зрения использование языка C++ предпочтительнее, чем, например, C#.

Qt – кроссплатформенная программная платформа для разработки программного обеспечения на языке C++. Данная платформа позволяет запускать написанное с её помощью программное обеспечение в большинстве современных операционных систем путём обычной компиляции программы для каждой операционной системы без изменения файлов исходного кода. Отличительной особенностью Qt является использование метаобъектного компилятора – системы предварительной обработки исходного кода. Метаобъектная система обеспечивает в C++ поддержку «сигналов» и «слотов» – подход, позволяющий реализовать шаблон «Наблюдатель» («Observer»), минимизируя написание повторяющегося кода. Использование сигналов и слотов упрощает коммуникацию между объектами в режиме реального времени, что особенно актуально для интерактивных приложений с графическим интерфейсом. Кроме того, объекты, соединённые посредством этой системы, не обязаны знать о структуре объектов, с которыми они соединены.

Таким образом, использование языка C++ вместе с платформой Qt для разработки даёт возможность, с одной стороны, получить более выраженную модульную структуру, а с другой – достаточное быстродействие, позволяющее производить большие объёмы вычислений.

5.2. Пакет CUDD

Пакет CUDD (Colorado University Decision Diagram) [5], разработанный в университете Колорадо, предоставляет набор функций для манипулирования двоичными диаграммами решений (Binary Decision Diagram (BDD)) [6], алгебраическими диаграммами решений (Algebraic Decision Diagram (ADD)) [7] и двоичными диаграммами решений с сокращенным нулем (Zero-suppressed Binary Decision Diagram (ZDD)) [8].

BDD или Reduced Ordered BDD (ROBDD)-графы используются для канонического представления булевых функций, что и определило их широкое применение при решении практических задач. ADD используются для представления т.н. «псевдобулевых» функций, реализующих отображение из множества $\{0,1\}^n$ в произвольное множество целых чисел. ZDD, подобно BDD, представляют булевы функции, но имеют более компактное представление в случае «разряженных» булевых функций (функций, у которых мощность множества единичных значений существенно меньше 2^n).

Пакет предоставляет большой набор операций для BDD, ADD и ZDD, функции преобразования их друг в друга, а также большое множество методов сортировки переменных. Пакет CUDD можно использовать тремя способами:

1. Как черный ящик. В этом случае прикладная программа, работающая с диаграммами решений, использует только публичные функции пакета. Богатый набор функций, включенных в пакет CUDD, позволяет писать с его помощью самые разные приложения. В приложениях, написанных с использованием функций пакета, нет необходимости заботиться о деталях переупорядочения переменных, которые происходят внутри этих функций.

2. Как белый ящик. При написании сложного приложения, использующего диаграммы принятия решений, для увеличения скорости вычислений часто требуется выполнение некоторых функций, которые напрямую манипулируют диаграммами, вместо использования заготовленных функций.

3. Через интерфейс. Объектно-ориентированные языки, такие как C++ и Perl5, могут освободить программиста от сложностей управления памятью. Интерфейс C++ включен в дистрибутив CUDD. Он позволяет автоматически освобождать память, которая больше не используется приложением, а также перегружает операторы.

6. Результаты бенчмарков

Для анализа алгоритма использовались бенчмарки ISCAS’89. Для каждого выхода каждой схемы было выбрано не менее 10 самых длинных (табл. 1) и коротких путей (табл. 2).

Таблица 1

Данные для самых длинных путей

Схема	Число входов	Число вентилей	Выбрано путей	Робастно тестируемые пути, %	Размер графов			Ранги тестов		
					min	avg	max	min	avg	max
s298	17	119	146	65.07	3	7.26	17	1	4.52	7
s344	24	160	159	69.81	4	10.5	45	1	9.01	12
s400	24	162	258	82.56	4	11.6	24	1	8.84	14
s444	24	181	237	59.92	4	9.04	15	2	7.84	13
s641	54	379	309	44.34	4	10.1	19	2	6.87	13
s820	23	289	232	99.14	6	8.96	28	2	7.01	13
s953	45	395	338	92.60	3	12.4	40	1	10.8	16
s1196	32	529	334	48.50	3	11.2	72	1	10.2	16
s1488	14	653	312	93.27	6	9.84	27	1	7.51	13
s1494	14	647	336	91.07	5	9.94	27	3	7.74	13

Таблица 2

Данные для самых коротких путей

Схема	Число входов	Число вентилей	Выбрано путей	Робастно тестируемые пути, %	Размер графов			Ранги тестов		
					min	avg	max	min	avg	max
s298	17	119	149	87.25	3	7.02	15	1	4.36	6
s344	24	160	161	91.30	4	13.3	77	1	8.92	12
s400	24	162	239	72.38	4	12.3	41	1	8.21	14
s444	24	181	234	67.09	4	11.4	30	2	8.74	13
s641	54	379	330	92.73	4	26.6	155	2	14.2	24
s820	23	289	229	100.00	6	26.4	928	2	16.3	20
s953	45	395	365	93.15	3	12.6	49	1	10.5	17
s1196	32	529	328	92.68	3	31.2	299	1	13.2	18
s1488	14	653	306	99.35	6	16.9	189	1	11	13
s1494	14	647	305	97.05	5	17.2	196	3	10.7	13

Из полученных данных можно сделать следующие выводы:

– Поскольку предложенный метод является точным, процент обнаруженных робастно тестируемых путей намного выше в сравнении с используемыми в настоящее время эвристическими методами сканирования.

– Среднее число вершин в графах не превышает нескольких десятков. Как следствие, затраты вычислительных ресурсов на их получение невелики.

– Небольшие ранги получившихся тестов дают больше возможностей для их доопределения с целью снижения количества перепадов при построении тестовой последовательности.

Заключение

Анализ результатов показывает, что реализация предложенного алгоритма с использованием пакета CUDD имеет значительное преимущество в количестве обнаруженных неисправностей по сравнению с используемыми в настоящее время эвристическими методами. Небольшие размеры получаемых графов дают возможность применения алгоритма для тестирования больших схем за приемлемое время. Также у данного алгоритма имеется большой потенциал для решения проблемы генерации соседних пар тестовых наборов, что в свою очередь позволит снизить разогрев схем при их тестировании. Целью дальнейшего исследования является разработка алгоритма, позволяющего строить тестовые последовательности, сокращая число перепадов между её соседними векторами. Объединение этих алгоритмов может найти широкое применение при тестировании неисправностей задержек путей.

ЛИТЕРАТУРА

1. *Agrawal V.D.,Cheng R.T.,Johnson D.D., Lin T.S.* Designing Circuits with Partial Scan // IEEE Design & Test of Computers. – 1988. – V. 5. – N 2. – P. 8–15.

2. *Hu Fu Y., Fan X., Fujiwara H.* Localized Random Access Scan: Towards Low Area and Routing Overhead // Proceedings of the 2008 Asia and South Pacific Design Automation Conference. – IEEE Computer Society Press, 2008. – P. 565–570.

3. *Adiga R.,Arpit G., Singh V., Satuja K.K., Fujivara H., Singh A.D.* On Minimization of Test Application Time for RAS // 2010 23 International Conference on VLSI design. – IEEE, 2010. – P. 393–398.

4. *Matrosova A.Yu., Andreeva V.V., Nikolaeva E.A.* Finding Test Pairs for PDFs in Logic Circuits Based on Using Operations on ROBDDs //Russian Physics Journal. – 2018. – V. 61. – № 5. – P. 994–999.

5. *Somenzi F.* CUDD: CU Decision Diagram Package. Release 3.0.0 [Электронный ресурс] : University of Colorado at Boulder, 2015. – URL: http://davidkebo.com/cudd (дата обращения: 29.12.2018).

6. *Bryant R.E.* Graph-Based Algorithms for Boolean Function Manipulation // IEEE Transactions on Computers. – 1986. – V. 35. – № 8. – P. 677–691.

7. *Bahar R.I.* Algebraic decision diagrams and their applications // Proceedings of the 1993 IEEE/ACM International Conference of CAD. – IEEE Computer Society Press. – 1993. – P. 188–191.

8. *Minato S.* Zero-suppressed BDDs for set manipulation in combinatorial problems // Proc. 30th ACM/IEEE Design Automation Conference. – 1993. – P. 272–277.

ПРИМЕНЕНИЕ РАСШИРЕННОГО АВТОМАТА ДЛЯ ОБУЧЕНИЯ АЛГОРИТМАМ С ВАРИАТИВНОЙ ТРАЕКТОРИЕЙ РЕШЕНИЯ

А.В. Васинев, Е.Н. Беккерман
Томский государственный университет

Введение

Автоматизированные системы обучения являются одним из действенных средств повышения эффективности учебного процесса и позволяют обеспечить адаптацию процесса обучения к индивидуальным характеристикам обучаемых по уровню знаний и разгрузку преподавателя от ряда трудоемких повторяющихся операций, связанных с предъявлением учебной информации и контролем знаний. Использование компьютерных средств обучения изменяет традиционные способы преподнесения информации. Основной идеей большинства автоматических обучающих систем является автоматизация тех или иных аспектов образовательного процесса.

Контролировать освоение обучаемым курса можно различными способами, такими как простое тестирование [1] или тестирование с градацией ценности ответа по сложности [2–5]. Помимо контроля процесса освоения материала, возможно также использовать автоматизированные системы обучения для преподавания материала курса студенту, выявления проблемных разделов курса и построения индивидуального плана обучения [6–9].

При построении подобных систем выгодно отличается автоматная модель. Она расширяет возможности автоматизированных систем обучения в вопросе выделения сильных и слабых сторон подготовки студента и создания индивидуальной траектории обучения [10,11].

Все вышеперечисленные примеры сосредотачивались на построении цельной системы обучения, предоставляющей студенту теоретический материал и проводящей тестирование или контрольную работу по нему.

В данной работе интерес авторов связан с применением автоматной модели для обучения решению конкретной задачи, а именно – задачи приведения матрицы к ступенчатому виду. Эта задача является одной из стандартных задач, возникающих при изучении дисциплины «Линейная алгебра» в разделе «Решение систем линейных уравнений».

Решение задачи приведения матрицы к ступенчатому виду легко алгоритмизируется для решения её при помощи компьютера. Однако при решении этой задачи человеком возникают некоторые особенности, которые требуется учитывать. Например, учебные задачи содержат, как правило, только целые числа. Используя элементарные гауссовы преобразования, можно решить задачу приведения матрицы к ступенчатому виду, не выходя за рамки множества целых чисел, тогда как машинный алгоритм в общем случае можно сформулировать как для использования рациональных чисел, так и только целых чисел. Мы решили ограничить возможности компьютера только целыми числами. В этом случае ход решения задачи приведения матрицы к ступенчатому виду может различаться у разных людей, и полученные в результате ступенчатые матрицы могут численно различаться. Таким образом, задача имеет вариативную траекторию решения и ответ нельзя сравнить с единственно верным.

В процессе решения студентами подобных задач могут возникать два типа ошибок: арифметические и алгоритмические. Арифметические ошибки связаны с неправильными вычислениями, в то время как алгоритмические – с неправильным применением алгоритма. Арифметические ошибки связаны с невнимательностью, в то время как алгоритмически ошибки могут значить, что студент не твёрдо знает изученный материал. В процессе изучения алгоритма студент должен сосредоточиться в первую очередь на изучении алгоритма, а не на арифметических вычислениях, т.е. логично перепоручить

компьютеру вычисления, сосредоточившись на изучении самого алгоритма. Проверка не только ответа студента, но и хода решения может помочь как преподавателю, так и студенту выявить пробелы в знаниях студента и помочь устранить их [12,13]. Кроме того, для обучения решению задач с вариативной траекторией решения необходимо создание обучающего диалога, который не будет давать конкретные подсказки, а будет лишь указывать цель, которой необходимо достичь студенту на данном этапе решения и комментировать прошлые действия. Использование автоматной модели позволяет облегчить реализацию обучающего диалога.

Все эти принципы легли в основу нашей работы. Задача данной работы состоит в том, чтобы построить управляющий расширенный автомат обучающего диалога, который позволит направлять ход решения студента в соответствии с корректным применением гауссовых преобразований, не ограничивая его действия машинным алгоритмом.

1. Организация обучающего диалога с использованием автомата

Задача студента состоит в последовательном приведении матрицы к ступенчатому виду, не разрушая уже построенного фрагмента решения. После каждого действия студента осуществляется анализ текущего состояния матрицы, для чего достаточно следить за изменениями двух переменных: k – номер строящейся ступени матриц и w – количество ненулевых элементов в столбце под первым ненулевым элементом лидер-строки. Изменение этих переменных говорит об эффективности и правильности действий студента. Увеличение переменной w или уменьшение переменной k говорит о фатальной ошибке в применении алгоритма, о чём должно быть сообщено студенту. Правильные действия также должны быть прокомментированы. Т.е. в зависимости от изменения переменных k и w из одного и того же состояния, по одному и тому же входному символу может совершиться переход в разные состояния и с разными выходными символами.

Таким образом, для реализации поставленных задач целесообразно использовать расширенный автомат. Помимо состояний, переходов, входных и выходных символов имеющихся в конечном автомате, расширенный автомат обладает также контекстными переменными и условиями перехода (предикатами).

Входной алфавит управляющего автомата должен состоять из символов, соответствующим элементарным гауссовым преобразованиям. Заметим, что входные символы в этом случае требуют дополнительных параметров:

- перестановка строк, параметрами являются номера переставляемых строк i, j;
- деление нацело строки на целое число, параметрами являются номер строки i и общий делитель элементов строки $p \neq 0$;
- линейная комбинация двух строк, параметрами являются номера строк i, j и коэффициенты линейной комбинации $p, g \neq 0$.

Выходной алфавит автомата должен соответствовать реакциям на действия, производимые студентом, или различным этапам решения, а именно:

- критическая ошибка, разрушившая частичное решение;
- нерезультативное действие, когда в результате действия студента не произошло никаких изменений, видимых из поведения переменных k и w, характеризующих состояние матрицы;
- лидер–строка была поставлена на первое (правильное) место среди неступенчатых строк и студенту необходимо уменьшить количество ненулевых элементов под первым ненулевым элементом лидер-строки;
- уменьшение количества ненулевых элементов под первым ненулевым элементом лидер-строки на единицу;
- построение очередной ступени завершено;
- матрица приведена к ступенчатому виду.

Состояния автомата также должны соответствовать различным этапам решения. На начальном этапе, или в нулевом состоянии, необходимо произвести анализ сгенерированной матрицы и выбрать следующее состояние. Если лидер-строка не находится на правильном месте, то автомат должен переходить в первое состояние. В этом состоянии необходимо поставить лидер-строку на правильное место. Иначе автомат должен переходить во второе состояние. В этом состоянии необходимо уменьшить количество ненулевых элементов под первым ненулевым элементом лидер-строки до нуля. Первое и второе состояния являются «рабочими», весь процесс приведения матрицы к ступенчатому виду осуществляется переходами между ними. Если же матрица сразу была сгенерирована ступенчатой или студент привёл её к ступенчатому виду, то автомат переходит в третье состояние. В этом состоянии студент может совершать любые действия, не разрушающие ступенчатую матрицу. Оно предназначено для того, чтобы студент мог каким либо образом улучшить готовую матрицу, например, избавиться от отрицательных элементов в строке. Если студент подтверждает готовность ступенчатой матрицы, то автомат переходит в четвёртое состояние. Это состояние является конечным, в нём студент не может совершать никаких действий. Если переход в это состояние был совершён, когда матрица ещё не была приведена к ступенчатому виду, то студенту выставляется неудовлетворительная оценка.

Большинство сообщений обучающего диалога должны состоять из двух частей. Первая часть характеризует предыдущее действие, например, если студент уменьшил количество ненулевых элементов под первым ненулевым элементом на два или построил очередную ступень. Вторая часть должна направлять дальнейшее решение, например, если студенту необходимо поставить лидер-строку на первое место среди неступенчатых строк или уменьшить количество ненулевых элементов под первым ненулевым элементом лидер-строки. Исключениями из этого правила должны быть только сообщения о критических ошибках и нерезультативных действиях. Нерезультативные действия не характеризуются никак, а сообщения о критических ошибках содержит только характеристику предыдущего действия.

2. Управляющий автомат обучающего модуля

Предлагаемый нами автомат представлен на рис. 1.

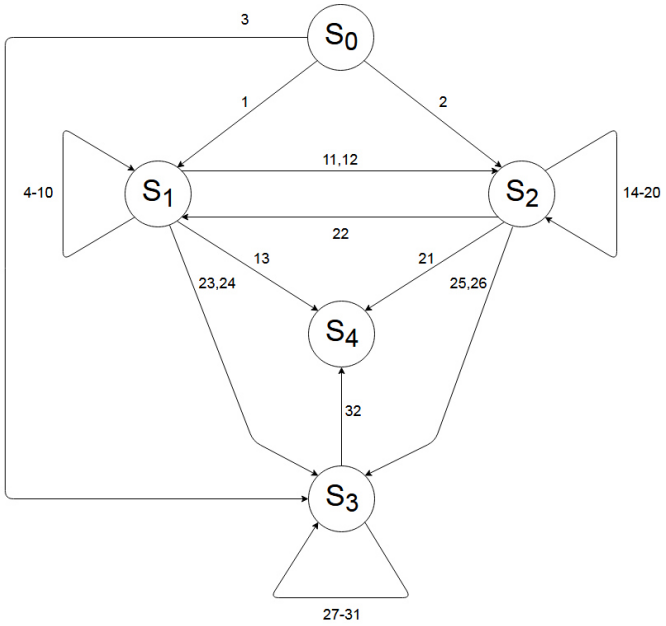


Рис. 1. Управляющий автомат обучающего диалога

Исходя из предшествующих предпосылок, представленный на рис. 1 автомат является расширенным и обладает следующими параметрами:

Множество контекстных переменных $V = \{k, w\}$:

- k – номер строящейся ступени матрицы (строки, в которой должна располагаться лидер-строка);
- w – число ненулевых элементов под первым ненулевым элементом лидер-строки, если она расположена на первом месте среди неступенчатых строк, иначе w не вычисляется и всегда равна нулю.

Символы входного алфавита автомата $X = \{x_0, x_1, x_2, x_3, x_4\}$:

- x_0 – пустой входной символ;
- x_1 – перестановка строк, входными параметрами являются номера строк i, j;
- x_2 – деление нацело строки на целое число, входными параметрами являются номер строки i и число $p \neq 0$;
- x_3 – линейная комбинация строк, входными параметрами являются номера строк i, j и числа $p, g \neq 0$, на которые они умножаются;
- x_4 – символ, обозначающий завершение работы.

Кроме того, входными параметрами всех символов, кроме x_0, являются номер текущей лидер-строки k и количество ненулевых элементом под первым ненулевым элементом лидер-строки w.

Выходной алфавит автомата $Y = \{y_0, y_1, y_2, y_3, y_4, y_5, y_6\}$ соответствует реакциям на действия, производимые студентом:

- y_0 – пустой выходной символ;
- y_1 – неправильное действие;
- y_2 – лидер–строка поставлена на правильное место;
- y_3 – появление нового нулевого элемента под первым ненулевым элементом лидер-строки;
- y_4 – построение ступени закончено;
- y_5 – построение ступенчатой матрицы закончено;
- y_6 – нерезультативное действие.

Кроме того, выходными параметрами всех символов, являются номер текущей лидер-строки k и количество ненулевых элементом под первым ненулевым элементом лидер-строки w.

Множество состояний $S = \{S_0, S_1, S_2, S_3, S_4\}$:

- S_0 – соответствует началу работы и является начальным состоянием. Предназначено для анализа сгенерированной матрицы и перехода в следующее состояние. Из этого состояния автомат переходит в состояние S_1, если лидер-строка не находится на правильном месте, и в состояние S_2, если находится. Если матрица уже ступенчатая, то переходит в состояние S_3;
- S_1 – соответствует началу построения очередной ступени, а именно – постановке лидер-строки с наименьшим номером первого ненулевого элемента на первое место среди неступенчатых строк;
- S_2 – соответствует непосредственно построению очередной ступени. В данном состоянии лидер-строка находится на правильном месте, но есть ненулевые элементы под первым ненулевым элементом лидер-строки;
- S_3 – соответствует построению ступенчатой матрицы, когда матрица приведена к ступенчатому виду, но студент не подтвердил её готовность. В этом состоянии пользователь может совершать действия, не разрушающие ступенчатую матрицу;
- S_4 – соответствует окончанию работы, когда студент подтвердил готовность ступенчатой матрицы. В этом состоянии пользователь не может совершать никаких действий.

Множество переходов между состояниями T можно разделить на следующие группы:

- Переходы 1–3 предназначены для выбора состояния, с которого студент начинает работу. Выбор перехода зависит от состояния сгенерированной матрицы.
- Переходы 4–10 и 22 предназначены для случая, когда в ходе приведения матрицы к ступенчатому виду возникает необходимость поставить лидер-строку на первое место среди неступенчатых строк.
- Переходы 11, 12, 14–20 предназначены для случая, когда в ходе приведения матрицы к ступенчатому виду возникает необходимость уменьшить количество ненулевых элементов под первым ненулевым элементом лидер-строки w до нуля.
- Переходы 13 и 21 предназначены для случая, когда матрица не приведена к ступенчатому виду, однако студент подтвердил её готовность.
- Переходы 23–26 предназначены для случая, когда матрица была приведена к ступенчатому виду.
- Переходы 27–31 предназначены для случая, когда матрица приведена к ступенчатому виду, но студент решает сделать какие-либо улучшения или он не заметил, что матрица уже приведена к ступенчатому виду.
- Переход 32 предназначен для случая, когда приведение матрицы к ступенчатому виду завершено и студент подтверждает её готовность.

Заключение

В работе рассмотрено применение автоматной модели для обучения решению задачи построения ступенчатой матрицы, являющейся задачей с вариативной траекторией решения.

Использование расширенного автомата позволяет контролировать процесс решения и оценивать знания студента не только по конечному результату, а по ходу решения задачи. Кроме того, использование расширенного автомата облегчает построение обучающего диалога, предоставляя достаточное для построения диалога количество информации. Используя эти данные, студент или преподаватель может выяснить, какая часть вызывает затруднения у студента.

ЛИТЕРАТУРА

1. *Янковская А.Е., Разин В.В.* Система управления обучением на основе смешанных диагностических тестов с использованием технологии semantic web // Вестник Томского государственного университета. Управление, вычислительная техника и информатика. – 2016. – № 2(35). – С. 78–88.

2. *Янковская А.Е., Семенов М.Е.* Принятие решений в интеллектуальных обучающе - тестирующих системах, основанное на смешанных диагностических тестах // Искусственный интеллект и принятие решений. – 2012. – № 1. – С. 47–56.

3. *Yankovskaya A., Dementyev Y., Yamshanov A.* Application of Learning and Testing Intelligent System with Cognitive Component Based on Mixed Diagnostics Tests // XV International Conference "Linguistic and Cultural Studies: Traditions and Innovations", 2015. – P. 254–261.

4. *Yankovskaya A., Levin I., Fuks I.* Assessment of teaching and learning by mixed diagnostic testing // Proceedings of the Frontiers in Mathematics and Science Education Research Conference, 2014. – P. 86–93.

5. *Тумбинская М.В.* Модели, алгоритмы и комплекс программ оценивания знаний в интерактивной обучающей системе // Автореф. дис. ... к-та техн. наук. – Казань: Изд-во Казанского государственного технического университета, 2010. – 16 с.

6. *Murray T.* Authoring Intelligent Tutoring Systems: An analysis of the state of the // International Journal of Artificial Intelligence in Education, 1999. – P. 98–129.

7. *Тумбинская М.В.* Автоматизированные информационные системы в образовании как инструмент человеко-машинного взаимодействия: история и перспективы // Труды SORUCOM-2014. Третья международная конференция развитие вычислительной техники и ее программного обеспечения в России и странах бывшего СССР: история и перспективы. – 2014. – С. 369–376.

8. *Гиря И.А.* Понятийный граф как основа ведения модели знаний // Наука и образование: научное издание. – 2011. – № 5. – С. 1–9.

9. *Зверев Г.Н., Зверева Н.Н.* Семиотические модели и автоматизация конструирования педагогических тестов // Открытое образование. – 2013. – № 2. – С. 61–71.

10. *Таземдинов А.Д.* Автоматный подход в построении компьютерных обучающих диалогов // Информационно-управляющие системы. 2008. – № 4. – С. 42–48.

11. *Ильин В.Д., Соколов И.А.* Символьная модель системы знаний информатики в человеко-автоматной среде // Информатика и её применения. 2007. – Т. 1. – № 1. – С. 66–78.

12. *Евтушенко Н.В., Паршина Н.А., Янковская А.Е.* Автоматная модель контроля знаний в интеллектуальных обучающих системах // Математическое моделирование. Кибернетика. Информатика. – Томск: Издательство Томского университета, 1999. – С. 60–66.

13. *Yankovskaya A., Yevtushenko N.* Finite State Machine (FSM) - based Knowledge Representation in a Computer Tutoring System // New Media and Telematic Technologies for Education in East European Countries, 1996. – P. 63–69.

СИСТЕМА СЖАТИЯ ИНФОРМАЦИИ С ВЫБОРОМ ОПТИМАЛЬНОГО АЛГОРИТМА

А.А. Кожемякина, Н.Б. Буторина
Томский государственный университет

Введение

Количество информации постоянно растёт, в связи с чем нам приходится иметь дело с огромным объёмом данных, которые необходимо хранить и обрабатывать. Возможность эффективно использовать дисковое пространство рассматривалось ещё при рождении электронно-вычислительных систем. На протяжении всего времени эта проблема решалась различными алгоритмами сжатия данных.

Главный принцип сжатия основывается на том, что в любом файле, который содержит случайные данные, информация частично повторяется, а используя статистико-математические модели, можно выявить вероятность повторного вхождения определённой комбинаций символов.

В настоящее время существует множество алгоритмов сжатия информации, а также множество программ, которые используют данные алгоритмы. Но среднестатистический пользователь не владеет информацией о том, какой алгоритм более других подходит для того или иного файла. Целью нашей работы является создание приложения, которое автоматически подбирает алгоритм для выбранного пользователем файла.

Очевидно, что эффективность алгоритма должна зависеть от типа и объёма файла. Основываясь на результатах тестов, мы разработали приложение, которое, в зависимости от этих двух параметров, рекомендует пользователю оптимальный алгоритм, также предоставляя выбор приоритета по скорости или наилучшей степени сжатия.

В данной работе рассматривается исключительно сжатие без потерь. При таком сжатии закодированные данные восстанавливаются однозначно с точностью до бита. К таким алгоритмам относятся: LZMA2, BZip2, Defalte, PPMd. Именно их мы взяли за основу программы, ввиду их распространённости в наши дни и превосходством над более ранними и менее распространёнными алгоритмами современности.

Создание визуальной и функциональной оболочки программы происходило на языке программирования Python с использованием набора привязок графического фреймворка PyQt в среде программирования PyCharm от компании по производству программного обеспечения JetBrains.

Выбор этого программного обеспечения обусловлен следующими факторами:

1. Python ориентирован на повышение производительности пользователей и читаемость кода: синтаксис ядра минималистичен в то время, как стандартная библиотека содержит огромное количество полезных методов.

2. Среда разработки PyCharm обеспечивает простую организацию файлов и быструю скорость работы по сравнению с другими средами для программирования на языке Python.

3. PyQt представляет собой смесь языка программирования Python и библиотеки Qt. В то время как Qt – одна из самых мощных библиотек, которая содержит более 300 классов и почти 6000 функций и методов. Также эта библиотека является кроссплатформенной.

1. Обзор алгоритмов

LZMA2

LZMA2 является новой версией LZMA. Отличие между данными двумя алгоритмами в том, что в первом реализована многопоточность и хранение одновременно сжатый и несжатых данных, что, в свою очередь, экономит биты.

Форматом файлов, предназначенным для размещения сжатых данных, который по умолчанию использует алгоритм LZMA2, является формат .xz. В настоящее время этот формат дает возможность выбора между LZMA и LZMA2.

Алгоритм LZMA включает в себя:

1. Дельта-кодирование.
2. Алгоритм LZ77.
3. Интервальное кодирование.

Дельта-кодирование представляет данные как разницу между последовательными символами [1].

LZ77 рассматривает последовательности символов, и если какая-то последовательность встречается более одного раза, то он заменяет её повторные вхождения ссылкой на её первый экземпляр [2]. Как и во всех алгоритмах семейства LZ, в LZ77 используется словарь, хранящий последовательности, которые встречались ранее, т.е. он использует принцип «скользящего окна».

Сам словарь хранит данные о смещении, длине серий и символ расхождения. Смещение – это расстояние фразы от начала файла. Длина серии – это количество символов, которое принадлежит фразе, если считать от смещения. Символ расхождения показывает на то, что на выход пришла фраза, которая похожа на ту, которая обозначена смещением и длиной [3].

При интервальном кодировании выделяется некоторый диапазон символов, и оценивается вероятность вхождения каждого. Таким образом, этот алгоритм преобразовывает все символы сообщения в одно число.

Название алгоритма происходит от «Lempel-Ziv Markov chain Algorithm».

BZip2

Bzip2 – это утилита командной строки для сжатия данных, разработанная Джулианом Стюардом в 1996 г. При сжатии к оригинальному расширению файла добавляется суффикс .bz2. Прежде чем приступить к кодированию, выполняются предварительные преобразования.

Алгоритм состоит из этапов:

1. Деление на блоки фиксированного размера.
2. Преобразование Барроуза-Уилера.
3. Преобразование MTF.
4. Кодирование Хаффмана.

Размер блока можно задать при помощи аргументов командной строки. Такие блоки сжимаются независимо друг от друга и записываются подряд. При этом в начале каждого блока используется последовательность 0x314159265359 (в кодировке ASCII при выравнивании на границу байта отображается как «1AY&SY»), что является записью первых цифр числа пи в формате BCD. А в конце каждого блока используется запись первых цифр корня из числа пи, т.е. последовательность 0x177245385090.

Преобразование Барроуза-Уилера (или BWT) – это алгоритм, который преобразовывает исходные данные так, что на выходе повторяющиеся подстроки образуют последовательность из повторяющихся символов [2], т.е.:

- • создаётся массив строк;
- • в этом массиве сохраняются всевозможные преобразования строки, которая подаётся на вход;

• • отсортировка массива;
• • на выходе получаем последний столбец.

Преобразование MTF(move-to-front) заменяет каждый входной символ на его номер в специальном стеке, в который заносятся символы, использованные недавно.

Алгоритм Хаффмана основан на частоте появления символа в последовательности. Символ, который встречается в строке чаще всего получает самый короткий код, а символ, который встречается реже остальных – наоборот, самый длинный код. Это нужно для того, чтобы самые частые символы занимали меньше места.

Сначала вычисляется частота каждого символа. Затем создаются узлы бинарного дерева для всех знаков и добавляются в очередь, используя частоту в качестве приоритета. После достаются два первых элемента из очереди и связываются, создавая новый узел дерева, в котором они оба будут потомками, а приоритет нового узла будет равен сумме их приоритетов. Далее получившийся новый узел отправляется обратно в очередь.

Повторяя шаги, описанные выше, последовательно, получим оптимальное кодовое дерево, на основе которого строится отображение код-символа. Чтобы получить код для любого символа, надо пройтись по дереву и для каждого перехода выбрать 1, если мы идём вправо, и 0, если мы идём влево.

Deflate

Алгоритм Deflate создан Жан-Лу Галли и Марком Адлером и впервые выпущен в 1992 году. В 1993 году была создана версия 1.0. Создаёт файлы в формате .gz и комбинирует алгоритм LZ77 и кодирование Хаффмана, которые были рассмотрены выше.

PPMd

Метод сжатия PPMd – одна из версий алгоритма PPM. PPM реализует сжатие без потерь, используя статистические методы. Он основан на контекстном моделировании и предсказании. Под контекстом подразумевается множество символов в несжатом потоке, предшествующих данному, чтобы предсказывать значение символа на основе статистических данных.

Длина контекста, использующегося при предсказании, зачастую сильно ограничена. Если предсказание символа по контексту из n символов не может быть произведено, то происходит попытка предсказать его с помощью $n-1$ символов. Рекурсивный переход к моделям с меньшим порядком происходит, пока предсказание не произойдёт в одной из моделей, либо когда контекст станет нулевой длины. В случае, когда контекст становиться нулевой длины, вероятность символа определяется исключительно из частоты его появления в сжимаемом потоке данных. Для получения наилучшего значения вероятности элемента необходимо учитывать контексты разных длин.

В алгоритме сжатия PPM реализуется вариант стратегии перемешивания, т.е. оценки вероятностей, сделанные на основе анализа контекстов разных длин, формируются в одну общую вероятность. Сформировавшаяся оценка кодируется любым алгоритмом энтропийного кодирования. На этом этапе, непосредственно, и происходит сжатие.

Рассматриваемый нами алгоритм PPMd увеличивает псевдосчётчик нового символа каждый раз, когда, действительно, в потоке появляется новый символ. Другими словами, PPMd оценивает вероятность появления нового символа как отношение числа уникальных символов к общему числу используемых символов. В то время, как PPM полагают счётчик нового символа равным фиксированной величине, например, единице [4].

2. Функционал приложения

Разработанное нами приложение анализирует выбранный пользователем файл, выявляет его размер и расширение и на основе этих двух параметров делает вывод о том,

какой метод сжатия наилучшим образом подходит для данного файла по одному из приоритетов – скорости сжатия или степени сжатия.

Ниже представлен интерфейс программы, разработанный с помощью PyQt.

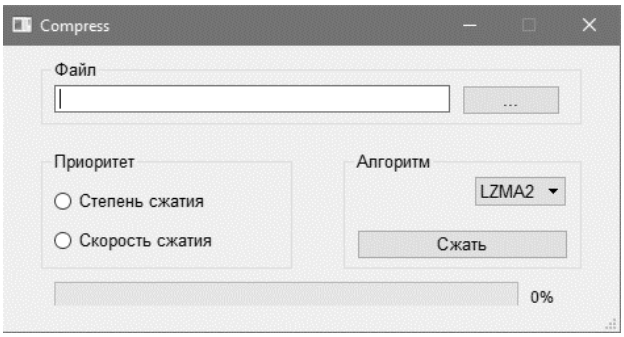


Рис. 1. Пользовательский интерфейс программы

Объекты распределены по контейнерам – элемент графического интерфейса, называемый groupBox.

Контейнер под названием «Файл» содержит кнопку, при нажатии на которую вызывается диалоговое окно выбора файла и строку, куда будет помещена директория этого файла.

Контейнер с именем «Приоритет» содержит кнопки-переключатели, которые дают пользователю выбор между приоритетами.

В контейнере «Алгоритм» расположены выпадающий список с различными методами сжатия и кнопка, которая непосредственно производит сжатие по выбранной директории и выбранному алгоритму. В выпадающем списке алгоритмы расположены в порядке убывания – от наиболее эффективному к менее эффективному.

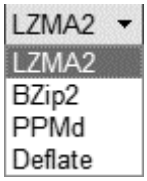


Рис. 2. Выпадающее меню

В самом низу интерфейса размещен индикатор прогресса – элемент графического интерфейса пользователя, который показывает, на сколько процентов в данный момент времени выполнена задача. В нашем случае под задачей подразумевается сжатие файла.

По завершении сжатия, программа не закрывается, предоставляя пользователю выбрать другой файл или закрыть её самостоятельно.

3. Анализ результатов тестирования по степени сжатия

Для тестирования использовались файлы разного типа: txt, doc, docx, xls, xlsx, ppt и pptx. Каждый тип условно разбили на маленький, средний и большой размер файла. Используя все алгоритмы произвели сжатие каждого файла с замером времени. Запоминался исходный размер файла и размер после компрессии, после чего на основе этих данных высчитывался коэффициент сжатия в процентах. Ключевую роль в работе играют два показателя – это скорость выполнения программы и пространство на диске, занимаемое данными. Поэтому при тестировании во внимание принимались два критерия:

1. Коэффициент сжатия.
2. Скорость сжатия.

В табл. 1 приведены результаты тестов, где первый столбец означает формат файла, второй – начальный размер файла и последующие столбцы – степень сжатия в процентах после применения алгоритмов, выполняющих компрессию.

Таблица 1

Сравнение степеней сжатия

		BZip2	Deflate	LZMA2	PPMd
txt	s	64,03	59,93	61,07	66,98
	m	71,96	65,58	69,39	74,97
	l	73,38	63,94	70,00	76,06
doc	s	24,49	26,53	27,55	25,51
	m	6,30	8,13	8,74	6,10
	l	3,49	5,08	5,48	3,39
docx	s	0,92	1,83	0,92	0,92
	m	-0,56	0,00	-1,11	-0,19
	l	-0,40	0,00	-1,20	0,00
xls	s	76,35	73,85	83,40	78,40
	m	81,75	75,91	91,97	82,48
	l	82,65	76,86	95,43	81,58
xlsx	s	2,97	9,14	10,13	6,94
	m	28,57	33,33	37,86	37,14
	l	50,81	48,11	57,84	57,30
ppt	s	36,21	39,09	41,15	38,68
	m	17,48	20,00	21,36	18,06
	l	10,02	11,61	12,40	9,62
pptx	s	6,71	9,15	9,15	6,10
	m	2,10	3,50	2,57	0,70
	l	0,66	1,55	0,55	0,00

Для более наглядного представления на основе таблицы 1 была построена следующая диаграмма:

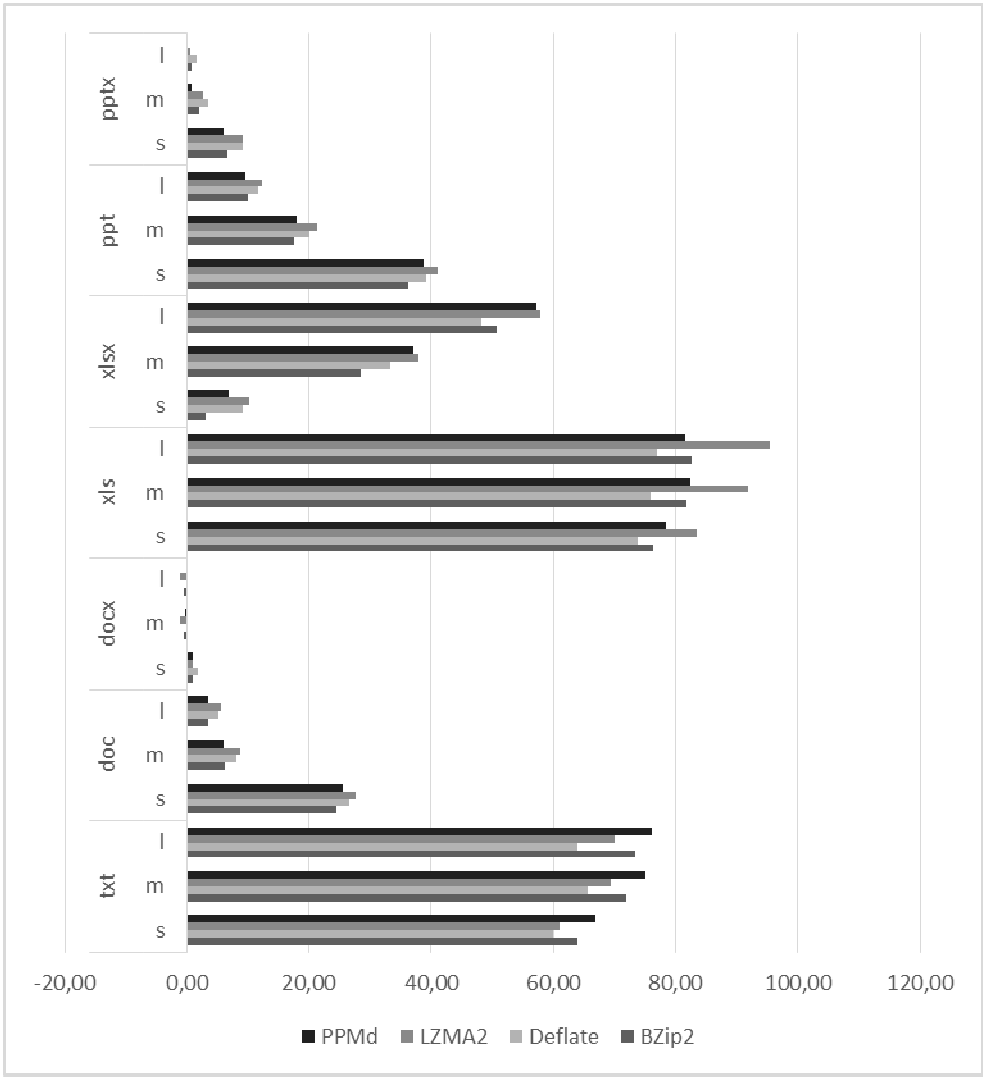


Рис. 3. Диаграмма степени сжатия

Здесь буквой «s» обозначается маленький объем файла, «m» – средний и «l» – большой. Для расширения .txt были взяты файлы размером 15, 150 и 700 килобайт. Файлы .doc и .docx тестировались на объёмах 100 килобайт, 500 килобайт и 1 мегабайт. В таблицы Microsoft Excel было занесено по 100, 1000 и 5000 строкам, что в формате .xls составило 20, 150 и 650 килобайт соответственно, а для более нового .xlsx – 10, 40 и 200 килобайт. Презентации .ppt и .pptx сжимались в размерах 250, 500 и 1000 килобайт.

Из таблицы и диаграммы мы видим, что для сжатия текстового редактора лучше всего подходит метод PPMd. Также очевидно, что LZMA2 лучше других справляется с такими файлами, как таблицы Ecxel, документы в формате .doc и презентации PowerPoint. С Microsoft Word более поздней версии, а именно в формате .docx, дела обстоят иначе – при больших размерах такие файлы могут сжать не все алгоритмы, а некоторые и вовсе увеличивают первоначальный размер. Такие алгоритмы были исключены из списка выбора метода сжатия при указании директории на подобные файлы.

4. Функция принадлежности

В процессе написания программы для определения степени принадлежности объёма файла к той или иной категории были использованы элементы нечёткой логики, а именно была использована функция принадлежности. Она получила свою популярность в связи с наличием нечётких приближённых рассуждений при описании человеком процессов, систем и объектов. Существует ряд формул для задания функции принадлежности, из которых мы выбрали треугольную как наиболее подходящую к данной задаче. Такая функция принадлежности определяется тройкой чисел (a, b, c), где $a \le b \le c$, и значение в точке x вычисляется по формуле:

$$\mu(x) = \begin{cases} 1 - \dfrac{b-x}{b-a}, & x \in [a,b], \\ 1 - \dfrac{x-b}{c-b}, & x \in [b,c], \\ 0, & x \notin [a,c]. \end{cases}$$

Здесь числа a и c означают границы интервалов объёмов, b – значение размера файла, на котором проводились тесты. Следовательно, чем ближе значение объёма выбранного файла к объёму тестированного, тем больше файл подходит под ту или иную категорию.

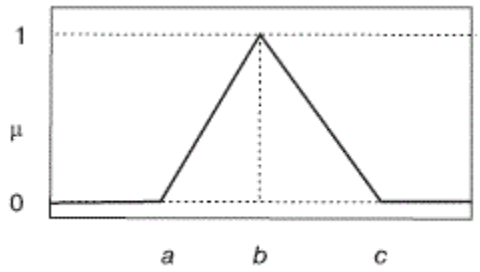


Рис. 4. Треугольная функция принадлежности

5. Анализ результатов тестирования по скорости сжатия

При рассмотрении второго приоритета – скорости сжатия, были построены следующая таблица и диаграмма, значения времени в которых указаны в секундах:

Таблица 2

Сравнение скорости сжатия

		BZip2	Deflate	LZMA2	PPMd
txt	s	0,1	0,24	0,02	0,30
	m	0,07	0,06	0,09	0,14
	l	0,11	0,1	0,25	0,34
doc	s	0,02	0,01	0,04	0,11
	m	0,1	0,03	0,11	0,27
	l	0,22	0,07	0,23	0,30
docx	s	0,04	0,01	0,05	0,07
	m	0,13	0,02	0,12	0,19
	l	0,22	0,05	0,2	0,34
xls	s	0,01	0,02	0,02	0,06
	m	0,01	0,04	0,06	0,12
	l	0,06	0,17	0,23	0,29
xlsx	s	0	0,02	0,03	0,08
	m	0,01	0,003	0,04	0,12

	l	0,17	0,35	0,14	0,40
ppt	s	0,06	0,03	0,07	0,10
	m	0,11	0,04	0,12	0,14
	l	0,44	0,5	0,37	0,56
pptx	s	0,05	0,01	0,05	0,07
	m	0,1	0,02	0,09	0,23
	l	0,27	0,07	0,19	0,38

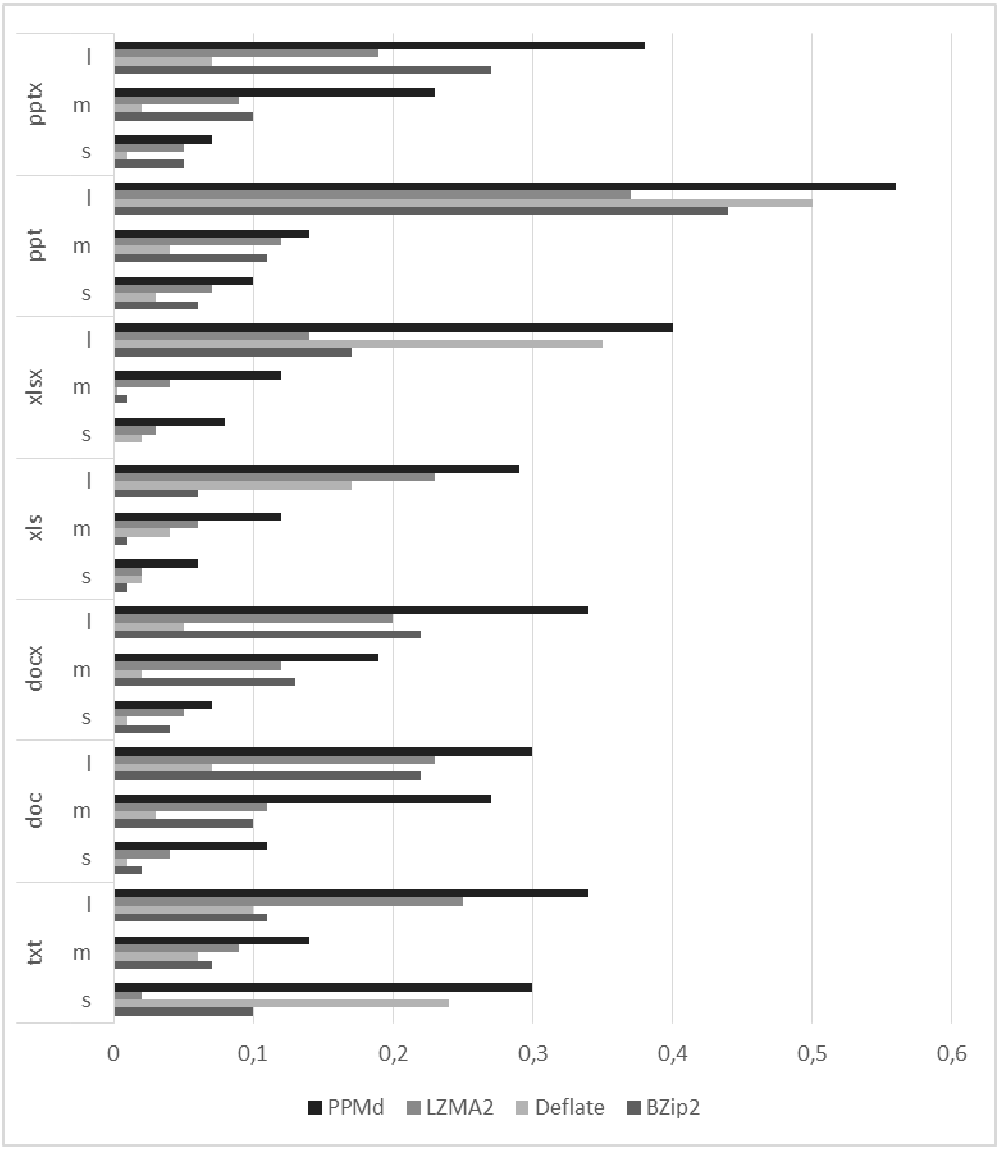


Рис. 5. Диаграмма скорости сжатия

На основе этих данных мы можем сделать вывод о том, что метод PPMd хоть и сжимает текстовые редакторы лучше остальных алгоритмов, но при этом имеет недостаток в скорости. По большей части алгоритм Deflate показывает наилучший результат по этому критерию.

Заключение

В настоящей работе были рассмотрены различные способы компрессии. Все они при тестировании дали разные показатели времени и степени сжатия. Основываясь на

результатах, представленных в таблице 1 и таблице 2, а также на рис. 3 и рис. 4, можно рекомендовать для фалов .txt алгоритм сжатия PPMd для наилучшей компрессии, но не лучшей скорости, и метод LZMA2, как лидер среди сжатия данных с таблицами, документами и презентациями. Как самый быстрый алгоритм, чаще всего показывает себя Deflate.

ЛИТЕРАТУРА

1. Академик [Электронный ресурс] : Дельта-кодирование – 2014. – URL: https://dic.academic.ru/dic.nsf/ruwiki/41976.
2. Habr [Электронный ресурс] : Алгоритмы сжатия данных без потерь, часть 2 – 2014. – URL: https://habr.com/post/235553/.
3. Habr [Электронный ресурс] : Простейшие алгоритмы сжатия: RLE и LZ77 – 2012. – URL: https://habr.com/post/141827/.
4. Wikipedia [Электронный ресурс] : Алгоритм сжатия PPM – 2018. – URL: https://ru.wikipedia.org/wiki/Алгоритм_сжатия_PPM.

ВИРТУАЛЬНАЯ МАШИНА EMERAVM

Е.П. Чалых, С.И. Самохина
Томский государственный университет

Введение

На данный момент в мире существует большое число языков программирования, каждый из них создан с определённой целью. Изначальной целью проекта было создание мощного и быстрого языка программирования с поддержкой математических вычислений, таких как дифференцирование, интегрирование, работа с матрицами и другие. Но первая проблема нового языка возникла при выборе его типа: интерпретируемый или компилируемый. Очевидный минус первого типа заключается в том, что процесс выполнения кода будет довольно долгим, особенно на очень больших файлах исходного кода. Но главное достоинство – интерпретатор языка можно сделать кроссплатформенным. Компилируемый язык программирования будет работать гораздо быстрей, т.к. из исходного кода мы получим исполняемый файл. А минус такого языка, естественно, платформозависимость.

Очевидно, для получения кроссплатформенного и быстрого языка, нужно использовать промежуточное состояние исходного кода и исполняемого файла. Такое состояние называется байт-кодом[1]. Байт-код – это последовательность байтов, каждый из которых характеризует конкретную инструкцию выполнения. Чтобы сформировать байт-код, нужен определенный транслятор, а для выполнения полученного байт-кода – определенный интерпретатор, называемый виртуальной машиной[2].

1. Постановка задачи

Первый этап создания языка программирования – это создание работоспособной, желательно кроссплатформенной, виртуальной машины. И поскольку язык будет использоваться в математических целях, виртуальная машина должна поддерживать большие математические вычисления.

Название проекта происходит от названия языка и сокращения от virtual machine – EmeraVM, или сокращенно EVM.

Кроссплатформенность можно достичь с помощью другого кроссплатформенного языка, в данном случае используется язык Kotlin[3], работающий на JVM[4] – виртуальной машине Java. Соответственно, EVM будет работать на любой операционной системе, где установлена Java. Инструментом создания EVM является мощная интегрированная среда разработки IntelliJ IDEA[5].

2. Архитектура EVM

Реализация EVM разделена на 3 основных модуля: lang, vm, asm (рис. 1).

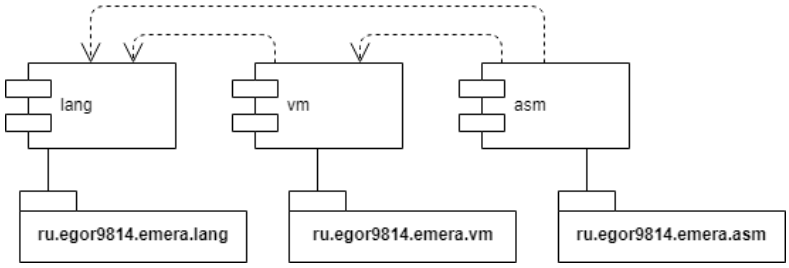


Рис. 1. Модули EVM и их зависимость

Модуль «lang» является главным, т. к. он отвечает за работу с типами данных. Рассмотрим его структуру (рис. 2).

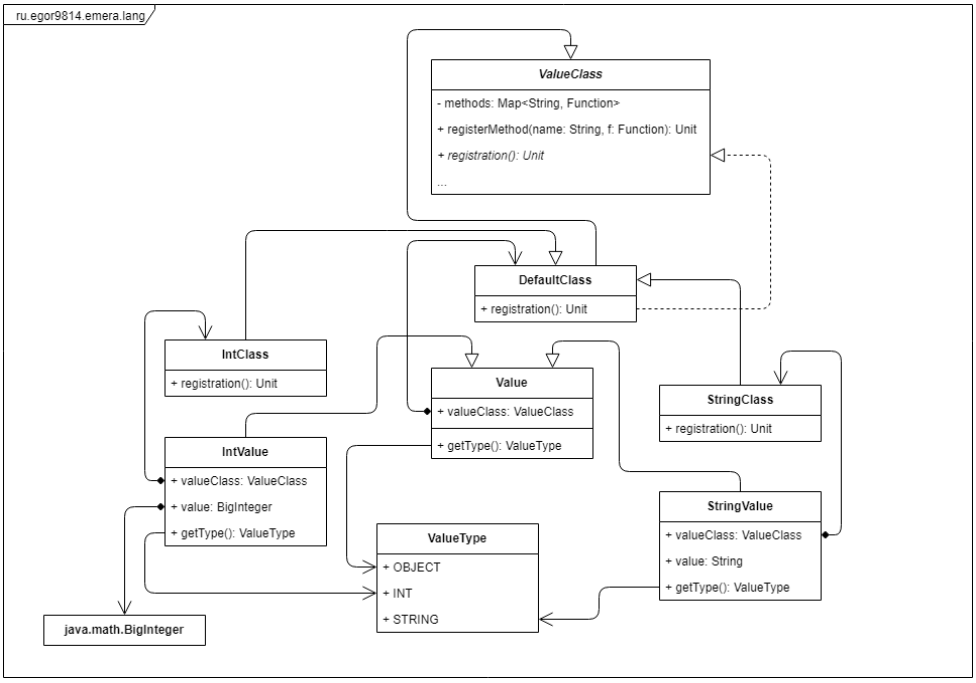


Рис. 2. Классы модуля «lang»

Класс «ValueType» определяет доступные типы данных: STRING – строка, INT – целое число, OBJECT – объект.

Класс «ValueClass» определяет доступные методы для конкретного класса. Класс «DefaultClass» регистрирует метод «toString()», который представляет типе в виде строки. Класс «StringClass» определяет методы работы со строками. Класс «IntClass» определяет методы работы с целыми числами.

Класс «Value» определяет тип данных и способ работы с ним. Класс «IntValue» содержит поле типа «BigInteger», что позволяет работать с большими целыми числами. Класс «StringValue» содержит поле типа «String», что позволяет работать со строками.

В данном модуле также есть класс для работы с объектами типа «Value» - EmeraEnvironment (рис. 3).

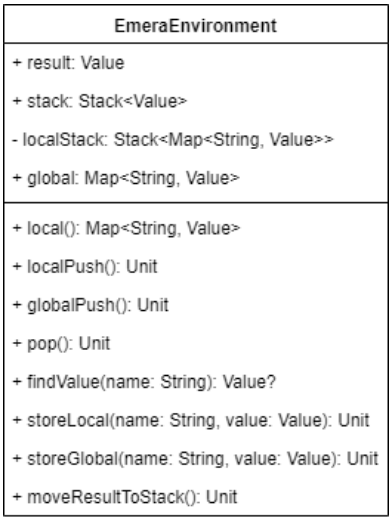


Рис. 3. Класс «EmeraEnvironment»

Этот класс позволяет управлять стеком значений, сохранять глобально или локально объекты по ключу, а также загружать в локальный стек. По умолчанию результат выполнения любой операции заносится в регистр результата «result», и если есть необходимость его использовать, то с помощью метода «moveResultToStack()» можно поместить результат в локальный стек.

Модуль «vm» (рис. 4, 5) позволяет прочитать байт-код и выполнить его. Интерфейс «IInstruction» предоставляет абстракцию, которую реализует конкретная инструкция, в зависимости от кода операции класса «ByteCodeV1». Класс «InstructionEnvironment» позволяет контролировать процесс добавления и выполнения инструкций. Класс «Interpreter» отвечает за чтение байт-кода и составление списка инструкций. Класс «ByteCodeLauncher» использует экземпляр класса «InstructionEnvironment», полученный из интерпретатора «Interpreter», для дальнейшего выполнения инструкций в методе «execute()». Если результат вызова этого метода окажется ложным (False), то с помощью метода «getError()» можно получить текст ошибки.

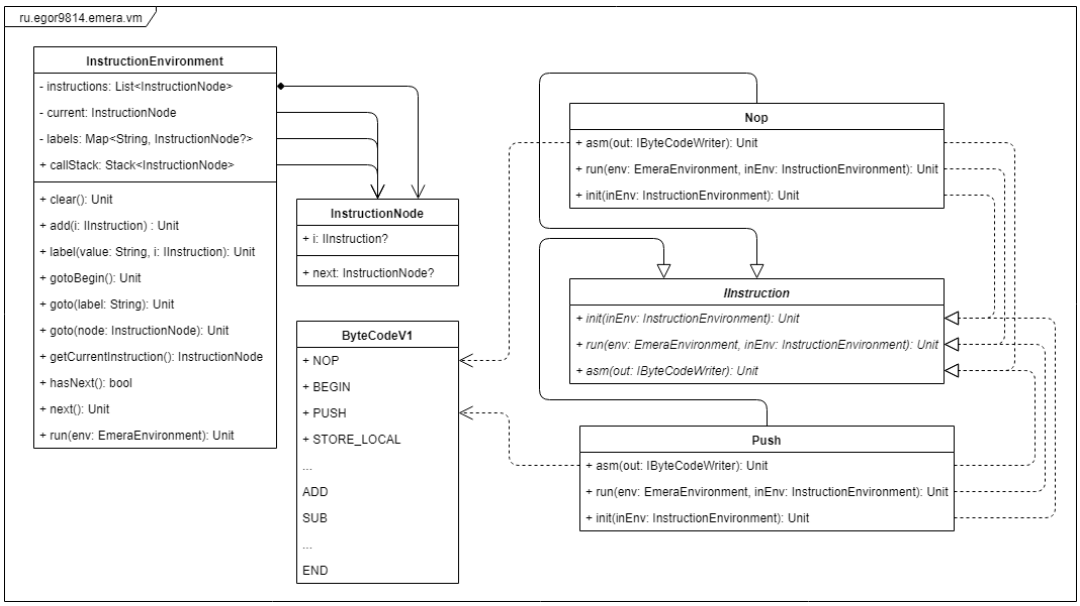


Рис. 4. Первая часть модуля «vm»

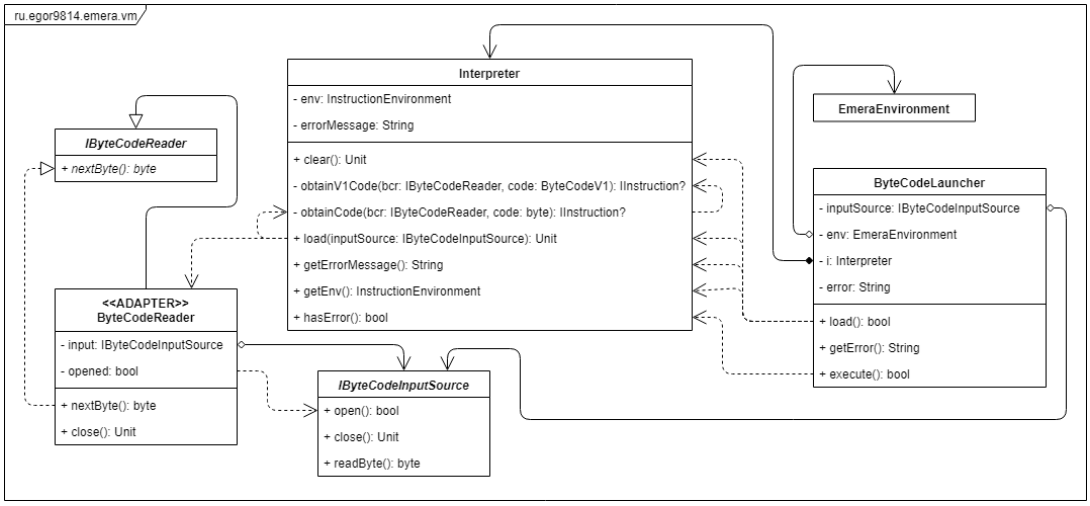


Рис. 5. Вторая часть модуля «vm»

На данный момент у нас есть способ запустить байт-код, но как сформировать его? Первый способ – это вручную прописывать байты в файле. Это очень неудобно, трудоёмко, и велика вероятность ошибки. Второй – использовать компилятор, который преобразует исходный код в байт-код. Это наилучший вариант, но языка, как и компилятора, ещё не существует. На этот случай был разработан модуль «asm» (рис. 6). Он позволяет преобразовать текстовые инструкции, понятные для человека, в байт-код. В качестве таких инструкций выступает низкоуровневый язык программирования похожий на Ассемблер (рис. 7).

Класс «Tokenizer» позволяет разбить одну строковую линю на специальные последовательности – токены (класс «Token»), а именно: ключевое слово, строка, число. Класс «Interpreter» проходит весь исходный код по строкам и с помощью «Tokenizer», определяет тип инструкции, после чего формирует список инструкций (класс «IInstruction»).

Класс «Assembler» – это генератор байт-кода. Первый метод «interpret()» анализирует исходный код с помощью класса «Interpreter». Если во время анализа произойдет ошибка, то результатом этой функции будет текст этой ошибки. Метод «asm()» проходит по каждой инструкции и вызывает метод «IInstruction:: asm(IByteCodeWriter)», который в свою очередь формирует и записывает фрагмент байт-кода в интерфейс вывода «IByteCodeWriter».

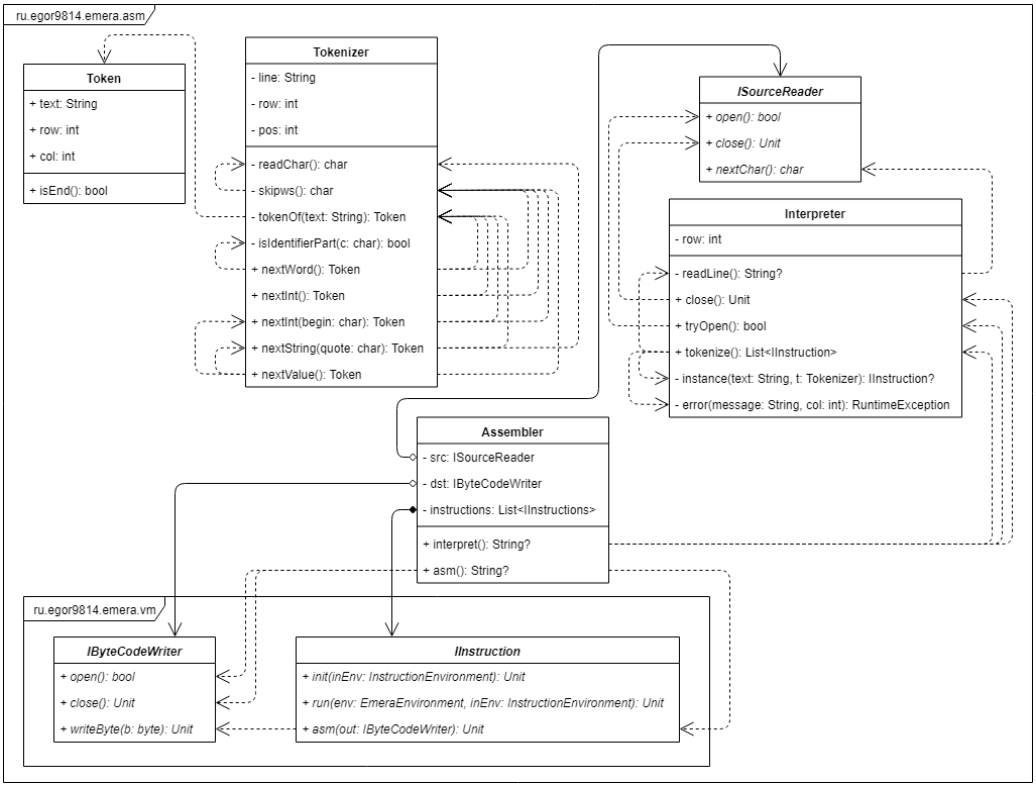


Рис. 6. Классы модуля «asm»

```
Open
 2 push 10
 3 global "a"
 4
 5 push 20
 6 global "b"
 7
 8 lbl __add
 9   load "b"
10   load "a"
11   add
12   mrs
13   global "add"
14
15 lbl __sub
16   load "b"
17   load "a"
18   sub
19   mrs
20   global "sub"
21
22 lbl __mul
23   load "b"
24   load "a"
25   mul
26   mrs
27   global "mul"
28
29 lbl __div
30   load "b"
31   load "a"
32   div
33   mrs
34   global "div"
35
36 lbl __pow
37   load "b"
38   load "a"
39   pow
40   mrs
41   global "pow"
```

Рис. 7. Пример исходного кода для EVM

3. Способ запуска виртуальной машины

Запустить EVM на данный момент можно через терминал с помощью скрипта «emera.bat» на Windows, и «./emera» на MacOS или Linux (рис. 8).

Команда «./emera –help» позволит увидеть доступные параметры для запуска EVM.

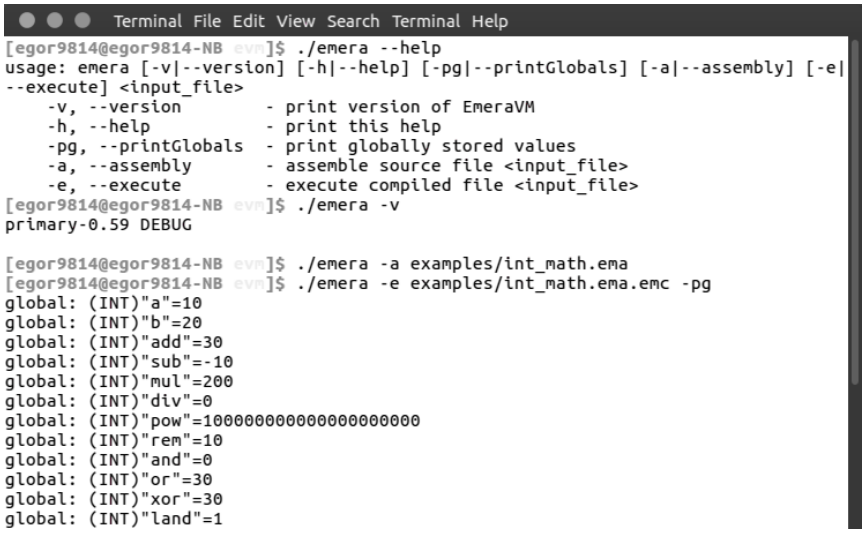


Рис. 8. Пример запуска EVM

Заключение

Данный проект был создан как компонент языка программирования «EmeraLang», в нём реализована поддержка больших целых чисел, работа со строками, условные и безусловные переходы и другое. Т. к. язык направлен на математические вычисления, то в виртуальную машину также будет добавлена поддержка больших чисел с плавающей точкой, комплексных чисел, матриц, а также булева алгебра.

ЛИТЕРАТУРА

1. Байт-код [Электронный ресурс]: сайт / Википедия – URL: https://ru.wikipedia.org/wiki/Байт-код (дата обращения: 21.05.2019).

2. Виртуальная машина [Электронный ресурс]: сайт / Википедия – URL: https://ru.wikipedia.org/wiki/Виртуальная_машина (дата обращения: 21.05.2019).

3. Язык программирования Kotlin [Электронный ресурс]: сайт / KotlinLang – URL: https://kotlinlang.org/ (дата обращения: 21.05.2019).

4. JVM [Электронный ресурс]: сайт / Википедия – URL: https://ru.wikipedia.org/wiki/Java_Virtual_Machine (дата обращения: 21.05.2019).

5. IntelliJ IDEA [Электронный ресурс]: сайт / JetBrains – URL: https://www.jetbrains.com/idea/, (дата обращения: 21.05.2019).

ВЛИЯНИЕ ВИБРАТО НА КАЧЕСТВО РАСПОЗНАВАНИЯ НОТ В ВОКАЛЬНОМ ИСПОЛНЕНИИ

А.Ю. Якимук

Томский государственный университет систем управления и радиоэлектроники

Введение

Применение программных средств в задачах обучения пользователей распространено во многих областях знаний. Точность работы таких программ крайне важна, поскольку ошибочно работающий алгоритм способен ввести в заблуждение ученика и сформировать компетенцию некорректно. Особенно сильно это проявляется в сфере речевых технологий. В случае, если пользователь неправильно выучит произношение слов при изучении языка или не сформирует музыкальный слух корректно, полученные в программе навыки окажутся бесполезными, а в некоторых ситуациях даже вредными. В данной работе рассматривается вопрос определения причин в получаемом проценте ошибок при идентификации спетой диктором ноты.

1. Постановка задачи

Программа для каждой аудиозаписи выводит график частот основного тона в каждый момент времени с сегментированием на ноты (рис. 1) и последовательность спетых нот в abc-нотации (рис. 2).

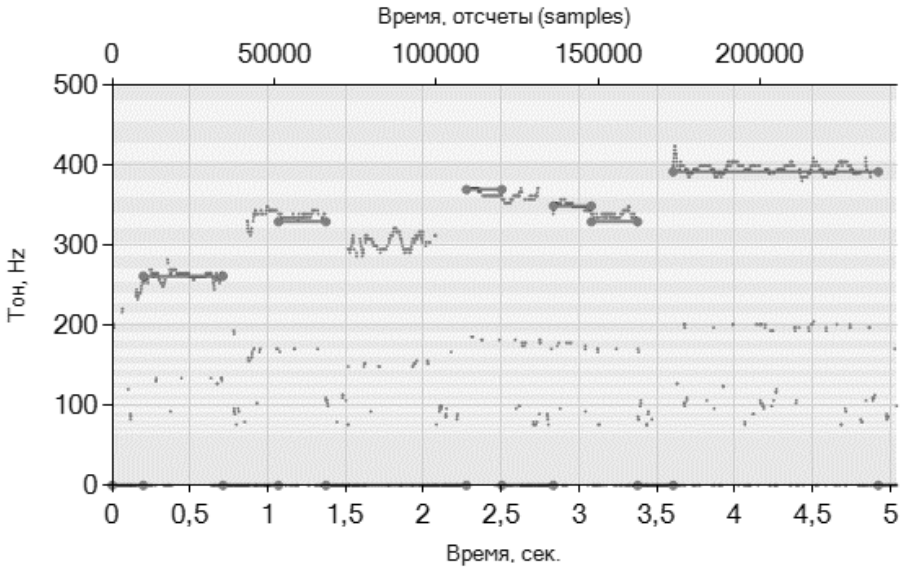


Рис. 1. Пример результатов алгоритма идентификации нот (частотная область)

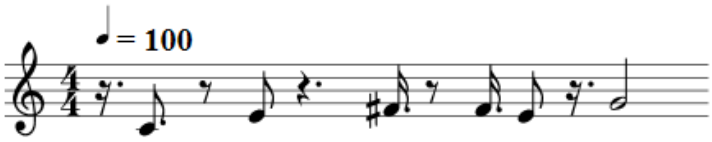


Рис. 2. Пример результатов алгоритма идентификации нот (полученные ноты)

В исследовании [1] было проведено тестирование программного комплекса по распознаванию нот, показавшее, что в диапазоне 70–400 Гц применяемый алгоритм способен верно идентифицировать 85% спетых диктором нот. Отметим, что используемый алгоритм измерения частоты основного тона (ЧОТ) [2], результаты которого используются при определении нот, имеет погрешность вычисления ЧОТ менее 1%, а надежность определения вокализованных участков при обработке речевого сигнала составляет 89–93%. Прогнозируемая точность алгоритма распознавания нот при вышеописанных характеристиках способна превысить 95%. Полученная разница в количестве ошибок стала основанием для дополнительного исследования по поиску источников ошибок в работе разработанного алгоритма, применяемого в программном комплексе, определяющем ноты в вокальном исполнении.

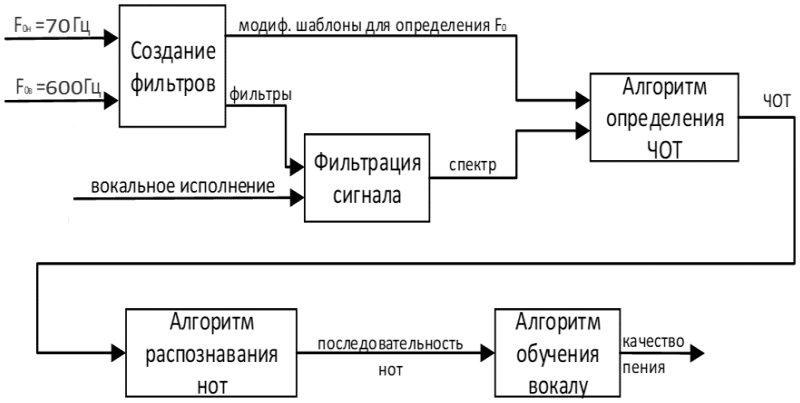


Рис. 3. Этапы работы программного комплекса

Было решено провести отдельно тестирование определения ЧОТ, сегментации и распознавания нот. Таким образом можно будет понять, на каком из этапов (рис. 3) возникают проблемы, приводящие к 10% ошибок в распознавании нот.

Ранее в рамках исследования уже проводилось изучение правильности работы некоторых из этапов работы программного комплекса. Так, в работе [3] уделяется внимание алгоритму, идентифицирующему частоту основного тона в каждый момент времени. По результатам проведенной работы было определено, что относительная ошибка определения частоты основного тона в пределах до 600 Гц не превышает ожидаемого 1% (табл. 1).

Таблица 1

Относительная ошибка определения частоты основного тона

Частота основного тона синусоидального сигнала, Гц	Относительная погрешность, %
150	0
200	0
250	0,4
300	0,33
350	0,29
400	0,5
450	0,44
500	0,6
550	0,73
600	0,83

Также в работе [4] было проведено исследование корректности метода определения границ для идентификации ноты. Целью работы стало определение влияния выбора метода шкалирования для идентификации нот на получаемые результаты (рис. 4).

Частота звучания ноты	Среднее											
	Гармоническое		Логарифмическое		Геометрическое		Арифметическое		Квадратическое		Кубическое	
	Ниж.гр.	Верх.гр.	Ниж.гр.	Верх.гр.	Ниж.гр.	Верх.гр.	Ниж.гр.	Верх.гр.	Ниж.гр.	Верх.гр.	Ниж.гр.	Верх.гр.
698,456	678,290	718,623	678,571	718,921	678,572	718,923	678,856	719,223	679,138	719,522	679,421	719,822
739,989	718,623	761,355	718,921	761,671	718,923	761,672	719,223	761,990	719,522	762,308	719,822	762,625
783,991	761,355	806,627	761,671	806,962	761,672	806,963	761,990	807,300	762,308	807,636	762,625	807,972
830,609	806,627	854,591	806,962	854,946	806,963	854,948	807,300	855,305	807,636	855,661	807,972	856,017

Рис. 4. Шкалы идентификации нот

Полученные в предыдущих работах результаты не дали ответ на вопрос о причинах возникновения ошибок в работе алгоритма. В представленной работе было принято решение уделить внимание алгоритму идентификации нот.

2. Влияние вибрато на идентификацию нот

Одним из положительных моментов в работе [5] стало увеличение охватываемого диапазона до 600 Гц. Это позволило использовать аудиозаписи с вокальным исполнением нот, спетых выше 400 Гц. Отметим, что на предыдущих этапах исследования качества распознавания нот [5] использовались аудиозаписи с пением, которое соответствует упражнениям, преподаваемым в музыкальной школе. За основу были взяты задания похожие на те, которые применяются при обучении вокалу. Для дальнейшей работы было решено придерживаться этой стратегии.

Ключевым моментом в определении качества идентификации нот стало обнаружение в вокале дикторов вибрато, воспринимаемого экспертами как правильно исполненная нота, а программой как шум. Спецификой работы одного из этапов алгоритма идентификации нот, отвечающего за определение принадлежности вокализованного участка к ноте, заключается в ориентированности на чистое исполнение.

Исполнить ноту так, чтобы все частоты в момент пения находились в пределах полосы, отведенной под нее, достаточно сложно. По этой причине учитываются соседние ноты выше и ниже исполняемой. На рис. 5, 6 спектр частот разделен на участки, соответствующие границам нот. Каждый участок выделен горизонтальной полосой на фоне графика частот. По оси X – время, по оси Y – частоты.

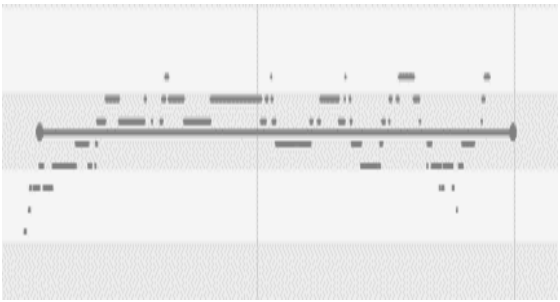


Рис. 5. Участок графика с нотой, спетой диктором без вибрато в голосе

Как можно видеть на рис. 5, сегмент, спетый в пределах одной ноты, был распознан программой. Наличие моментов, попавших в соседние ноты, незначительно велико и не оказало влияние на результат идентификации. Преобладающее число распознанных частот основного тона оказалось в диапазоне спетой ноты. С учетом прописанных в алгоритме требований по точности исполнения, программа смогла сделать вывод о спетой ноте. В случае, если количество фрагментов в каждом из 3 участков оказалось приблизительно одинаковым, алгоритм воспринимает весь вокализованный сегмент как шум.

Рис. 6. Участок графика с нотой, спетой диктором с вибрато в голосе

Эту ситуацию можно увидеть на рис. 6. Спетая диктором нота была исполнена с вибрато в голосе. По определению признаком присутствия в пении вибрато являются колебания в пределах полутона (соседние ноты) с частотой колебаний от 5 до 7 Гц [6]. Как можно заметить на участке, спетом с применением вибрато, присутствуют колебания в 4 соседних нотах. Кроме того, оценка вклада в каждую из нот у спетого сегмента меньше требуемого показателя.

Считается, что к наличию вибрато в голосе склонны высокие голоса, и наиболее распространен прием среди обладательниц сопрано. Этот факт объясняет отсутствие таких колебаний среди аудиозаписей в диапазоне от 70 до 400 Гц.

Следует отметить, что наличие колебаний в голосе при пении может свидетельствовать не только о профессионализме певца. В случае, если колебания происходят чаще (от 8 до 9 Гц) – этот эффект считается тремоляцией. Также недостатком является слишком медленное изменение частоты (3–4 Гц) в пении, воспринимаемое как «качание звука». Одной из задач повышения качества пения может стать доработка алгоритма с целью определения частоты изменения в пении ноты. Наблюдение учеником графика с информацией о скорости колебаний в голосе с рекомендациями по достижению желаемого эффекта способно создать положительную биологическую обратную связь [7].

Заключение

В рамках проведенного исследования было определено, что в работе алгоритма идентификации нот в диапазоне от 70 до 400 Гц отсутствуют видимые причины, которые могут способствовать в расхождении ожидаемой точности распознавания нот и получаемом результате работы комплекса. Авторами был обнаружен вклад в снижение процента правильно распознанных нот в диапазоне от 400 до 600 Гц, вызванный наличием колебаний в голосе диктора. Была определена задача по дальнейшему развитию программного комплекса по распознаванию нот, в которую входит определение типа колебаний в исполнении ноты. Для решения поставленной задачи необходимо реализовать алгоритм, определяющий количество изменений в частоте основного тона в секунду и принимающий решение по отнесению колебаний к одной из трех групп: «качание голоса», вибрато или тремоляция.

ЛИТЕРАТУРА

1. *Якимук А.Ю., Конев А.А., Осипов А.О.* Программный комплекс для автоматизации моделирования сегментации речевых сигналов и вокальных исполнений // Вестник Иркутского государственного технического университета. – 2017. – Т. 21. – № 10 (129). – С. 53–64.

2. *Конев А.А.* Модель и алгоритмы анализа и сегментации речевого сигнала: автореф. дис. … канд. техн. наук. – Томск, 2007.

3. *Якимук А.Ю., Егошин Н.С., Осипов А.О., Боков И.М.* Повышение качества идентификации нот в автоматизированной системе распознавания вокала // Электронные средства и системы управления. – 2016. – № 1–2. С. 29–32.

4. *Томская К.М.* Определение метода шкалирования для идентификации нот с помощью частот основного тона // Российская наука в современном мире: сборник статей XVI международной научно-практической конференции. – 2018. – С. 88–91.

5. *Якимук А.Ю.* Генерация фильтров для одновременной маскировки // Электронные средства и системы управления. – 2018. – № 1–2. – С. 29–31.

6. *Leydon C., Bauer J.J., Larson C.R.* The role of auditory feedback in sustaining vocal vibrato // Acoustical Society of America. – 2003. – V. 114(3). – P. 1575–1581.

7. *Ferguson S., Moere A.V., Cabrera D.* Seeing sound: real-time sound visualisation in visual feedback loops used for training musicians // Ninth International Conference on Information Visualisation (IV'05). London, UK. – 2005. – P. 97–102.

СОЗДАНИЕ API-КЛИЕНТОВ С ПОМОЩЬЮ КОД-ГЕНЕРАТОРОВ

А.Д. Никифоров, С.И. Самохина
Томский государственный университет

Введение

В современном web-программировании отдельное место заняла разработка API. API (application programming interface) – это язык, регламентированный способ общения одной компьютерной программы с другой, для исполнения какой-нибудь общей задачи, когда одна программа выполняет запросы другой [1]. Если говорить о web-программировании, то также к аббревиатуре API необходимо добавить аббревиатуру REST(Representational State Transfer) – это архитектурное решение взаимодействия компонентов информационной системы расположенных в сети интернет. REST определяет ряд правил, которые должны соблюдаться при проектировании программного продукта [2].

1. Обозначение проблемы

В процессе разработке новых интернет приложений программисту очень часто приходится обращаться к чужим API, использовать их интерфейс и внедрять в свою программу. Существует несколько ситуаций, в которых создание API будет целесообразным:

1. Мобильное приложение. Множество мобильных приложений для различных сервисов работают при использовании API этих самых сервисов. Если описано API, сделано мобильное приложение, то клиент со смартфоном будет получать информацию в свое устройство именно через API.

2. Открытое программное обеспечение. Создаётся API, при помощи которого пользователи при желании смогут создать новые клиенты для существующего приложения или новые сервисы на его основе.

3. Максимальное разделение frontend и backend. Например, при использовании frontend -фреймворков.

При интеграции с платформой, предлагающей API REST, разработчик может использовать сгенерированную или написанную человеком специальную библиотеку – SDK, также у программиста есть возможность самому писать HTTP-запросы для взаимодействия с чужим API.

Выбирая из этих двух вариантов, с уверенностью можно сказать, что первый – наиболее удобный, так как необходимо только обратиться к уже существующим методам и классам, чтобы вызвать функцию определенного API. Второй вариант тоже имеет место быть при отсутствии уже сделанных библиотек или же при незнании популярных языков программирования, на которых обычно и делаются API-клиенты [3].

Теперь необходимо разобраться в том, как сделать библиотеку или SDK, которую можно будет подключить в любой проект.

Существует несколько путей решения этой проблемы. Первый способ – вручную. Вся специфика подключения к API и генерация всех классов будет происходить с помощью программиста-разработчика. Этот подход имеет ряд преимуществ, одно из ко-

торых то, что человек может учесть все нюансы уже созданного API и реализовать их в SDK, которое будет распространятся для разработки новых сервисов. Но в таком подходе есть и большой минус – время, затрачиваемое на разработку таких пакетов разработчиком. Другой путь – это генераторы, которые на основе документации к API могут автоматически создать код на необходимом языке программирования.

Таким образом, имея в арсенале код-генератор и правильно написанную документацию API, есть возможность создать SDK, которое можно свободно использовать другим разработчикам. Однако, возникает несколько вопросов. Во-первых, где взять документацию API. Во-вторых, как сформулировать общий регламент спецификации? И в-третьих, как сделать этот регламент доступным для большинства код-генераторов? На помощь в этой ситуации пришла спецификация OpenAPI.

2. Спецификация OpenAPI

Спецификация OpenAPI – это формат описания API или язык определения API. Файл спецификации OpenAPI позволяет описывать API, включая (среди прочего): общую информацию об API, доступные пути (/ресурсы), доступные операции по каждому пути (get / resources), вход / выход для каждой операции.

Спецификацию OpenAPI Specification можно найти в репозитории github OpenAPI Initiative. В этом документе описываются все аспекты спецификации OpenAPI. Использование языка определения API, такого как спецификация OpenAPI, позволяет легко и быстро описать API. Это особенно полезно, когда идёт процесс разработки. Будучи простым текстовым файлом, файл спецификации OpenAPI можно совместно использовать и управлять в любом VCS. После написания файла спецификации его можно использовать в нескольких случаях: исходный материал для документации; спецификация для разработчиков; частичное или полное генерирование кода [4]. Файл OpenAPI может быть написан либо в JSON, либо в YAML. Пример спецификации на рис. 1.

Рис. 1. Пример описания API по спецификации OpenApi в формате json

3. Создание документации OpenAPI

Для того, чтобы создать документацию OpenAPI для WebAPI проекта необходим nuget пакет – Swashbuckle. Swashbuckle знает, как создать спецификацию OpenAPI для WebAPI.

Чтобы добавить Swashbuckle к проекту, необходимо ввести в командную строку следующую команду: dotnet add package Swashbuckle.

Теперь необходимо подключить Swashbuckle к проекту и создать документацию OpenAPI.

Базовая конфигурация для Swashbuckle довольно проста. Во-первых, нужно добавить следующую процедуру в ConfigureServices – метод Startup класса, чтобы зарегистрировать генератор Swagger и определить документ OpenAPI (рис. 2).

```
services.AddSwaggerGen(o => {
    o.SwaggerDoc(
        "v1",
        new Info
        {
           Title = "swaggerPlayground API",
           Version = "v1"
        });
});
```

Рис. 2. Регистрация swagger в ConfigureServices

Во-вторых, добавить промежуточное программное обеспечение, в котором сгенерированный документ Swagger будет сервером. Для этого необходимо добавить следующие строки в Configure метод (рис. 3).

```
app.UseSwagger();
app.UseSwaggerUI(o =>
{
    o.SwaggerEndpoint("/swagger/v1/swagger.json", "API");
});
```

Рис. 3. Добавление swagger в приложение, как middlware сервис

На этом этапе можно запустить приложение и перейти к: http://localhost:5000/swagger/v1/swagger.json. Должен получиться JSON документ, как на рис. 4. Это спецификация OpenAPI – на основе неё, с помощью код генератора, можно создать клиентскую библиотеку.

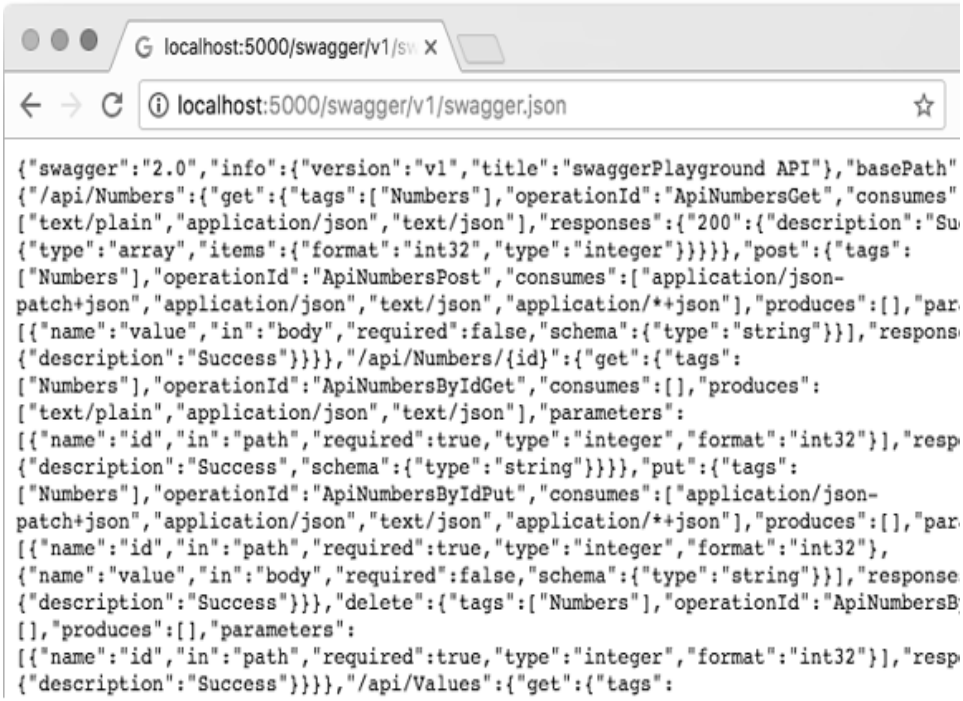


Рис. 4. Спецификация OpenApi

4. Генерация SDK

После того, как создана спецификация OpenAPI, из неё можно генерировать SDK, которое потом будет использоваться разработчиками для создания новых приложений. Вот несколько примеров программ, разработанных специально для автоматического создания SDK: APIMATIC, Restlet Studio, Swagger Codegen, REST United, AutoRest. У каждой из этих программ есть свои особенности предоставления инструментов для генерации кода, но суть у всех одна и та же. Для примера рассмотрим программу AutoRest. AutoRest – это инструмент для создания клиентских библиотек с использованием файла спецификации, который описывает REST API. Используя autorest, есть возможность создавать клиентские библиотеки для следующих языков: C#, NodeJS, Python, Java, Ruby, Go. Теперь необходимо взять OpenAPI файл и создать клиентскую библиотеку. Для этого необходимо предоставить AutoRest входной файл, выходную папку и язык, на котором необходимо создать клиентскую библиотеку. Например, на рис. 5 показано, как из командной строки Windows можно начать генерацию на языке C#.

```
autorest --input-file=swagger.json --output-folder=generated_csharp
--csharp
```

Рис. 5. Команда Autorest для создания кода

Аналогичным образом, по необходимости, можно генерировать клиентскую библиотеку для любого другого языка, заменяя «--csharp», например, на «python», «go» или другим поддерживаемым языком. Если служба защищена, и, например, требуется токен OAuth, необходимо также добавить команду «--add-credentials». Это создаст несколько разные конструкторы для принятия клиентом учетных данных.

После проделанных манипуляций AutoRest создаст файлы с кодом на выбранном языке, в которых будут обращения к API в соответствии с описанной документацией

OpenAPI. Когда есть сгенерированный код, необходимо создать пользовательскую библиотеку, которую можно подключить к стороннему проекту.

В данный момент AutoRest не создаёт файл csproj вместе с сгенерированным кодом, поэтому нужно вручную создать проект в Microsoft Visual Studio типа Class Library с использованием технологии .NET Framework. Далее необходимо добавить сгенерированные классы в созданный проект, после чего собрать его. Теперь в папке bin проекта есть библиотека, которую можно добавить в nuget пакет и выложить его на хостинг nuget.org. После этих действий сгенерированную библиотеку можно добавлять в любой проект.

Заключение

Автоматическая генерация SDK – сложная перспектива; есть много вещей, которые необходимо учитывать при попытке автоматически генерировать код. Отсутствующие данные, такие как определения моделей, коды ошибок и информация аутентификации, могут привести к тому, что автоматически сгенерированный код будет иметь очень низкое качество, а в некоторых случаях код не будет работать. Дизайн API также является основным фактором при попытке автоматически генерировать код. Нет единого, широко принятого, стандартизованного в отрасли стандарта проектирования для API. Не говоря уже о том, что некоторые API разработаны плохо, возвращая некорректно отформатированный JSON и производя неожиданные результаты.

ЛИТЕРАТУРА

1. Tilda API. Автоматическая интеграция проекта на Тильде с собственным сайтом [Электронный ресурс]. – URL: http://help-ru.tilda.ws/api (дата обращения 11.03.2019)

2. *Masse M.* REST API Design Rulebook. – O'Reilly Media, Inc., 1005 Gravenstein Highway North, Sebastopol, CA 95472, 2012. – 94 с.

3. API — Википедия [Электронный ресурс]. – URL: https://ru.wikipedia.org/wiki/API (дата обращения 11.03.2019)

4. About Swagger Specification [Электронный ресурс]. – URL: https://swagger.io/docs/specification/about/ (дата обращения: 11.03.2019).

IV. ТЕОРИЯ МАССОВОГО ОБСЛУЖИВАНИЯ И ТЕЛЕТРАФИКА

МЕТОДЫ РАСЧЕТА ПОКАЗАТЕЛЕЙ ЭФФЕКТИВНОСТИ СМО, ФУНКЦИОНИРУЮЩЕЙ В СЛУЧАЙНОЙ СРЕДЕ[1]

К.И. Аду, Е.В. Маркова, И.А. Гудкова
Российский университет дружбы народов

Введение

Рассмотрим некоторую случайную среду, способную оказывать влияние на показатели качества обслуживания пользователей в рамках одной соты беспроводной сети передачи данных [1,2]. Предположим, что радиус соты равен R, и все пользователи находятся на одинаковом расстоянии от базовой станции (БС), равном d. Рассмотрим систему, в которую поступают запросы пользователей на предоставление услуг, генерирующих потоковый трафик [3]. Данный тип трафика характеризуется фиксированным временем и гарантированной скоростью передачи данных r_0. Под влиянием случайной среды мощность передачи данных может меняться [4], что, в свою очередь, отражается на достижимой скорости передачи данных и максимальном количестве обслуживаемых пользователей [5]. В работе исследуется модель, в которой мощность передачи данных может быть изменена трижды. Для расчета показателей эффективности модели предложен мультипликативный вид стационарного распределения вероятностей состояний модели, а также рекуррентный алгоритм.

1. Построение математической модели

Пусть входящий поток запросов имеет пуассоновское распределение с параметром λ, а среднее время обслуживания одного запроса равно μ^{-1}. Случайная среда может находиться в трёх состояниях $s = 0,1,2$, где $s = 2$ означает, что передача данных осуществляется на максимальной мощности p_2; $s = 0$ – передача данных осуществляется на минимальной мощности p_0.

Под влиянием внешней среды мощность передачи данных снижается интенсивностью α_s, $s = 1,2$, при условии, что число обслуживаемых запросов n, соответствующее состоянию среды s, $s = 1,2$, становится равным максимальному числу запросов, обслуживаемых в состоянии $s - 1$, т.е. $n = N_{s-1}$, $s = 1,2$, и восстанавливается с интенсивностью β_s, $s = 0,1$.

Максимальное число N_s, $s = 0,1,2$ обслуживаемых в системе пользователей может быть рассчитано как отношение достижимой скорости передачи данных r_s, $s = 0,1,2$, определяемой согласно формуле Шеннона и в соответствии с текущим состоянием среды, к гарантированной скорости r_0.

$$N_s = \left\lfloor \frac{\omega \cdot \ln\left(\frac{Gp_s}{d^{\kappa} N_0} + 1\right)}{r_0} \right\rfloor, \; s = 0,1,2,$$

где ω – ширина полосы частот канала, G – коэффициент затухания сигнала, κ – степень затухания сигнала, N_0 – мощность шума.

Граф интенсивностей переходов представлен на рис. 1.

[1] Публикация подготовлена при финансовой поддержке РФФИ в рамках научного проекта №18-37-00231

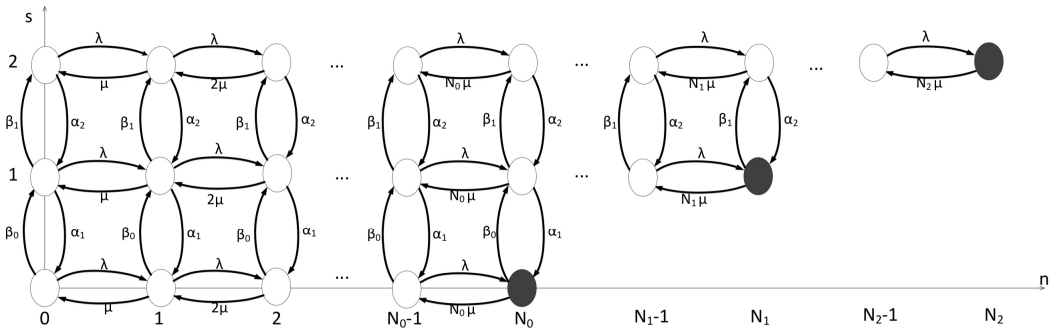


Рис. 1. Диаграмма интенсивностей переходов

Состояние модели описывает двумерный вектор (n,s) над пространством состояний $\mathbf{X} = \{(n,s)\colon 0 \le n \le N_s,\ s = 0,1,2\}$, где n – число обслуживаемых запросов пользователей, s – состояние случайной среды.

Случайный процесс (СП), описывающий поведение системы, характеризуется следующей системой уравнений равновесия (СУР):

$$\begin{cases}
(\lambda+\beta_0)\,p(0,0) = \mu p(1,0) + \alpha_1 p(0,1);\\
(\lambda+\alpha_1+\beta_1)\,p(0,1) = \mu p(1,1) + \alpha_2 p(0,2) + \beta_0 p(0,0);\\
(\lambda+\alpha_2)\,p(0,2) = \mu p(1,2) + \beta_1 p(0,1);\ n = \overline{1, N_0 - 1}\\
(\lambda+n\mu+\beta_0)\,p(n,0) = \lambda p(n-1,0) + (n+1)\mu p(n+1,0) + \alpha_1 p(n,1),\ n = 1,\dots,N_0,\\
(\lambda+n\mu+\alpha_1+\beta_1)\,p(n,1) = \lambda p(n-1,1) + (n+1)\mu p(n+1,1) + \alpha_2 p(n,2) + \beta_0 \cdot p(n,0);\\
n = 1,\dots,N_1,\\
(\lambda+n\mu+\alpha_2)\,p(n,2) = \lambda p(n-1,2) + (n+1)\mu p(n+1,2) + \beta_1 p(n,1);\\
n = N_0+1,\dots,N_1-1,\\
(\lambda+n\mu+\beta_1)\,p(n,1) = \lambda p(n-1,1) + (n+1)\mu p(n+1,1) + \alpha_2 p(n,2);\ n = N_1+1,\dots,N_2-1,\\
(\lambda+n\mu)\,p(n,2) = \lambda p(n-1,2) + (n+1)\mu p(n+1,2);\\
(N_0\mu+\beta_0)\,p(N_0,0) = \lambda p(N_0-1,0) + \alpha_1 p(N_0,1);\\
(N_1\mu+\beta_1)\,p(N_1,1) = \lambda p(N_1-1,1) + \alpha_2 p(N_1,2);\\
N_2\mu p(N_2,2) = \lambda p(N_2-1,2),
\end{cases} \qquad (1)$$

где $p(n,s)$, $(n,s) \in \mathbf{X}$ – стационарное распределение вероятностей состояний системы.

2. Расчет стационарного распределения вероятностей

СП, описывающий поведение системы, является обратимым Марковским СП, в связи с чем выполняется частичный баланс, а стационарное распределение вероятностей $p(n,s)$, $(n,s) \in \mathbf{X}$ состояний системы может быть представлено в мультипликативном виде

$$p(n,s) = \left(\sum_{(n,s)\in \mathbf{X}\setminus\{0,0\}} \prod_{i=1}^{s} \frac{\beta_{i-1}}{\alpha_i} \frac{\rho^n}{n!} + 1 \right)^{-1} \prod_{i=1}^{s} \frac{\beta_{i-1}}{\alpha_i} \frac{\rho^n}{n!},\ (n,s) \in \mathbf{X} \setminus \{0,0\}.$$

Также стационарное распределение вероятностей $p(n,s)$, $(n,s)\in\mathbf{X}$ состояний системы может быть найдено и при помощи рекуррентного алгоритма. Введём ненормированные вероятности $q(n,s)$, $(n,s)\in\mathbf{X}$, которые рассчитываются по формулам

$$q(0,0)=1,\ q(0,1)=x,\ q(0,2)=y,\ q(n,s)=\gamma_{ns}+\delta_{ns}x+\theta_{ns}y. \qquad (2)$$

В (2) коэффициенты γ_{ns}, δ_{ns} и θ_{ns} вычисляются по рекурсивным выражениям

$$\gamma_{00}=1,\ \gamma_{00}=1,\ \delta_{00}=0,\ \theta_{00}=0;\ \gamma_{01}=0,\ \delta_{01}=1,\ \theta_{01}=0;\ \gamma_{02}=0,\ \delta_{02}=0,\ \theta_{02}=1;$$

$$\gamma_{10}=\frac{\lambda+\beta_0}{\mu},\ \delta_{10}=-\frac{\alpha_1}{\mu},\ \theta_{10}=0;$$

$$\gamma_{11}=-\frac{\beta_0}{\mu},\ \delta_{11}=\frac{\lambda+\alpha_1+\beta_1}{\mu},\ \theta_{11}=-\frac{\alpha_2}{\mu};\ \gamma_{12}=0,\ \delta_{12}=-\frac{\beta_1}{\mu},\ \theta_{12}=\frac{\lambda+\alpha_2}{\mu},\ n=\overline{2,N_0-1};$$

$$\gamma_{n0}=\frac{1}{n}\Big[\big(\gamma_{10}+(n-1)\big)\gamma_{n-1,0}+\delta_{10}\gamma_{n-1,1}-\rho\gamma_{n-2,0}\Big],$$

$$\delta_{n0}=\frac{1}{n}\Big[\big(\gamma_{10}+(n-1)\big)\delta_{n-1,0}+\delta_{10}\delta_{n-1,1}-\rho\delta_{n-2,0}\Big],$$

$$\theta_{n0}=\frac{1}{n}\Big[\big(\gamma_{10}+(n-1)\big)\theta_{n-1,0}+\delta_{10}\theta_{n-1,1}-\rho\theta_{n-2,0}\Big];\ n=\overline{2,N_0},$$

$$\gamma_{n1}=\frac{1}{n}\Big[\big(\rho+(n-1)-\delta_{10}-\delta_{12}\big)\gamma_{n-1,1}-\rho\gamma_{n-2,1}+\theta_{11}\gamma_{n-1,2}+\gamma_{11}\gamma_{n-1,0}\Big],$$

$$\delta_{n1}=\frac{1}{n}\Big[\big(\rho+(n-1)-\delta_{10}-\delta_{12}\big)\delta_{n-1,1}-\rho\delta_{n-2,1}+\theta_{11}\delta_{n-1,2}+\gamma_{11}\delta_{n-1,0}\Big],$$

$$\theta_{n1}=\frac{1}{n}\Big[\big(\rho+(n-1)-\delta_{10}-\delta_{12}\big)\theta_{n-1,1}-\rho\theta_{n-2,1}+\theta_{11}\theta_{n-1,2}+\gamma_{11}\theta_{n-1,0}\Big],\ n=\overline{2,N_1};$$

$$\gamma_{n2}=\frac{1}{n}\Big[\big(\rho+(n-1)-\theta_{11}\big)\gamma_{n-1,2}-\rho\gamma_{n-2,2}+\delta_{12}\gamma_{n-1,1}\Big],$$

$$\delta_{n2}=\frac{1}{n}\Big[\big(\rho+(n-1)-\theta_{11}\big)\delta_{n-1,2}-\rho\delta_{n-2,2}+\delta_{12}\delta_{n-1,1}\Big],$$

$$\theta_{n2}=\frac{1}{n}\Big[\big(\rho+(n-1)-\theta_{11}\big)\theta_{n-1,2}-\rho\theta_{n-2,2}+\delta_{12}\theta_{n-1,1}\Big],\ n=\overline{N_0+1,N_1-1};$$

$$\gamma_{n1}=\frac{1}{n}\Big[\big(\rho+(n-1)-\delta_{12}\big)\gamma_{n-1,1}-\rho\gamma_{n-2,1}+\theta_{11}\gamma_{n-1,2}\Big],$$

$$\delta_{n1}=\frac{1}{n}\Big[\big(\rho+(n-1)-\delta_{12}\big)\delta_{n-1,1}-\rho\delta_{n-2,1}+\theta_{11}\delta_{n-1,2}\Big],$$

$$\theta_{n1}=\frac{1}{n}\Big[\big(\rho+(n-1)-\delta_{12}\big)\theta_{n-1,1}-\rho\theta_{n-2,1}+\theta_{11}\theta_{n-1,2}\Big],\ n=\overline{N_1+1,N_2-1};$$

$$\gamma_{n2}=\frac{1}{n}\Big[\big(\rho+(n-1)\big)\gamma_{n-1,2}-\rho\gamma_{n-2,2}+\delta_{12}\gamma_{n-1,1}\Big],$$

$$\delta_{n2}=\frac{1}{n}\Big[\big(\rho+(n-1)\big)\delta_{n-1,2}-\rho\delta_{n-2,2}+\delta_{12}\delta_{n-1,1}\Big],$$

$$\theta_{n2}=\frac{1}{n}\Big[\big(\rho+(n-1)\big)\theta_{n-1,2}-\rho\theta_{n-2,2}+\delta_{12}\theta_{n-1,1}\Big].$$

Используя СУР (1), определим x и y

$$x=(PQ-RL)/(PT-UL),\ y=(RT-UQ)/(PT-UL),$$

где

$$P = \left(\alpha_1 + N_1\mu + \beta_1\right)\theta_{N_1,1} - \alpha_2\theta_{N_2,2} + \lambda\theta_{N_1-1,1},\ Q = \lambda\gamma_{N_2-1,2} - N_2\mu\gamma_{N_2,2},$$
$$R = \alpha_2\gamma_{N_2,2} + \lambda\gamma_{N_1-1,1} - \left(\alpha_1 + N_1\mu + \beta_1\right)\gamma_{N_1,1},\ L = N_2\mu\theta_{N_2,2} - \lambda\theta_{N_2-1,2},$$
$$T = N_2\mu\delta_{N_2,2} - \lambda\delta_{N_2-1,2},\ U = \left(\alpha_1 + N_1\mu + \beta_1\right)\delta_{N_1,1} - \alpha_2\delta_{N_1,2} + \lambda\delta_{N_1-1,1}.$$

Рассчитав ненормированные вероятности $q(n,s)$, $(n,s) \in \mathbf{X}$, определим стационарное распределение вероятностей состояний модели

$$p(n,s) = \frac{q(n,s)}{\sum_{(i,j)\in\mathbf{X}} q(i,j)},\ (n,s) \in \mathbf{X}.$$

3. Вероятностные характеристики модели

Основными вероятностными характеристиками модели являются: вероятость блокировки запроса на передачу данных

$$B = p\left(N_0,0\right) + p\left(N_1,1\right) + p\left(N_2,2\right),$$

среднее число запросов, обслуживаемых в системе

$$\overline{N} = \sum_{s=0}^{2}\overline{N}_s = \sum_{s=0}^{2}\sum_{n=1}^{N_s} np\left(n,s\right),$$

средняя скорость передачи данных в системе

$$\overline{r} = \frac{\sum_{s=0}^{2} r\left(d,p_s\right)\cdot\sum_{n=1}^{N_s} p\left(n,s\right)}{1-\sum_{s=0}^{2} p\left(0,s\right)}.$$

Заключение

В работе рассмотрена модель соты беспроводной сети передачи данных, описанная в виде системы массового обслуживания, работающей в случайной среде с тремя состояниями. Предложен расчёт стационарного распределения вероятностей состояний модели двумя способами – получена аналитическая формула и рекуррентный алгоритм. Представлены формулы для расчёта основных вероятностных характеристик модели – вероятности блокировки, среднего числа обслуживаемых запросов, средней скорости передачи данных.

ЛИТЕРАТУРА

1. *Gudkova I.A., Samouylov K.E., Ostrikova D.Y., Mokrov E.V., Ponomarenko-Timofeev A.A., Andreev S.D., Koucheryavy Y.A.* Service failure and interruption probability analysis for Licensed Shared Access regulatory framework // Proc. of the 7th International Congress on Ultra Modern Telecommunications and Control Systems ICUMT–2015. – 2015. – P. 123131.

2. *Maule M., Moltchanov D., Kustarev P., Komarov M., Andreev S., Koucheryavy Y.* Delivering Fairness and QoS Guarantees for LTE/Wi-Fi Coexistence under LAA Operation // IEEE Access, IET–2018. – Vol. 6. – P. 7359–7373.

3. *Башарин Г.П.* Лекции по математической теории телетрафика: Учеб. пособие. Изд. 3-е, испр. и доп. – М.: РУДН, 2009. – 342 с.

4. *Ahmadian A., Galinina O., Gudkova I., Andreev S., Shorgin S., Samouylov K.* On capturing spatial diversity of joint M2M/H2H dynamic uplink transmissions in 3GPP LTE cellular system // Lecture Notes in Computer Science. – 2014. – Vol. 9247. – P. 407–421.

5. *Markova E., Gudkova I., Ometov A., Dzantiev I., Andreev S., Koucheryavy Ye., Samouylov K.* Flexible Spectrum Management in a Smart City within Licensed Shared Access Framework // IEEE Access. – 2017. – Vol. 5. — P. 22252–22261.

ИССЛЕДОВАНИЕ БЕСКОНЕЧНОЛИНЕЙНЫХ СМО С ГРУППОВЫМ ОБСЛУЖИВАНИЕМ

А.А. Бояркина, С.П. Моисеева
Томский государственный университет

Введение

Несмотря на активное развитие теории массового обслуживания, остаются реальные системы, для описания процесса функционирования которых не находится соответствующих математических моделей. В таком контексте возникают трудности при исследовании многих реальных объектов, которые функционируют по сложному механизму и состояния которых трудно предсказать.

Существующие в настоящее время основные математические модели систем массового обслуживания (СМО) получены для ординарного потока требований. В таком потоке вероятность наступления двух и более событий за бесконечно малое время пренебрежимо мала в сравнении с вероятностью наступления одного события, но данное условие не всегда соответствует реальным физическим явлениям.

В последнее время интенсивно развивается научное направление, связанное с исследованием сетей массового обслуживания с неординарным входящим потоком требований и групповым обслуживанием [1–3], которые эффективно используются в качестве математических моделей сложных систем.

В данной работе демонстрируется применение метода динамического просеивания для решения задачи математического моделирования бесконечнолинейных систем массового обслуживания с групповым обслуживанием заявок.

1. Постановка задачи

Пусть S — открытая система массового обслуживания с неординарным входящим потоком и групповым обслуживанием. Из неординарности входящего потока и группового обслуживания следует, что исследуемая система обладает отличительной особенностью, заключающейся в том, что поступившая на обслуживание группа заявок после обслуживания в том же составе покидает систему. Рассматриваемая система массового обслуживания имеет неограниченное число обслуживающих приборов и является системой без отказов в обслуживании.

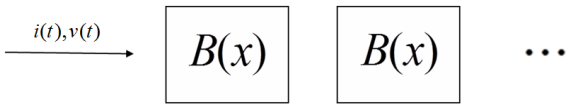


Рис. 1. Бесконечнолинейная ресурсная СМО

На вход системы поступает неординарный поток заявок. Входящая заявка занимает любой свободный прибор, где обслуживается в течении случайного времени $\tau > 0$ с функцией распределения вероятностей $B(x) = P\{\tau < x\}$. Каждая поступающая заявка $i(t)$ в составе группы объёма $v > 0$ занимает ν приборов и освобождает их одновременно.

Предположим, что в начальный момент времени t_0 система пуста. Обозначим $K(t)$ – число занятых групп в системе в момент времени t, $V(t) = \sum_{i=1}^{K(t)} v_i$ – суммарный объём занимаемого ресурса в момент времени $t \geq t_0$. Необходимо найти стационарное распределение вероятностей двумерного случайного процесса $\{K(t), V(t)\}$. Ввиду того, что рассматриваемый двумерный процесс не является марковским, будет использоваться

модифицированный метод динамического просеивания [4], который позволяет учитывать занятые ресурсы заявками, которые находятся в системе к моменту времени *t*.

Зафиксируем некоторый момент времени $T > t_0$. Отметим моменты времени наступления событий входящего потока на оси 0 (см. рис. 2). Далее будем учитывать объёмы занятых ресурсов только тех заявок, которые к моменту времени *T* не закончат свое обслуживание (прошли процедуру просеивания).

Пусть функция $S(t)$, согласно [5], определяет динамическую вероятность просеивания на ось 1 по формуле: $S(t) = 1 - B(T - t)$. Таким образом, $S(t)$ – вероятность того, что заявка, поступившая в систему в момент времени $t < T$, к моменту времени *T* не закончит обслуживание и будет занимать некоторый объём ресурса.

Следовательно, с вероятностью $1 - S(t)$ заявка закончит своё обслуживание и освободит объём ресурса, который занимала при поступлении в систему, поэтому в просеянном потоке она не рассматривается.

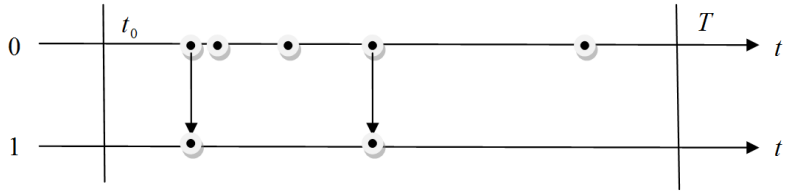


Рис. 2. Просеивание заявок входящего потока

Обозначим $n(t)$ – число групп событий просеянного потока, наступивших на промежутке $[t_0, T)$, $W(t)$ – суммарный объём занимаемого ресурса. Отметим, что построенный процесс не является марковским в случае непуассоновского входящего потока.

Между процессами $n(t)$ и $K(t)$ в [5] установлена связь, а именно – утверждается, что в момент времени $t = T$ законы распределения вероятностей данных случайных величин совпадают: $P\{K(T) = m\} = P\{n(T) = m\}$ для любых *m*. Из данной формулы очевидно следует:

$$P\{K(T) = m, V(T) < x\} = P\{n(t) = m, W(t) < x\}. \quad (1)$$

Отметим, что в теории равенство (1) называется основной формулой метода динамического просеивания для ресурсных систем массового обслуживания с неограниченным числом приборов.

В данной работе рассматривается применение метода для исследования бесконечнолинейных СМО с групповым обслуживанием.

2. Исследование бесконечнолинейных СМО с групповым обслуживанием

Пусть на вход системы с групповым обслуживанием и дискретным временем, процесс функционирования которой изображен на рис. 1, поступает неординарный входящий потоком требований. Построим двумерный случайный процесс $\{n(t), W(t)\}$. Введём обозначение распределение вероятностей: $P\{n(t) = n, W(t) < w\} = P(n, w, t)$ для $n, w > 0$. По формуле полной вероятности запишем равенство

$$P(n, w, t + \Delta t) = P(n, w, t)(1 - \lambda\Delta t) + P(n, w, t)\lambda\Delta t(1 - S(t)) + \sum_{k=0}^{w} P(n-1, w-k, t)\, p_k \lambda\Delta t S(t),$$

откуда получаем систему дифференциальных уравнений Колмогорова

$$\frac{\partial P(n, w, t + \Delta t)}{\partial t} = \lambda S(t)\left(\sum_{k=0}^{w} P(n-1, w-k, t) p_k - P(n, w, t)\right) \quad (2)$$

с начальным условием $P(n,w,t_0)=\begin{cases}1, & n=0,\\ 0, & \text{иначе.}\end{cases}$

Введём производящую функцию $F(x,y,t)=\sum_{n=0}^{\infty}\sum_{w=0}^{\infty}x^n y^w P(n,w,t)$. С учётом производящей функции уравнение (2) перепишется в виде

$$\frac{\partial F(x,y,t)}{\partial t}=\lambda S(t)\left(\sum_{n=0}^{\infty}\sum_{w=0}^{\infty}\sum_{k=0}^{w}x^n y^w P(n-1,w-k,t)\,p_k-F(x,y,t)\right)=$$

$$=\lambda S(t)\left(x\sum_{n=0}^{\infty}\sum_{w=0}^{\infty}\sum_{k=0}^{w}x^{n-1}y^{w-k}y^k P(n-1,w-k,t)\,p_k-F(x,y,t)\right)=$$

$$=\lambda S(t)\left(x\sum_{k=0}^{\infty}y^k p_k\sum_{n=0}^{\infty}\sum_{w=k}^{\infty}x^{n-1}y^{w-k}P(n-1,w-k,t)-F(x,y,t)\right)=\lambda S(t)F(x,y,t)\left(xf(y)-1\right).$$

Таким образом, уравнение (2) окончательно примет вид:

$$\frac{\partial F(x,y,t)}{\partial t}=\lambda S(t)F(x,y,t)\left(xf(y)-1\right) \tag{3}$$

с начальным условием $F(x,y,t_0)=1$, при этом решение (3) примет вид

$$F(x,y,t)=\exp\left\{\frac{\lambda}{\mu}e^{\mu(t-T)}\left(xf(y)-1\right)\right\}.$$

При $t=T$ для производящей функции двумерного процесса числа занятых групп и суммарного объёма занятого ресурса в стационарном режиме получим $F(x,y,t)=\exp\left\{\frac{\lambda}{\mu}(xf(y)-1)\right\}$, где $f(y)=\sum_{k=0}^{\infty}y^k p_k$ является производящей функцией числа заявок в группе.

Производящая функция процесса суммарного объёма занятого ресурса в стационарном режиме имеет вид $F(x,y,t)\big|_{x=1}=\exp\left\{\frac{\lambda}{\mu}(f(y)-1)\right\}$.

Таким же образом, полагая $y=1$, получим производящую функцию распределения вероятностей числа занятых групп в системе в стационарном режиме $F(x,y,t)\big|_{y=1}=\exp\left\{\frac{\lambda}{\mu}(x-1)\right\}$.

Найденные характеристики суммарного объёма занятого ресурса и распределения вероятностей числа занятых групп в стационарном режиме позволяют исследовать функционирование системы.

Найдём математическое ожидание суммарного объема занятого ресурса в системе: $M\{W\}=F'(y)=e^{\frac{\lambda}{\mu}(f(y)-1)}\frac{\lambda}{\mu}f'(y)\big|_{y=1}=\frac{1}{\mu}\overline{k}$, где $\overline{k}$ – среднее число заявок в группе.

Анализируя полученные результаты, можно утверждать об их достоверности.

Заключение

В данной работе была исследована математическая модель бесконечнолинейных систем с групповым обслуживанием и неординарным потоком требований в дискретном времени. Посредством модифицированного метода динамического просеивания были получены основные характеристики системы.

Исследуемую систему можно рассматривать в рамках многих сложных реальных систем, функционирующие сложным образом и требующие оценки будущих состояний.

ЛИТЕРАТУРА

1. *Митрофанов Ю.И., Рогачко Е.С., Станкевич Е.П.* Анализ сетей массового обслуживания с неординарными потоками: Учебное пособие. — Саратов: Издат. центр "Наука", 2011. – 109 с.

2. *Митрофанов Ю.И., Рогачко Е.С., Станкевич Е.П.* Динамическое распределение нагрузки в замкнутых сетях массового обслуживания с групповыми переходами требований // Изв. Сарат. ун-та. Нов. сер. Серия Математика. Механика. Информатика. – 2012. – Т. 12. – Вып. 1. – С. 22–28.

3. *Тананко И.Е., Фокина Н.П.* Анализ замкнутых ненадежных сетей массового обслуживания с групповыми переходами требований // Изв. Сарат. ун-та. Нов. сер. Серия Математика. Механика. Информатика. – 2013. – Т. 13. – Вып. 2. – Ч. 1. – С. 111–117.

4. *Назаров А.А., Моисеева С.П.* Методы асимптотического анализа в теории массового обслуживания – Томск: Изд-во НТЛ, 2006. – 112 с.

5. *Назаров А.А., Терпугов А.Ф.* Теория массового обслуживания. – Томск: Изд-во НТЛ, 2005. – 228 с.

ВЕРОЯТНОСТНАЯ МОДЕЛЬ НАРЕЗКИ РЕСУРСОВ БЕСПРОВОДНОЙ СЕТИ МЕЖДУ ДВУМЯ ВИРТУАЛЬНЫМИ ОПЕРАТОРАМИ

П.И. Дымова, В.Н. Савич, Н.А. Поляков, А.С. Власкина

Российский университет дружбы народов

Введение

Сообщения, передаваемые по сетям, характеризуются двумя принципиально различными видами трафика: потоковым и эластичным. В конце 20-го века, когда наиболее актуальны были телефонные сети, большое внимание уделялось изучению потокового трафика, чувствительного к задержкам [1]. В начале 21-го века, с ростом потребности в передаче не только речи, но и других медиафайлов, внимание сдвинулось в сторону эластичного трафика, допускающего большие задержки, но чувствительного к потерям данных.

В работе исследована система массового обслуживания (СМО) с двумя очередями с нетерпеливым эластичным трафиком и минимальными требованиями к скорости обслуживания заявки. Трафик способен адаптироваться к изменениям доступной ему пропускной способности, но чувствителен к потерям. Система предусматривает равенство вероятностей попадания заявок обоих типов на прибор. Построена модель такой системы обслуживания, рассмотрены ее возможные состояния, получена матрица интенсивностей переходов и формулы для вычисления показателей эффективности системы.

1. Постановка задачи

Рассматриваемая система имеет заявки двух типов, n_1 и n_2. Интенсивность входящего потока заявок первого типа обозначена через λ_1, а второго через λ_2. Для каждого типа заявок предусмотрена своя очередь, m_1 и m_2 соответственно [2]. На обслуживание каждой заявки выделяется одинаковое количество ресурсов. Если заявка обслужилась, то вероятность попадания на прибор заявки первого типа равна вероятности попадания на прибор заявки второго типа, т.е. $1/2$ (рис. 1).

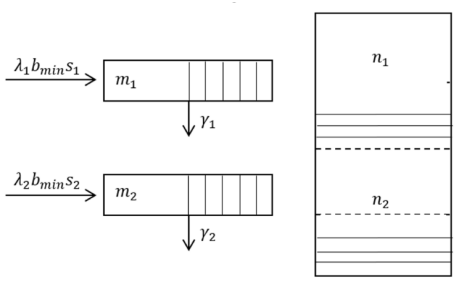


Рис. 1. Система массового обслуживания с динамической сегментацией ресурсов по активности пользователей

Здесь $b_{\min}$ – минимальная гарантированная скорость для обслуживания заявки, а s_1, s_2 – объем заявки соответствующего типа. Интенсивность ухода заявки по причине нетерпеливости обозначена через γ_1 из очереди первого типа и γ_2 из очереди второго. Пунктиром на рисунке обозначено динамическое изменение количества заявок двух типов на обслуживании [3]. Пространство состояний для данной системы выглядит следующим образом:

$$X = \begin{Bmatrix} (n_1, n_2, m_1, m_2) : n_1 \geq 0, n_2 \geq 0, m_1 \geq 0, m_2 \geq 0; \\ (n_1, n_2, 0, 0) : n_1 + n_2 \leq N; \\ (n_1, n_2, m_1, m_2) : n_1 + n_2 = N, m_1 \leq M, m_2 \leq M \end{Bmatrix},$$

где N –максимальное число заявок, находящихся на обслуживании одновременно, M – максимальная длина очереди.

Пространство состояний системы можно так же представить графически (рис. 2).

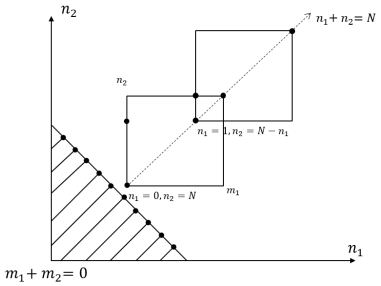


Рис. 2. Графическое представление пространства состояний системы с двумя операторами

2. Матрица интенсивностей переходов

Для вывода формул для расчёта показателей эффективности системы, необходимо составить матрицу интенсивности переходов. Матрица интенсивностей переходов составляется на основе диаграммы интенсивностей переходов, которая может быть представлена либо в виде графа с состояниями – вершинами и интенсивностями – ребрами, либо в виде таблицы.

Рассматриваемая система имеет 15 состояний. В табл. 1 представлены состояния системы, записанные в общем виде, где n_1 и n_2 – количество заявок в системе, а m_1 и m_2 – количество заявок в очереди.

Таблица 1

Состояния и интенсивности переходов системы с динамической сегментацией ресурсов

Состояние	Интенсивность	Комментарии
$n_1+1, n_2, 0, 0$	λ_1	Очереди нет. Заявка 1-го оператора поступает на 1-ый прибор
$n_1, n_2+1, 0, 0$	λ_2	Очереди нет. Заявка 1-го оператора поступает на 1-ый прибор

Состояние	Интенсивность	Комментарии
$n_1-1, n_2, 0, 0$	c/s_1	Очереди нет. Заявка 1-го оператора обслуживается
$n_1, n_2-1, 0, 0$	c/s_2	Очереди нет. Заявка 2-го оператора обслуживается
n_1, n_2, m_1+1, m_2	λ_1	Все приборы заняты. Заявка 1-го оператора пришла в очередь
n_1, n_2, m_1, m_2+1	λ_2	Все приборы заняты. Заявка 2-го оператора пришла в очередь
n_1, n_2, m_1-1, m_2	$c/2s_1 + m_1\gamma_1$	На приборе заявки разных операторов. В очереди каждого есть заявки. 1) Заявка обслужилась. На ее место встает заявка из 1-ой очереди 2) Нетерпеливая заявка
n_1, n_2, m_1, m_2-1	$c/2s_2 + m_2\gamma_2$	На приборе заявки разных операторов. В очереди каждого есть заявки. 1) Заявка обслужилась. На ее место встает заявка из 2-ой очереди 2) Нетерпеливая заявка
$n_1, n_2, m_1-1, 0$	$c/s_1 + m_1\gamma_1$	На приборе заявки 2-х типов. Очередь 2-го оператора пуста. 1) Обслужится заявка из очереди 1-го оператора 2) Нетерпеливая заявка
$n_1, n_2, 0, m_2-1$	$c/s_2 + m_2\gamma_2$	На приборе заявки 2-х типов. Очередь 1-го оператора пуста. 1) Обслужится заявка из очереди 2-го оператора 2) Нетерпеливая заявка
$n_1-1, n_2+1, 0, m_2-1$	c/s_1	На приборе заявки 2-х типов. Очередь 1-го оператора пуста. Обслужилась заявка 1-го оператора, на прибор поступила заявка 2-го оператора, так как его очередь не пуста
$n_1+1, n_2-1, m_1-1, 0$	c/s_2	На приборе заявки 2-х типов. Очередь 2-го оператора пуста. Обслужилась заявка 2-го оператора, на прибор поступила заявка 1-го оператора, так как его очередь не пуста
n_1-1, n_2+1, m_1, m_2-1	$c/2s_1$	На приборе заявки 2-х типов. В очереди каждого есть заявки. Заявка 1-го типа обслуживается и на ее место приходит заявка 2-го типа из очереди, так как вероятность попадания из каждой очереди – 1/2
n_1+1, n_2-1, m_1-1, m_2	$c/2s_2$	На приборе заявки 2-х типов. В очереди каждого есть заявки. Заявка 2-го типа обслуживается и на ее место приходит заявка 1-го типа из очереди, так как вероятность попадания из каждой очереди – 1/2

3. Численный анализ показателей эффективности обслуживания трафика операторов

На основе матрицы интенсивностей переходов были выведены формулы, необходимые для численного анализа эффективности рассматриваемой системы:

– среднее число заявок на приборе;
– среднее число заявок в очереди;
– среднее число заявок в системе;
– среднее время обслуживания на приборе;
– среднее время пребывания заявки в очереди;
– среднее время пребывания заявки в системе;
– средняя скорость обслуживания заявки;
– доля занятости ресурса.

Формулы для вычисления характеристик представлены в табл. 2.

Таблица 2

Показатели эффективности модели динамической сегментации ресурсов

Показатель	Заявки 1-го типа	Заявки 2-го типа	Суммарное количество заявок
Среднее число заявок на приборе	$\sum_{\vec{x}\in x_1} n_1 p(\vec{x})$	$\sum_{\vec{x}\in x_2} n_2 p(\vec{x})$	$\sum_{\vec{x}\in x_1} n_1 p(\vec{x}) + \sum_{\vec{x}\in x_2} n_2 p(\vec{x})$
Среднее число заявок в очереди	$\sum_{\vec{x}\in x_1} m_1 p(\vec{x})$	$\sum_{\vec{x}\in x_2} m_2 p(\vec{x})$	$\sum_{\vec{x}\in x_1} m_1 p(\vec{x}) + \sum_{\vec{x}\in x_2} m_2 p(\vec{x})$
Среднее число заявок в системе	$\sum_{\vec{x}\in x_1} n_1 p(\vec{x}) + \sum_{\vec{x}\in x_1} m_1 p(\vec{x})$	$\sum_{\vec{x}\in x_2} n_2 p(\vec{x}) + \sum_{\vec{x}\in x_2} m_2 p(\vec{x})$	$\sum_{\vec{x}\in x_1} n_1 p(\vec{x}) + \sum_{\vec{x}\in x_1} m_1 p(\vec{x}) + \sum_{\vec{x}\in x_2} n_2 p(\vec{x}) + \sum_{\vec{x}\in x_2} m_2 p(\vec{x})$
Среднее время обслуживания заявок на приборе	$\frac{\sum_{\vec{x}\in x_1} n_1 p(\vec{x})}{(\lambda_1+\lambda_2)/2}$	$\frac{\sum_{\vec{x}\in x_2} n_2 p(\vec{x})}{(\lambda_1+\lambda_2)/2}$	–
Среднее время пребывания заявок в очереди	$\frac{\sum_{\vec{x}\in x_1} m_1 p(\vec{x})}{\lambda_1}$	$\frac{\sum_{\vec{x}\in x_2} m_2 p(\vec{x})}{\lambda_2}$	–
Среднее время пребывания заявок в системе	$\frac{\sum_{\vec{x}\in x_1} n_1 p(\vec{x})}{(\lambda_1+\lambda_2)/2} + \frac{\sum_{\vec{x}\in x_1} m_1 p(\vec{x})}{\lambda_1}$	$\frac{\sum_{\vec{x}\in x_2} n_2 p(\vec{x})}{(\lambda_1+\lambda_2)/2} + \frac{\sum_{\vec{x}\in x_2} m_2 p(\vec{x})}{\lambda_2}$	–
Средняя скорость обслуживания	$\frac{1-\rho_1^N/N!}{\sum_{n_1=0}^{N}\rho_1^N/n_1!}$, где $\rho_1=\frac{\lambda_1}{c/s_1}$	$\frac{1-\rho_2^N/N!}{\sum_{n_2=0}^{N}\rho_2^N/n_2!}$, где $\rho_2=\frac{\lambda_2}{c/s_2}$	–
Доля занятости ресурса	$\frac{C}{n_1+n_2}*n_1$	$\frac{C}{n_1+n_2}*n_2$	–

Перейдем непосредственно к численному анализу. Для моделирования используем некоторые исходные данные, представленные в табл. 3.

Таблица 3

Исходные данные для численного анализа характеристик системы с двумя виртуальными операторами

Параметр	Характеристика	Значение
$b_{\min}$	Минимальная гарантированная скорость	6 [кб/с]
C	Объем ресурса	10 [Мб/с]
λ_1	Интенсивность входящего потока 1-го типа	1 заявка в минуту
λ_2	Интенсивность входящего потока 2-го типа	2 заявки в минуту
γ_1	Интенсивность ухода из системы по причине нетерпеливости заявки 1-го типа	0.001
γ_2	Интенсивность ухода из системы по причине нетерпеливости заявки 2-го типа	0.001
s_1	Объем заявки 1-го типа	1
s_2	Объем заявки 2-го типа	1

С помощью программного кода, написанного на языке Java, были получены числовые значения необходимых характеристик для каждого типа заявки. Они представлены в табл. 4.

Таблица 4

Показатели эффективности модели динамической сегментации ресурсов

Показатель	Заявки 1-го типа	Заявки 2-го типа
Среднее число заявок на приборе	0.0962001	0.2259799
Среднее число заявок в очереди	0.0163979	0.0484574
Среднее число заявок в системе	0.0722421	0.1439886
Среднее время обслуживания заявок на приборе	0.0532014	0.1183483
Среднее время пребывания заявок в очереди	0.0163979	0.0242287
Среднее время пребывания заявок в системе	0.0695994	0.1425771
Средняя скорость обслуживания	0.909090	0.833333
Доля ресурса на приборе	90	90

Заключение

В данной работе рассмотрена система массового обслуживания с двумя очередями и нетерпеливым эластичным трафиком с заданной минимальной скоростью обслуживания. Для рассматриваемой модели была построена матрица интенсивностей переходов, и получены формулы для проведения численного анализа основных характеристик системы.

ЛИТЕРАТУРА

1. *Башарин Г.П.* Лекции по математической теории телетрафика. – М.: Российский университет дружбы народов, 2007. 268 с.

2. *Вишневский В.М, Семенова О.В.* Системы поллинга теория и применение в широкополосных беспроводных сетях. – М.: Техносфера, 2007. 308 с.

3. *Samouylov K.E., Gudkova I.A.* Recursive computation for a multi-rate model with elastic traffic and minimum rate guarantees (2010) // 2010 International Congress on Ultra Modern Telecommunications and Control Systems and Workshops. ICUMT, art. no. 5676509, 2010. – P. 1065–1072.

ПРИМЕНЕНИЕ ОТРИЦАТЕЛЬНОГО БИНОМИАЛЬНОГО РАСПРЕДЕЛЕНИЯ ДЛЯ АППРОКСИМАЦИИ СТАЦИОНАРНОГО РАСПРЕДЕЛЕНИЯ ЧИСЛА ЗАЯВОК В СИСТЕМЕ HIGI|GI|∞

М.П. Зимин, А.Н. Моисеев

Томский государственный университет

Введение

В работе рассматривается система массового обслуживания HIGI|GI|∞ с неограниченным числом приборов, входящим рекуррентным потоком заявок и рекуррентным обслуживанием. В [1] была получена гауссовская аппроксимация для стационарного распределения числа заявок в данной системе, полученная в условии высокой интенсивности входящего потока, установлена область применимости указанной аппроксимации. В [2] на основе асимптотического анализа третьего порядка получена более точная аппроксимация для указанного распределения с более широкой областью применимости. При этом применение последней значительно усложнилось по причине необходимости проводить обратное преобразование Фурье.

В ходе поиска аппроксимации с еще более широкой областью применимости, но более «легким» применением, было замечено, что в [3–5] в качестве аппроксимирующего распределения получено гамма-распределение, дискретным аналогом которого является отрицательное биномиальное распределение [6] (далее – ОБР) с действительными положительными значениями параметра, задающего число ожидаемых успехов. Исходя из этого, было принято решение использовать ОБР для аппроксимации стационарного распределения числа заявок в различных типах систем массового обслуживания, где получение аналитических неасимптотических результатов не представляется

возможным. В частности, в настоящей работе проводится анализ применимости ОБР в качестве аппроксимации стационарного распределения числа заявок в системе HIGI|GI|∞.

1. Постановка задачи

Рассмотрим систему массового обслуживания с неограниченным числом приборов, на вход которой поступает высокоинтенсивный рекуррентный (HIGI) поток заявок [7]. Длины τ интервалов между последовательным поступлением заявок из этого потока независимы и одинаково распределены. Функция распределения значений τ описывается следующим образом. Представим τ в виде $\tau = \xi / N$, где ξ – некоторая неотрицательная случайная величина с функцией распределения $A(z)$, а параметр $N > 0$ имеет смысл большой величины (в теоретических исследованиях предполагается, что $N \to \infty$). Тогда $P\{\tau < x\} = A(Nx)$. Время обслуживания каждой заявки является случайной величиной с заданной функцией распределения $B(t)$.

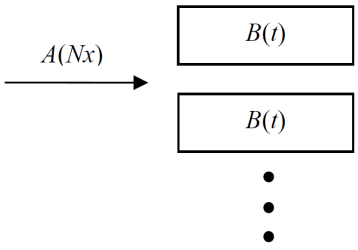


Рис. 1. Схематичное изображение системы HIGI|GI|∞

Ставится задача определения распределения числа заявок в системе в стационарном режиме ее функционирования.

2. Результаты асимптотического анализа

В [1,2] с применением асимптотического анализа установлено, что в условиях высокой интенсивности входящего потока стационарное распределение числа заявок в системе является гауссовским, получены параметры этого гауссовского распределения. Характеристическую функцию $h_2(u)$ стационарного распределения числа заявок в системе можно записать следующим образом:

$$h_2(u) = \exp\left\{ juN\lambda b + \frac{(ju)^2}{2}\left[N\lambda b + N\kappa\beta\right]\right\}, \qquad (1)$$

где $\lambda = \dfrac{1}{\mathrm{M}\{\xi\}} = \dfrac{1}{\int_0^\infty [1 - A(z)]\, dz}$, $b = \int_0^\infty [1 - B(\tau)] d\tau$ – среднее время обслуживания, $\beta = \int_0^\infty [1 - B(\tau)]^2 d\tau$, величина κ вычисляется как $\kappa = \lambda^3 (\sigma^2 - a^2)$, где a и σ^2 – математическое ожидание и дисперсия случайной величины с функцией распределения $A(x)$.

Указанные результаты получены в условии высокой интенсивности входящего потока, т.е. при $N \to \infty$. Таким образом, функцию (1) можно рассматривать в качестве аппроксимации для характеристической функции стационарного распределения числа заявок в системе HIGI|GI|∞ при достаточно больших значениях параметра *N*, т.е. в указанных условиях в качестве аппроксимации стационарного распределения числа заявок

можно использовать нормальное распределение с математическим ожиданием $N\lambda b$ и дисперсией $\left[N\lambda b + N\kappa b\right]$ (точнее – его дискретный неотрицательный аналог).

В [2] также было проведено исследование данной системы с применением асимптотического анализа третьего порядка. В результате исследования была получена функция $h_3(u)$ вида

$$h_3(u) \approx \exp\left\{ juN\lambda b + \frac{(ju)^2}{2}\left[N\lambda b + N\kappa\beta\right] + \frac{(ju)^3}{6}\left[N\lambda b + 3N\kappa\beta + N\mu\gamma\right]\right\},$$

которая, как было установлено, даёт более точную, по сравнению с (1), аппроксимацию для характеристической функции числа заявок в системе в стационарном режиме. Здесь $\mu = 6\lambda + 3\lambda^2 a_2\kappa - \lambda^4 a_3$, a_2 и a_3 – начальные моменты соответственно второго и третьего порядков случайной величины с функцией распределения $A(x)$, $\gamma = \int_0^\infty \left(1 - B(\tau)\right)^3 d\tau$, а остальные параметры определены выше.

3. Аппроксимация отрицательным биномиальным распределением

Отрицательное биномиальное распределение – это распределение дискретной случайной величины ξ, равной количеству неудач, произошедших в последовательности испытаний Бернулли с вероятностью успеха p до наступления r-го успеха. В оригинале распределения выглядит следующим образом:

$$\mathrm{P}\{\xi = k\} = \frac{(k+r-1)!}{k!(r-1)!}\, p^r (1-p)^k .$$

Для случаев действительных значений параметра r его можно переписать с использованием гамма-функции Эйлера в следующем виде:

$$\mathrm{P}\{\xi = k\} = \frac{\Gamma(r+k)}{k!\Gamma(r)}\, p^r (1-p)^k .$$

Для использования ОБР в качестве аппроксимации для целей настоящей работы нам необходимо определить его параметры r и p таким образом, чтобы полученное распределение было максимально близко к настоящему стационарному распределению числа заявок в системе. Для получения таких значений параметров очень удобно воспользоваться формулами для математического ожидания и дисперсии:

$$\mathrm{M}\{\xi\} = \frac{r(1-p)}{p},\ \mathrm{D}\{\xi\} = \frac{r(1-p)}{p^2},$$

приравняв их к соответствующим значениям матожидания и дисперсии числа заявок в системе HIGI|GI|∞ в стационарном режиме. Таким образом, полученная аппроксимация будет гарантировать совпадение с реальным распределением как минимум в двух первых моментах. Однако, к сожалению, указанные параметры получить аналитически не представляется возможным (по крайней мере, дисперсию), поэтому в качестве таких «опорных» значений будем использовать матожидание и дисперсию приведенной в предыдущем разделе гауссовской аппроксимации (известно, что математическое ожидание для реального распределения совпадает с указанным). Таким образом, можно записать следующую систему уравнений относительно неизвестных параметров ОБР r и p:

$$\begin{cases} N\lambda b = \dfrac{r(1-p)}{p}, \\ N\lambda b + N\kappa\beta = \dfrac{r(1-p)}{p^2}, \end{cases}$$

где параметры N, λ, b, κ, β описаны в предыдущем разделе. Решая эту систему, получаем следующие выражения для вычисления параметров r и p:

$$r = \frac{\lambda b}{\lambda b + \kappa\beta}, \quad p = \frac{N(\lambda b)^2}{\kappa\beta}. \tag{2}$$

4. Анализ области применимости ОБР-аппроксимации

Выполним исследование области применимости полученной аппроксимации в виде отрицательного биномиального распределения с параметрами (2). Выражения (2) получены на основе гауссовской аппроксимации, которая, в свою очередь, была получена в асимптотических условиях высокой интенсивности входящего потока, т.е. при $N \to \infty$. Таким образом, можно предполагать, что при увеличении параметра N в выражениях (2) ОБР-аппроксимация будет становиться точнее.

Ввиду отсутствия допредельных аналитических выражений для стационарного распределения числа заявок в системе HIGI|GI|∞ анализ точности ОБР-аппроксимации будем выполнять путем её сравнения с эмпирическим рядом распределения, построенным на основе результатов имитационного моделирования соответствующей системы массового обслуживания. Сравнение будем производить на основе вычисления расстояния Колмогорова между функциями распределений:

$$D_q = \sup_x \left| F_q(x) - F(x) \right|.$$

Здесь $F(x)$ – эмпирическая функция распределения, $F_q(x)$ – функция распределения для аппроксимации. Допустимой границей применимости аппроксимации будем считать значения $D_q \le 0.05$. Также выполним аналогичный анализ области применимости и для описанных аппроксимаций второго (гауссовской) и третьего порядка [1,2].

Пусть система HIGI|GI|∞ задана следующими параметрами. Интервалы между наступлением событий во входящем потоке являются гамма-величинами с параметрами распределения $k = 1/3$, $\theta = 3/N$. Длительности обслуживания на каждом приборе также имеют гамма-распределение с параметрами $k = 0.5$, $\theta = 2$ для характеристической функция гамма-распределения, записанной в виде $(1 - \theta it)^{-k}$.

Моделирование выполнено при различных значениях параметра высокой интенсивности N, результаты представлены на рис. 2, а сравнение полученных значений расстояний Колмогорова для разных аппроксимаций приведены в табл. 1 и на рис. 3. В таблице полужирным шрифтом выделены допустимые значения для расстояния Колмогорова согласно выбранному критерию ($D_q \le 0.05$).

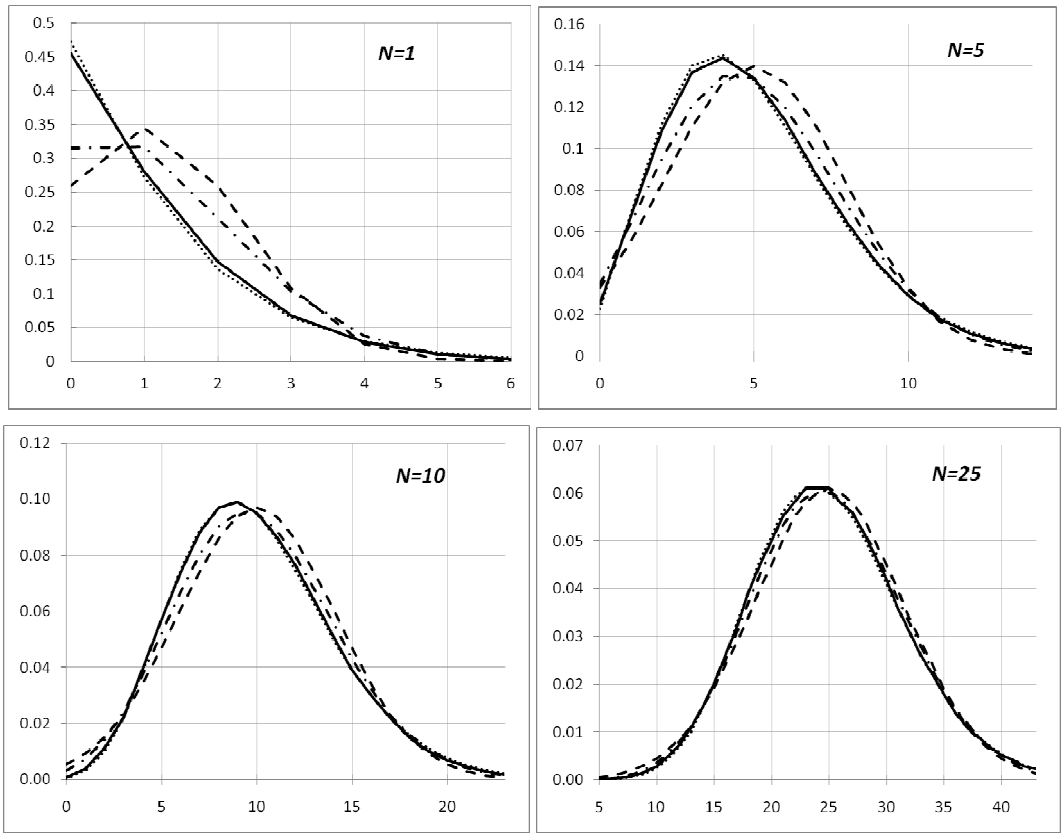


Рис. 2. Сравнение эмпирического ряда распределения (сплошная линия), ОБР-аппроксимации (точечная линия) и асимптотических распределений второго (пунктирная линия) и третьего (пунктирная линия с точкой) порядков для стационарного распределения числа заявок системе при различных значениях интенсивности входящего потока N

Таблица 1

Расстояние Колмогорова для различных аппроксимаций

Интенсивность входящего потока N	1	5	10	25
Гауссовская аппроксимация (второго порядка)	0.23839	0.05532	**0.03315**	**0.01487**
Аппроксимация третьего порядка	0.1627	**0.02986**	**0.017**	**0.00643**
ОБР-аппроксимация	**0.02183**	**0.00682**	**0.00407**	**0.00344**

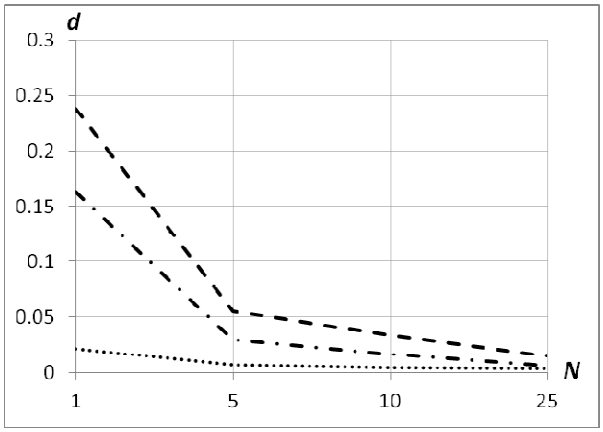


Рис. 2. Значения расстояния Колмогорова между эмпирическим распределением и аппроксимациями: ОБР (точечная линия), второго (пунктирная линия) и третьего (пунктирная линия с точкой) порядков

Как видим, аппроксимация, полученная с помощью ОБР, является более точной, чем аппроксимации второго (гауссовская) и третьего порядков. Получается, что ОБР-аппроксимация применима и при малых значениях параметра N, т.е может быть использована для систем типа HIGI|GI|∞ при небольшой интенсивности входящего потока ($N \geq 1$). Анализ применимости ОБР-аппроксимации при меньших значениях интенсивности входящего потока, а также анализ зависимости области применимости данной аппроксимации относительно дисперсий (длин интервалов между заявками входящего потока и длительности обслуживания) представляет собой тему для отдельного исследования.

Заключение

В работе выполнено исследование применимости отрицательного биномиального распределения в качестве аппроксимации для стационарного распределения числа заявок в системе массового обслуживания HIGI|GI|∞, которое не представляется возможным получить аналитически без использования асимптотических методов. В работе предлагается метод вычисления параметров указанной ОБР-аппроксимации, основанный на гауссовской аппроксимации, полученной методом асимптотического анализа. Проведён сравнительный анализ областей применимости различных аппроксимаций и установлено, что ОБР является наилучшей среди них. Более того, она оказалась применимой даже при небольших значениях интенсивности входящего потока, что делает её чрезвычайно полезной для практического применения.

ЛИТЕРАТУРА

1. *Моисеев А.Н., Назаров А.А.* Исследование системы массового обслуживания HIGI/GI/∞ // Вестник Томского госуниверситета. Управление, вычислительная техника и информатика. – 2013 – № 2 (23). – С. 75–83.

2. *Моисеев А.Н.* Исследование математических моделей систем и сетей массового обслуживания с высокоинтенсивными непуассоновскими входящими потоками: Дисс. … докт. физ.-мат. наук: Томск: ТГУ. 2016. 333 с.

3. *Fedorova E.* Quasi-geometric and gamma approximation for retrial queueing systems // Communications in Computer and Information Science. – 2014. – Vol. 487. – P. 123–136.

4. *Fedorova E.A., Nazarov A.A., Paul S.V.* Discrete gamma probability distribution approximation in retrial queues // Аналитические и вычислительные методы в теории вероятностей и ее приложениях (АВМТВ-2017): материалы Междунар. науч. конф. Россия, Москва, 23–27 окт. 2017 г. – М.: Изд-во РУДН, 2017. – С. 560–565.

5. *Fedorova E., Voytikov K.* Retrial Queue M/G/1 with Impatient Calls Under Heavy Load Condition // Communications in Computer and Information Science. – 2017. – Vol. 800. – P. 347–357.

6. *Вадзинский Р.Н.* Справочник по вероятностным распределениям. – СПб.: Наука, 2001. – 295 с.

7. *Moiseev A., Nazarov A.* Investigation of high intensive general flow // Proc. IV International Conference “Problems of Cybernetics and Informatics” (PCI’2012), September 12−14, 2012, Baku, Azerbaijan. – Baku: ANAS, 2012. – P. 161–163.

АСИМПТОТИЧЕСКИЙ АНАЛИЗ RQ-СИСТЕМЫ MMPP|M|1 С N ТИПАМИ ВЫЗЫВАЕМЫХ ЗАЯВОК В УСЛОВИИ ПРЕДЕЛЬНО РЕДКИХ ИЗМЕНЕНИЙ СОСТОЯНИЙ ВХОДЯЩЕГО ПОТОКА[*]

О.Д. Лизюра, А.А. Назаров, С.В. Пауль

Томский государственный университет

Введение

RQ-системы характеризуются следующей отличительной особенностью: заявка нашедшая прибор занятым при поступлении в систему пробует занять прибор снова после осуществления некоторой случайной задержки. Необслуженные заявки после неудачной попытки занять прибор поступают на орбиту, откуда через некоторое время снова обращаются к прибору. RQ-системы являются адекватными моделями телекоммуникационных систем, компьютерных сетей. В монографиях [1,2] приводятся обширные исследования моделей с данной особенностью.

[*] Работа выполнена при финансовой поддержке гранта РФФИ № 18-01-00277.

В RQ-системах время, в течение которого прибор находится в незанятом состоянии, является временем простоя системы. В целях увеличения эффективности работы системы нелишним будет уменьшить время простоя. Мы рассматриваем системы, где прибор не только принимает заявки на обслуживание, поступающие в систему из входящего потока, но и вызывает заявки извне, когда находится в свободном состоянии. В теории массового обслуживания уже рассматривались модели, где заявки вызывались прибором извне [3], однако в рассмотренных моделях не было учтено повторное обращение заявок с орбиты. Модели с повторным обращением и с вызываемыми заявками имеют приложение в работе таких систем, как call-центры, где оператор в свободное время совершает звонки или занимается альтернативными видами деятельности. О приложениях RQ-систем в моделировании call-центров подробно изложено в работах [4,5]. RQ-системы с данной особенностью поведения прибора будем называть системами с вызываемыми заявками.

RQ-системы с вызываемыми заявками активно изучаются в последнее время [6,7]. В работах упомянутых выше исследуются марковские RQ-системы с вызываемыми заявками. Модель RQ-системы с несколькими типами вызываемых заявок была рассмотрена в статье [8]. Для такой модели был получен численный алгоритм нахождения стационарного распределения вероятностей числа заявок на орбите. Также рассматривалась многолинейная RQ-система [9].

В данной работе в качестве метода исследования рассматриваемых моделей предлагается метод асимптотического анализа. В работе [10] метод асимптотического анализа применяется для исследования RQ-системы M|M|1 с вызываемыми заявками в условии большой задержки заявок на орбите.

В данной работе рассматривается RQ-система MMPP|M|1 с N типами вызываемых заявок. Мы предполагаем, что интенсивности вызова и длительности обслуживания разнотипных заявок различны. Цель данного исследования рассмотреть предложенную систему в условии предельно редких изменений состояний входящего потока.

1. Постановка задачи

Рассмотрим RQ-систему с одним обслуживающим прибором, на вход которой поступает MMPP-поток заявок. Заявки, поступающие в систему из потока, занимают прибор для обслуживания и обслуживаются экспоненциальное случайное время с параметром μ_1, если прибор свободен. Если поступившая заявка застает прибор занятым, она отправляется на орбиту, где осуществляет случайную задержку. Длительность задержки экспоненциальная с параметром σ.

Когда прибор свободен, он вызывает заявки извне. В системе N типов вызываемых заявок. Прибор вызывает заявки типа n с интенсивностью α_n. Вызванная заявка мгновенно занимает прибор для обслуживания. Время обслуживания вызываемой заявки типа n распределено по экспоненциальному закону с параметром μ_n. Для удобства исследования пронумеруем типы вызываемых заявок от 2 до $N+1$. Рис. 1 иллюстрирует структуру описанной модели.

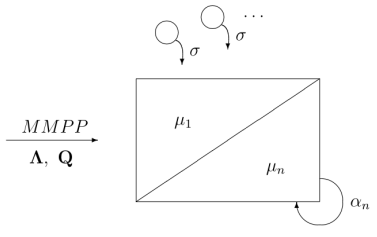


Рис. 1. RQ-система MMPP|M|1 с N типами вызываемых заявок

Пусть $i(t)$ — число поступивших заявок в системе в момент времени t, без учета вызванной заявки, если она обслуживается на приборе. Ставится задача нахождения стационарного распределения вероятностей данного процесса методом асимптотического анализа в условии предельно редких изменений состояний входящего потока.

2. Система уравнений Колмогорова

Случайный процесс $i(t)$ не является марковским, поэтому введем дополнительный процесс $k(t)$, характеризующий состояние прибора в момент времени t следующим образом: 0, если прибор свободен, 1, если прибор занят обслуживанием поступившей заявки, n, если прибор занят обслуживанием вызванной заявки типа n, где $n = \overline{2, N+1}$. Также введем процесс $m(t)$ – цепь Маркова, управляющую ММРР-потоком. Процесс $m(t)$ задается матрицей инфинитезимальных характеристик **Q**. Когда $m(t) = m$, интенсивность входящего потока равна λ_m, $m = \overline{1, M}$.

Трехмерный процесс $\{i(t), k(t), m(t)\}$ образует цепь Маркова с непрерывным временем. Обозначим $P\{i(t) = i, k(t) = k, m(t) = m\} = P_k(i, m, t)$ – вероятность того, что в момент времени t прибор находится в состоянии k, в системе находится i поступивших заявок и управляющая ММРР-потоком цепь Маркова $m(t)$ принимает значение m.

Запишем систему уравнений Колмогорова в стационарном режиме:

$$-\left(\lambda_m + i\sigma + \sum_{n=2}^{N+1}\alpha_n\right)P_0(i,m) + \mu_1 P_1(i+1,m) + \sum_{n=2}^{N+1}\mu_n P_n(i,m) + \sum_{\nu=1}^{M}P_0(i,\nu)q_{\nu m} = 0,$$

$$-(\lambda_m + \mu_1)P_1(i,m) + \lambda_m P_1(i-1,m) + \lambda_m P_0(i-1,m) + i\sigma P_0(i,m) + \sum_{\nu=1}^{M}P_1(i,\nu)q_{\nu m} = 0,$$

$$-(\lambda_m + \mu_n)P_n(i,m) + \lambda_m P_n(i-1,m) + \alpha_n P_0(i,m) + \sum_{\nu=1}^{M}P_n(i,\nu)q_{\nu m} = 0,\ n = \overline{2, N+1}.$$

Перейдем к частичным характеристическим функциям

$$H_k(u,m) = \sum_{i=0}^{\infty} e^{jui}P_k(i,m),\ k = \overline{0, N+1},$$

где $j = \sqrt{-1}$. Домножим уравнения системы на e^{jui} и просуммируем по i, чтобы получить систему уравнений для частичных характеристических функций

$$\begin{gathered}
-\left(\lambda_m + \sum_{n=2}^{N+1}\alpha_n\right)H_0(u,m) + j\sigma\frac{\partial H_0(u,m)}{\partial u} + \mu_1 e^{-ju}H_1(u,m) + \sum_{n=2}^{N+1}\mu_n H_n(u,m) + \\
+\sum_{\nu=1}^{M}H_0(u,\nu)q_{\nu m} = 0, \\
-(\lambda_m + \mu_1)H_1(u,m) + \lambda_m e^{ju}H_1(u,m) + \lambda_m e^{ju}H_0(u,m) - j\sigma\frac{\partial H_0(u,m)}{\partial u} + \\
+\sum_{\nu=1}^{M}H_1(u,\nu)q_{\nu m} = 0, \\
-(\lambda_m + \mu_n)H_n(u,m) + \lambda_m e^{ju}H_n(u,m) + \alpha_n H_0(u,m) + \\
+\sum_{\nu=1}^{M}H_n(u,\nu)q_{\nu m} = 0,\ \ n = \overline{2, N+1}.
\end{gathered} \tag{1}$$

3. Метод асимптотического анализа

В системе уравнений (1) введем следующие замены:

$$q_{\nu m} = \varepsilon s_{\nu m},\ H_k(u,m) = F_k(u,m,\varepsilon),\ k = \overline{0, N+1},$$

и получим систему уравнений

$$\begin{gathered}
-\left(\lambda_m + \sum_{n=2}^{N+1}\alpha_n\right)F_0(u,m,\varepsilon) + \mu_1 e^{-ju}F_1(u,m,\varepsilon) + \sum_{n=2}^{N+1}\mu_n F_n(u,m,\varepsilon) + \\
+\varepsilon\sum_{\nu=1}^{M}F_0(u,\nu,\varepsilon)s_{\nu m} + j\sigma\frac{\partial F_0(u,m,\varepsilon)}{\partial u} = 0, \\
-\left(\lambda_m + \mu_1\right)F_1(u,m,\varepsilon) + \lambda_m e^{ju}F_1(u,m,\varepsilon) + \lambda_m e^{ju}F_0(u,m,\varepsilon) - \\
-j\sigma\frac{\partial F_0(u,m,\varepsilon)}{\partial u} + \varepsilon\sum_{\nu=1}^{M}F_1(u,\nu,\varepsilon)s_{\nu m} = 0, \\
-\left(\lambda_m + \mu_n\right)F_n(u,m,\varepsilon) + \lambda_m e^{-ju}F_n(u,m,\varepsilon) + \alpha_n F_0(u,m,\varepsilon) + \\
+\varepsilon\sum_{\nu=1}^{M}F_n(u,\nu,\varepsilon)s_{\nu m} = 0,\ n = \overline{2, N+1}.
\end{gathered} \tag{2}$$

Теорема 1. Если существуют пределы $\lim\limits_{\varepsilon\to 0}F_k(u,m,\varepsilon) = F_k(u,m)$, то

$$F(u) = \sum_{k=0}^{N+1}\sum_{m=1}^{M}F_k(u,m) = \sum_{m=1}^{M}C_m\frac{\mu_1}{\mu_1-\lambda_m}\left(1+\sum_{n=2}^{N+1}\frac{\alpha_n}{\mu_n+\lambda_m\left(1-e^{ju}\right)}\right)\times$$

$$\times\left(\frac{1-p_{1m}}{1-p_{1m}e^{ju}}\right)^{\frac{\lambda_m}{\sigma}(1+\nu_m)+1}\prod_{n=2}^{N+1}\left(\frac{1-p_{nm}}{1-p_{nm}e^{ju}}\right)^{\frac{\alpha_n(\theta_{nm}-\lambda_m)}{\sigma\theta_{nm}}},$$

где $p_{1m} = \dfrac{\lambda_m}{\mu_1}$, $\theta_{nm} = \mu_n + \lambda_m - \mu_1$, $\nu_m = \sum\limits_{n=2}^{N+1}\dfrac{\alpha_n}{\theta_{nm}}$, $p_{nm} = \dfrac{\lambda_m}{\mu_n+\lambda_m}$, $n = \overline{2, N+1}$.

Доказательство. В системе уравнений (2) перейдем к пределу при $\varepsilon\to 0$, получим

$$-\left(\lambda_m + \sum_{n=2}^{N+1}\alpha_n\right)F_0(u,m) + \mu_1 e^{-ju}F_1(u,m) + \sum_{n=2}^{N+1}\mu_n F_n(u,m) + j\sigma\frac{\partial F_0(u,m)}{\partial u} = 0,$$

$$-\left(\lambda_m + \mu_1\right)F_1(u,m) + \lambda_m e^{ju}F_1(u,m) + \lambda_m e^{ju}F_0(u,m) - j\sigma\frac{\partial F_0(u,m)}{\partial u} = 0,$$

$$-\left(\lambda_m + \mu_n\right)F_n(u,m) + \lambda_m e^{ju}F_n(u,m) + \alpha_n F_0(u,m) = 0,\ n = \overline{2, N+1}.$$

Данная система уравнений сводится к системе из одного дифференциального и $N+1$ алгебраических уравнений

$$F_1(u,m) = \frac{\lambda_m}{\mu_1+\lambda_m\left(1-e^{ju}\right)}F_0(u,m) - \frac{j\sigma}{\mu_1+\lambda_m\left(1-e^{ju}\right)}\cdot\frac{\partial F_0(u,m)}{\partial u},$$

$$F_n(u,m) = \frac{\alpha_n}{\mu_n+\lambda_m\left(1-e^{ju}\right)}F_0(u,m),\ n = \overline{2, N+1}, \tag{3}$$

$$\frac{\partial F_0(u,m)}{\partial u} = j\frac{\lambda_m}{\sigma}\left[\frac{\lambda_m e^{ju}}{\mu_1-\lambda_m e^{ju}} + \frac{\mu_1+\lambda_m\left(1-e^{ju}\right)}{\mu_1-\lambda_m e^{ju}}\sum_{n=2}^{N+1}\frac{\alpha_n}{\mu_n+\lambda_m\left(1-e^{ju}\right)}\right]F_0(u,m).$$

Проинтегрируем обе части третьего уравнения системы (3) по u:

$$\ln F_0(u,m) - \ln C_m = \int_0^u j\frac{\lambda_m}{\sigma}\left[\frac{\lambda_m e^{jx}}{\mu_1-\lambda_m e^{jx}} + \frac{\mu_1+\lambda_m\left(1-e^{jx}\right)}{\mu_1-\lambda_m e^{jx}}\sum_{n=2}^{N+1}\frac{\alpha_n e^{jx}}{\mu_n+\lambda_m\left(1-e^{jx}\right)}\right]dx,$$

где C_m – константа интегрирования. Решая дифференциальное уравнение, получим функцию $F_0(u,m)$ в явном виде:

$$F_0(u,m) = C_m \left(\frac{\mu_1 - \lambda_m}{\mu_1 - \lambda_m e^{ju}} \right)^{\frac{\lambda_m}{\sigma}(1+\nu_m)} \prod_{n=2}^{N+1} \left(\frac{\mu_n}{\mu_n + \lambda_m \left(1 - e^{ju}\right)} \right)^{\frac{\alpha_n(\theta_{nm} - \lambda_m)}{\sigma\theta_{nm}}},$$

где $\theta_{nm} = \mu_n + \lambda_m - \mu_1$, $\nu_m = \sum_{n=2}^{N+1} \frac{\alpha_n}{\theta_{nm}}$. С помощью соотношений (3) получим функции $F_k(u,m)$:

$$F_1(u,m) = C_m \frac{\lambda_m}{\mu_1 - \lambda_m e^{ju}} \left(\frac{\mu_1 - \lambda_m}{\mu_1 - \lambda_m e^{ju}} \right)^{\frac{\lambda_m}{\sigma}(1+\nu_m)} \prod_{n=2}^{N+1} \left(\frac{\mu_n}{\mu_n + \lambda_m \left(1 - e^{ju}\right)} \right)^{\frac{\alpha_n(\theta_{nm} - \lambda_m)}{\sigma\theta_{nm}}} \times$$

$$\times \left(1 + \sum_{n=2}^{N+1} \frac{\alpha_n}{\mu_n + \lambda_m \left(1 - e^{ju}\right)} \right),$$

$$F_n(u,m) = C_m \frac{\alpha_n}{\mu_n + \lambda_m \left(1 - e^{ju}\right)} \left(\frac{\mu_1 - \lambda_m}{\mu_1 - \lambda_m e^{ju}} \right)^{\frac{\lambda_m}{\sigma}(1+\nu_m)} \times$$

$$\times \prod_{n=2}^{N+1} \left(\frac{\mu_n}{\mu_n + \lambda_m \left(1 - e^{ju}\right)} \right)^{\frac{\alpha_n(\theta_{nm} - \lambda_m)}{\sigma\theta_{nm}}}, \; n = \overline{2, N+1}.$$

Просуммируем полученные функции по *k*:

$$\sum_{k=0}^{N+1} F_k(u,m) = F(u,m) = C_m \frac{\mu_1}{\mu_1 - \lambda_m} \left(+ \sum_{n=2}^{N+1} \frac{\alpha_n}{\mu_n + \lambda_m \left(1 - e^{ju}\right)} \right) \left(\frac{\mu_1 - \lambda_m}{\mu_1 - \lambda_m e^{ju}} \right)^{\frac{\lambda_m}{\sigma}(1+\nu_m)+1} \times$$

$$\times \prod_{n=2}^{N+1} \left(\frac{\mu_n}{\mu_n + \lambda_m \left(1 - e^{ju}\right)} \right)^{\frac{\alpha_n(\theta_{nm} - \lambda_m)}{\sigma\theta_{nm}}}.$$

Просуммируем полученные выражения для частичных характеристических функций по *m*:

$$F(u) = \sum_{m=1}^{M} C_m \frac{\mu_1}{\mu_1 - \lambda_m} \left(+ \sum_{n=2}^{N+1} \frac{\alpha_n}{\mu_n + \lambda_m \left(1 - e^{ju}\right)} \right) \left(\frac{1 - p_{1m}}{1 - p_{1m} e^{ju}} \right)^{\frac{\lambda_m}{\sigma}(1+\nu_m)+1} \prod_{n=2}^{N+1} \left(\frac{1 - p_{nm}}{1 - p_{nm} e^{ju}} \right)^{\frac{\alpha_n(\theta_{nm} - \lambda_m)}{\sigma\theta_{nm}}},$$

где $p_{1m} = \frac{\lambda_m}{\mu_1}$, $p_{nm} = \frac{\lambda_m}{\mu_n + \lambda_m}$, $n = \overline{2, N+1}$. **Теорема доказана.**

В ходе доказательства теоремы 1 найдена характеристическая функция числа поступивших заявок в системе с помощью метода асимптотического анализа в условии предельно редких изменений состояний входящего потока.

Заключение

В настоящей работе рассмотрена математическая модель системы MMPP|M|1 с *N* типами вызываемых заявок. С помощью метода асимптотического анализа в условии предельно редких изменений входящего потока получена характеристическая функция числа поступивших заявок в системе.

Для дальнейшего исследования данной системы планируется построить имитационную модель, с помощью которой можно определить границы применимости асимптотического результата. Также планируется изучение RQ-систем M|GI|1 и GI|M|1 с N типами вызываемых заявок с помощью методов асимптотического анализа в различных предельных условиях.

ЛИТЕРАТУРА

1. *Artalejo J.R., Gomez-Corral A.* Retrial queueing systems: a computational approach. – Springer, Berlin, 2008. – 320 p.

2. *Falin G., Templeton J.* Retrial queues. – CRC Press, 1997.

3. *Башарин Г.П., Самуйлов К.Е.* Об однофазной системе массового обслуживания с двумя типами заявок и относительным приоритетом // Техническая кибернетика. – 1983. – Т. 3. – С. 48–56.

4. *Bhulai S., Koole G.* A queueing model for call blending in call centers // IEEE Transactions on Automatic Control. – 2003. – Vol. 48. – № 8. – P. 1434–1438.

5. *Deslauriers A., L'Ecuyer P., Pichitlamken J., Ingolfsson A., Avramidis A.N.* Markov chain models of a telephone call center with call blending // Computers & operations research. – 2007. – Vol. 34. – № 6. P. 1616–1645.

6. *Artalejo J.R., Phung-Duc T.* Markovian retrial queues with two-way communication // Journal of industrial and management optimization. – 2012. – Vol. 12. – № 4. – P. 781–806.

7. *Artalejo J.R., Phung-Duc T.* Single server retrial queues with two-way communication // Applied Mathematical Modelling. – 2013. – Vol. 37. – № 4. – P. 1811–1822.

8. *Sakurai H., Phung-Duc T.* Two-way communication retrial queues with multiple types of outgoing calls // Top. – 2015. – Vol. 23. – № 2. – P. 466–492.

9. *Phung-Duc T., Kawanishi K.* An efficient method for performance analysis of blended call centers with redial // Asia-Pacific Journal of Operational Research. – 2014. – Vol. 31. – № 02. – P. 144–148.

10. *Nazarov A.A.* Asymptotic analysis of markovial retrial queue with two-way communication under low rate of retrials condition // In Proceedings 31st European Conference on Modelling and Simulation. – 2017. – P. 678–693.

ЧИСЛЕННЫЙ АНАЛИЗ СУММАРНОГО ОБЪЕМА ЗАНЯТОГО РЕСУРСА НА БЛОКАХ СИСТЕМЫ С ПАРАЛЛЕЛЬНЫМ ОБСЛУЖИВАНИЕМ И ВХОДЯЩИМ MMPP-ПОТОКОМ

Е.Н. Чернышова, Е.Ю. Лисовская

Томский государственный университет

Введение

В настоящее время наблюдается увеличение объема трафика беспроводных сетей, что обусловлено развитием технологий мобильного широкополосного доступа четвертого поколения (4G), а также постоянно растущим запросом на трафик в системах пятого поколения (5G) [2,3]. Анализ источников, рекомендаций и стандартов международных организаций установил, что необходимы комплексные модели, которые адекватно описывали бы особенности методов управления доступом и механизмами резервирования ресурсов сетей [1,4].

Для увеличения эффективности механизма управления ресурсами беспроводной сети, необходимо разрабатывать новые модели и методы анализа эффективности широкополосных беспроводных сетей. Решение данной проблемы сдерживается недостатками в теоретических основах, в том числе, в теории массового обслуживания из-за отсутствия адекватных моделей обслуживания и методов их анализа.

В данной работе предлагается построение и исследование математической модели системы массового обслуживания со случайными требованиями к объему занимаемого ресурса беспроводной сети с расщеплением (копированием) заявок [5]. Рассматриваемая модель позволит сделать оценку необходимых объемов резервируемых ресурсов для трафика и выработать стратегию их распределения.

1. Математическая модель

Рассмотрим систему массового обслуживания, состоящую из двух блоков, каждый из которых имеет неограниченное число приборов; на вход поступает MMPP поток

требований, управляемый цепью Маркова $k(t)$, заданный матрицами инфинитезимальных характеристик Q и условных интенсивностей Λ (рис. 1).

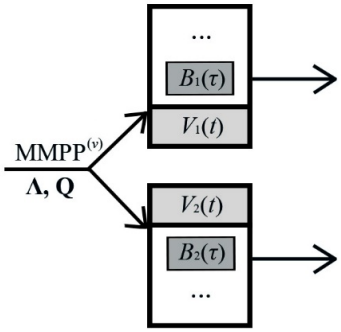


Рис. 1. Ресурсная СМО с MMPP потоком и параллельным обслуживанием

Дисциплина обслуживания определяется следующим образом. Заявка, пришедшая в СМО, попадает в каждый блок и мгновенно занимает любой свободный прибор, где обслуживается в течение случайного времени. Функция распределения вероятностей времени обслуживания на первом блоке – $B_1(\tau) = P\{\xi_1 < \tau\}$, на втором – $B_2(\tau) = P\{\xi_2 < \tau\}$. Функции распределения вероятностей для случайного объема ресурсов $G_1(y) = P\{v_1 < y\}$, $G_2(y) = P\{v_2 < y\}$ соответственно. При завершении обслуживания, требование покидает систему, освобождая прибор и все занимаемые ресурсы. Количество занимаемого ресурса и время обслуживания не зависят друг от друга. Пусть $V_1(t)$ и $V_2(t)$ – суммарный объём занятого ресурса на первом и втором блоках соответственно. Ставится задача исследования случайного двумерного процесса $\{V_1(t), V_2(t)\}$.

Ранее была сформулирована и доказана следующая теорема.

Теорема. Совместное асимптотическое распределение вероятностей двумерного процесса суммарных объемов занятого ресурса на блоках СМО с параллельным обслуживанием в условиях растущей интенсивности входящего потока и предельно частых изменениях состояний управляющей MMPP-потоком цепи Маркова является двумерным гауссовским распределением вероятностей с параметрами: вектор математических ожиданий

$$\mathbf{a} = \left[N\lambda a_1^{(1)} b_1, N\lambda a_2^{(1)} b_2 \right],$$

матрица ковариаций:

$$\mathbf{K} = \begin{bmatrix} N\lambda a_1^{(2)} b_1 + N\kappa \left(a_1^{(1)}\right)^2 \beta_1 & N a_1^{(1)} a_2^{(1)} b_{12} (\lambda + \kappa) \\ N a_1^{(1)} a_2^{(1)} b_{12} (\lambda + \kappa) & N\lambda a_2^{(2)} b_2 + N\kappa \left(a_2^{(1)}\right)^2 \beta_2 \end{bmatrix},$$

где $\lambda = \mathbf{r}\overline{\mathbf{\Lambda}}\mathbf{e}$, $\kappa = 2\mathbf{g}\left(\overline{\mathbf{\Lambda}} - \lambda \mathbf{I}\right)\mathbf{e}$, $a_i^{(1)}$, $a_i^{(2)}$, $i = 1,2$ – первые и вторые начальные моменты объема занимаемого ресурса каждой заявкой,

$$b_i = \int_0^\infty \left(1 - B_i(\tau)\right) d\tau, \; \beta_i = \int_0^\infty \left(1 - B_i(\tau)\right)^2 d\tau, \; i = 1,2,$$

$$b_{12} = \int_0^\infty \left(1 - B_1(\tau)\right)\left(1 - B_2(\tau)\right) d\tau.$$

Цель работы – с помощью имитационной модели получить эмпирическое распределение вероятностей суммарных объемов занятых ресурсов в системе, сравнить их с

полученными асимптотическими распределениями, определить область применимости результатов.

2. Имитационное моделирование и численный анализ

Эмпирическое распределение вероятностей построим на основе имитационной модели [6], для этого в существующей программе, учтем возможность «копирования» каждого поступающего требования на второй блок.

Исходные данные для рассматриваемой ресурсной системы с параллельным обслуживанием зададим следующим образом:

1. Входящий MMPP-поток определяется матрицами:

$$\mathbf{\Lambda} = N\begin{bmatrix} 0.5 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1.5 \end{bmatrix},\ \mathbf{Q} = N\begin{bmatrix} -0.8 & 0,4 & 0,4 \\ 0.3 & -0.6 & 0,3 \\ 0.4 & 0.4 & -0.8 \end{bmatrix},$$

где N – параметр растущей интенсивности входящего потока и предельно частых изменений состояний управляющей цепи Маркова. Таким образом, получим интенсивность потока $\lambda = N$;

2. Время обслуживания имеет гамма-распределение вероятностей с параметрами $\alpha_1 = \beta_1 = 0.5$ и $\alpha_2 = \beta_2 = 1.5$ для первого и второго блоков, соответственно, тогда среднее время обслуживания $b_1 = 1$ и $b_2 = 1$.

3. Количество ресурса необходимого для обслуживания одной заявки имеет равномерное на [1,2] и на [0,1] распределение вероятностей, соответственно для первого и второго блоков.

В табл. 1 представлен текст программы для вычисления параметров асимптотического распределения вероятностей суммарных объемов занятого ресурса в системе с параллельным обслуживанием и входящим MMPP-потоком, реализованной в виде рабочих листов системы MathCad.

Таблица 1

Блок вычисления параметров асимптотического распределения вероятностей

Параметры потока	
$Q := \begin{pmatrix} -0.8 & 0.4 & 0.4 \\ 0.3 & -0.6 & 0.3 \\ 0.4 & 0.4 & -0.8 \end{pmatrix}$ $\Lambda := \begin{pmatrix} 0.5 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1.5 \end{pmatrix}$	
Параметры обслуживания	
$\alpha 1 := 0.5$ $\beta 1 := \alpha 1$ $B1(x) := pgamma(x \cdot \beta 1, \alpha 1)$	$\alpha 2 := 1.5$ $\beta 2 := \alpha 2$ $B2(x) := pgamma(x \cdot \beta 2, \alpha 2)$
$b1 := \int_0^\infty (1 - B1(x))\,dx$ $\beta 1 := \int_0^\infty (1 - B1(x))^2\,dx$	$b2 := \int_0^\infty (1 - B2(x))\,dx$ $\beta 2 := \int_0^\infty (1 - B2(x))^2\,dx$
$b12 := \int_0^\infty (1 - B1(x)) \cdot (1 - B2(x))\,dx$	
Параметры занимаемого ресурса	
$a1_1 := \frac{1}{b_1 - a_1} \cdot \int_{a_1}^{b_1} y\,dy$ $a2_1 := \frac{1}{b_1 - a_1} \cdot \int_{a_1}^{b_1} y^2\,dy$	$a1_2 := \frac{1}{b_2 - a_2} \cdot \int_{a_2}^{b_2} y\,dy$ $a2_2 := \frac{1}{b_2 - a_2} \cdot \int_{a_2}^{b_2} y^2\,dy$
Первый блок	**Второй блок**
Асимптотическое среднее	

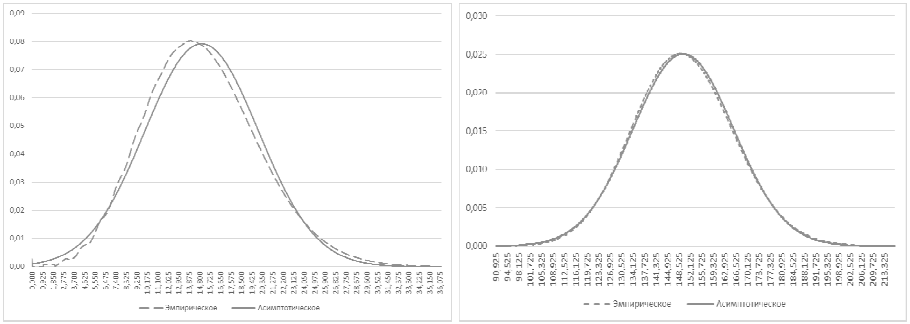
$N \cdot \lambda \cdot a1_1 \cdot b1 = 1.5N$	$N \cdot \lambda \cdot a1_2 \cdot b2 = 0.5N$
Асимптотическая дисперсия	
$N \cdot \left[\lambda \cdot a2_1 \cdot b1 + \kappa \cdot (a1_1)^2 \cdot \beta 1\right] = 2.538N$	$N \cdot \left[\lambda \cdot a2_2 \cdot b2 + \kappa \cdot (a1_2)^2 \cdot \beta 2\right] = 0.369N$
Асимптотический коэффициент корреляции	
$\frac{N \cdot a1_1 \cdot a1_2 \cdot b12 \cdot (\lambda + \kappa)}{N \cdot \left[\left[\lambda \cdot a2_1 \cdot b1 + \kappa \cdot (a1_1)^2 \cdot \beta 1\right] \cdot \left[\lambda \cdot a2_2 \cdot b2 + \kappa \cdot (a1_2)^2 \cdot \beta 2\right]\right]^{\frac{1}{2}}} = 0.434$	

Была проведена серия имитационных экспериментов с увеличением значений N. В табл. 2 приведены значения расстояний Колмогорова между функциями распределения вероятностей суммарного объема занятого ресурса на первом и втором блоках системы – Δ_1 и Δ_2, соответственно, а также между функциями распределения двумерных распределений вероятностей суммарных объемов на обоих блоках Δ_{12} для различных значений параметра N. На рис. 2, 3 изображены графики распределения вероятностей суммарного объема занятого ресурса на первом и втором блоках системы, соответственно.

Таблица 2

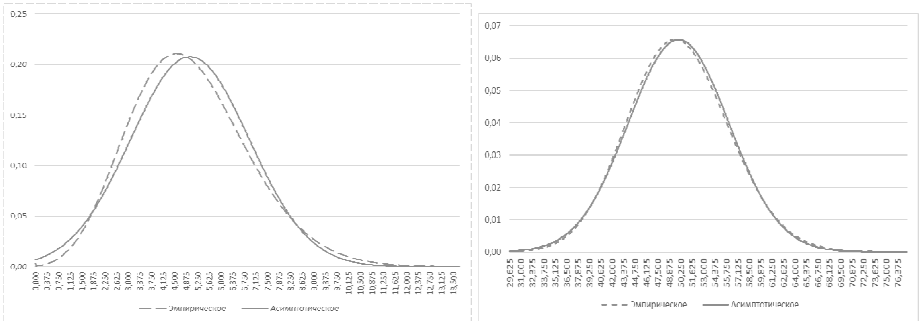
Расстояния Колмогорова между асимптотической и эмпирической функциями распределения вероятностей

N	1	3	5	7	10	20	50	100
Δ_1	0.373	0.111	0.061	0.041	0.028	0.018	0.011	0.008
Δ_2	0.373	0.101	0.053	0.036	0.025	0.016	0.010	0.007
Δ_{12}	0.369	0.110	0.059	0.040	0.028	0.018	0.011	0.008



а) $N = 10$ б) $N = 100$

Рис. 2. Распределение вероятностей суммарного объема занятого ресурса на первом блоке системы



а) $N = 10$ б) $N = 100$

Рис. 3. Распределение вероятностей суммарного объема занятого ресурса на втором блоке системы

Из табл. 2 можем сделать вывод о том, что точность аппроксимации возрастает при увеличении параметра N (при увеличении интенсивности входящего потока), а

рис. 2, 3демонстрируют это. Принимая допустимое значение расстояния Колмогорова меньше 0.03, сделаем следующий вывод: суммарный объем занятого ресурса на блоках системы с параллельным обслуживанием имеет гауссовское распределение, когда интенсивность обслуживания больше интенсивности обслуживания в 10 и более раз.

Заключение

В данной работе найдены основные вероятностные характеристики ресурсной системы с параллельным обслуживанием и входящим MMPP потоком. С помощью математического пакета программ MathCAD найдены параметры асимптотического (гауссовского) распределения вероятностей суммарных объемов занятого ресурса на блоках системы. Представлен численный анализ точности аппроксимации, а также определена область применимости асимптотических результатов.

ЛИТЕРАТУРА

1. *Boccardi F., Heath Jr R.W., Lozano A., Marzetta T.L., Popovski P.* Five disruptive technology directions for 5G // IEEE Communications Magazine. – 2014. – № 52(2). – P. 74–80.

2. Cisco, 2017. The Zettabyte Era: Trends and analysis. White Paper. Avaiable at: http://wikiurls.com/?http://www.cisco.com/c/en/us/solutions/collateral/service-provider/visual-networking-indexvni/vni-hyperconnectivity-wp.html (accessed June 17, 2017).

3. *Marsch P, Raaf B, Szufarska A, Mogensen P, Guan H, Farber M, Redana S, Pedersen K, Kolding T.* Future mobile communication networks: challenges in the design and operation // IEEE Vehicular Technology Magazine. – 2012. – №7(1). – P. 16–23.

4. *Бегишев В.О.* Вероятностные модели резервирования ресурсов и анализ показателей эффективности беспроводных сетей с трафиком взаимодействия устройств: Дис. ... канд. физ.-мат. наук. Москва, 2019. 115 с.

5. *Лисовская Е.Ю.* Асимптотические методы исследования ресурсных СМО с непуассоновскими входящими потоками: Дис. ... канд. физ.-мат. наук. Томск, 2018. 138 с.

6. *Лисовская Е.Ю., Моисеева С. П., Pagano M.* Имитационная модель бесконечнолинейной системы обслуживания требований случайного объема с MMPP входящим потоком // Свидетельство № 2017615302 от 12.05.2017 г. о государственной регистрации программы для ЭВМ. – 2017.

МОДЕЛЬ ОБСЛУЖИВАНИЯ ТРАФИКА МУЛЬТИВЕЩАНИЯ ТОЧКОЙ ДОСТУПА МИЛЛИМЕТРОВОГО ДИАПАЗОНА С УЧЕТОМ ФОНОВОГО ТРАФИКА*

В.А. Просвиров, И.С. Ярцева, А.А. Платонова, Т.Д. Фомченко, Ю.В. Гайдамака

Российский Университет Дружбы Народов

Введение

Стандарт 5G New Radio был принят некоммерческим консорциумом 3GPP в марте 2017 года, а уже летом 2019 в Южной Корее планируется запуск первых сетей 5G [1,2]. Особенностью сетей пятого поколения является использование миллиметрового диапазон длин волн для обеспечения высокой спектральной эффективности и скорости передачи данных. Достоинствами такого подхода является наличие дополнительного радиоресурса в спектре и малые размеры антенн. К недостаткам можно отнести сильное затухание в атмосфере и обязательное наличие прямой видимости (Line of Sight, LoS) между передающим и принимающим устройством.

В работе исследована модель точки доступа в миллиметровом диапазоне длин волн, которая обслуживает запросы на установление одноадресных и многоадресных соединений. Построение модели в виде марковского процесса выполнено с учетом наличия случайных требований к ресурсам, зависящих от расположения пользователей, а также учтена возможность временного нарушения прямой видимости между устройствами из-за так называемых блокираторов. Матрица интенсивностей переходов процесса представлена в блочном виде. В качестве основных показателей производитель-

* Публикация подготовлена при поддержке Программы РУДН «5-100» и при финансовой поддержке РФФИ в рамках научных проектов № 17-00-0845 и № 18-00-0156.

ности системы исследованы вероятности сброса одноадресных и многоадресных сессий, а также средний объем занятого ресурса.

1. Постановка задачи

При установлении соединения одноадресного типа узконаправленные антенны создают высокий коэффициент усиления сигнала на стороне пользовательского оборудования и на стороне точки доступа. Такая особенность позволяет устанавливать точки доступа на большом расстоянии друг от друга без значительных потерь мощности сигнала из-за затухания на расстоянии. Для многоадресных соединений характерным является использование антенн с широкой диаграммой направленности, т.к. в соответствии с принципом мультивещания передача данных для всех пользователей в зоне покрытия осуществляется однократно, что позволяет существенно экономить ресурс системы, затрачиваемый на обслуживание многоадресных соединений. Однако, при передаче данных с широкой диаграммой направленности радиус сектора, соответствующего зоне покрытия, значительно меньше, чем при узконаправленной передаче. Таким образом, при проектировании точки доступа в миллиметровом диапазоне длин волн задачей является достижение баланса между противоречащими друг другу требованиями двух типов обслуживаемого трафика для эффективного функционирования системы. Результатом данного исследования является построение математической модели системы беспроводной связи для обслуживания комбинации двух типов трафика, позволяющей проводить анализ таких показателей производительности системы как вероятность сброса запросов на установление соединений одноадресного и многоадресного типа, а также коэффициента использования ресурса [3].

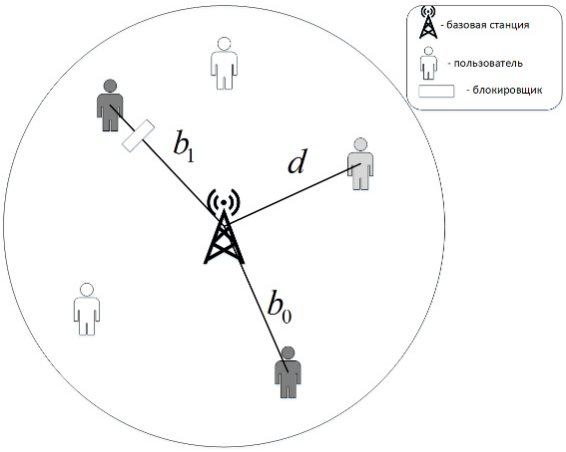


Рис. 1. Модель системы

Зона покрытия имеет форму круга, радиус которого определяется моделью распространения сигнала (рис. 1). Для обслуживания запросов пользователей, случайным образом распределенных внутри круга, точкой доступа выделяется радиоресурс, при этом объем ресурса зависит от местоположения пользователя. В данном исследовании модель построена в предположении о постоянных требованиях к ресурсу: для обслуживания каждого запроса на установление одноадресного соединения требуется d единиц ресурса из R имеющихся в системе; для обслуживания многоадресных запросов требуется b_0 единиц ресурса, при условии наличия прямой видимости между устройствами и b_1 единиц ресурса в случае блокировки прямой видимости. Модель учитывает возможность блокировки прямой видимости препятствиями для пользователей, уже установивших многоадресное соединение. При появлении препятствия на линии прямой видимости, соответствующей соединению между точкой доступа и пользовательским устройством, соединение не прерывается, но количество ресурса, выделенного точкой

доступа для его поддержания, увеличивается. Для запросов на одноадресные соединения может возникнуть отказ в установлении нового соединения (сброс сессии), если количество свободного ресурса недостаточно для установления нового одноадресного соединения, т.е. нет свободных d единиц ресурса. Также сброс запроса на одноадресные соединения происходит, если в момент появления блокатора прямой видимости для уже установленного многоадресного соединения нет свободного $b_1 - b_0$ единиц ресурса, который необходимо выделить дополнительно для обслуживания многоадресных сессий в состоянии nLoS. В этом случае сбрасывается несколько случайным образом выбранных одноадресных соединений, чтобы освободить необходимое количество $b_1 - b_0$ единиц ресурса для обслуживания многоадресных соединений в состоянии nLoS. Для запросов на многоадресные сессии может произойти сброс, если в момент его поступления нет ни одной сессии такого типа и свободного ресурса недостаточно для установления нового многоадресного соединения.

2. Построение математической модели системы

Состояние системы описывается случайным процессом $\mathbf{X}(t)$ над пространством состояний

$$X = \left\{ (s,m,n) : s = \{0,1\}, n = \{0,1\}, 0 \le n \le \left\lfloor \frac{R}{d} \right\rfloor,\ c(s,m,n) \le R \right\},$$

где занятый ресурс в состоянии $(s,m,n) \in X$ определяет функция

$$c(s,m,n) = mb_0 + sm\left(b_1 - b_0\right) + nd\ .$$

Здесь $s \in \{0,1\}$ – индикатор наличия блокираторов (0 – нет, 1 – есть), $m \in \{0,1\}$ – индикатор наличия заявок многоадресного типа в системе (0 – нет ни одной, 1 – есть одна или более), а число одноадресных соединений в системе равно $n \in \left\{ 0,1,...,\left\lfloor \frac{R-b_1}{d} \right\rfloor,...,\left\lfloor \frac{R-b_0}{d} \right\rfloor,...,\left\lfloor \frac{R}{d} \right\rfloor \right\}$.

Диаграмма интенсивностей переходов для такой системы представлена на рис. 2.

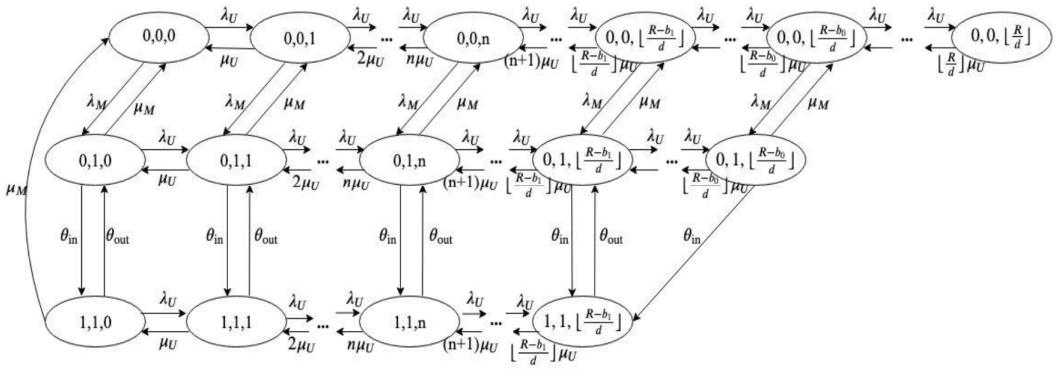


Рис. 2. Диаграмма интенсивностей переходов СП $\mathbf{X}(t)$

Вероятности сброса одноадресных и многоадресных заявок вычисляются по формулам

$$B_U = p\left(0,0,\left\lfloor \frac{R}{d} \right\rfloor\right) + \frac{\theta_{in}}{\lambda_U + n\mu_U + \mu_M + \theta_{in}} \sum_{n=\left\lfloor \frac{R-b_1}{d} \right\rfloor+1}^{\left\lfloor \frac{R-b_0}{d} \right\rfloor-1} p(0,1,n) +$$

$$+ \frac{\theta_{in}\, p\left(0,1,\left\lfloor \frac{R-b_0}{d} \right\rfloor\right)}{n\mu_U + \mu_M + \theta_{in}} + p\left(1,1,\left\lfloor \frac{R-b_1}{d} \right\rfloor\right),$$

$$B_M = \sum_{n=\left\lfloor \frac{R-b_0}{d} \right\rfloor+1}^{\left\lfloor \frac{R}{d} \right\rfloor} p(0,0,n)\,.$$

Среднее количество занятого ресурса одноадресными и многоадресными заявками имеет вид

$$\overline{R_U} = \sum_{s=0}^{1}\sum_{m=0}^{1}\sum_{n=0}^{\left\lfloor \frac{R-mb_s}{d} \right\rfloor} n\cdot d\cdot p(s,m,n)\,,\ \overline{R_M} = \sum_{s=0}^{1}\sum_{n=0}^{\left\lfloor \frac{R-b_s}{d} \right\rfloor} b_s \cdot p(s,1,n)\,.$$

Для нахождения стационарных вероятностей системы необходимо решить систему уравнений глобального баланса (СУГБ):

$$\left(\lambda_U + \lambda_M\right) p(0,0,0) = \mu_U\, p(0,0,1) + \mu_M\,(p(0,1,0) + p(1,1,0))\,,$$

$$\left(\lambda_U + \lambda_M + n\mu_U\right) p(0,0,n) =$$

$$= \lambda_U\, p(0,0,n-1) + (n+1)\mu_U\, p(0,0,n+1) + \mu_M\, p(0,1,n),\ \ n = \overline{1,\left\lfloor \frac{R}{d} \right\rfloor - 1},$$

$$\left(\lambda_U + \mu_M + n\mu_U + \theta_{in}\right) p(0,1,n) =$$

$$= \lambda_U\, p(0,1,n-1) + \lambda_M\, p(0,0,n) + (n+1)\mu_U\, p(0,1,n+1) + \theta_{out}\, p(1,1,n),\ \ n = \overline{0,\left\lfloor \frac{R-b_0}{d} \right\rfloor - 1},$$

$$\left(\mu_M + \theta_{out}\right) p(1,1,0) = \mu_U\, p(1,1,1) + \theta_{in}\, p(0,1,0)\,,$$

$$\left(n\mu_M + \theta_{out}\right) p(1,1,n) = (n+1)\mu_U\, p(1,1,n+1) + \theta_{in}\, p(0,1,n),\ \ n = \overline{1,\left\lfloor \frac{R-b_1}{d} \right\rfloor - 1}\,,$$

$$\left\lfloor \frac{R}{d} \right\rfloor \mu_U\, p\left(0,0,\left\lfloor \frac{R}{d} \right\rfloor\right) = \lambda_U\, p\left(0,0,\left\lfloor \frac{R}{d} \right\rfloor - 1\right),$$

$$\left(\mu_M + \left\lfloor \frac{R-b_1}{d} \right\rfloor \mu_U + \theta_{in}\right) p\left(0,1,\left\lfloor \frac{R-b_0}{d} \right\rfloor\right) = \lambda_M\, p\left(0,0,\left\lfloor \frac{R-b_0}{d} \right\rfloor\right) + \lambda_U\, p\left(0,1,\left\lfloor \frac{R-b_0}{d} \right\rfloor - 1\right),$$

$$\left(\left\lfloor \frac{R-b_1}{d} \right\rfloor \mu_U + \theta_{out}\right) p\left(1,1,\left\lfloor \frac{R-b_1}{d} \right\rfloor\right) = \theta_{in} \sum_{n=\left\lfloor \frac{R-b_1}{d} \right\rfloor}^{\left\lfloor \frac{R-b_0}{d} \right\rfloor} p(0,1,n)\,.$$

Матрица **A** в СУГБ $\vec{p}^T\mathbf{A}=\vec{0}^T$ имеет блочно-диагональный вид:

A	X_1	X_2	X_3
X_1	$\mathbf{D_1}$	$\mathbf{U_1}$	**0**
X_2	$\mathbf{L_1}$	$\mathbf{D_2}$	$\mathbf{U_2}$
X_3	$\mathbf{L_3}$	$\mathbf{L_2}$	$\mathbf{D_3}$

Здесь X_1, X_2 и X_3 – подмножества пространства состояний:

$$X_1=\left\{(s,m,n)\in X: s=0, m=0, n\in\left\{0,1,\ldots,\frac{R}{d}\right\}\right\},$$

$$X_2=\left\{(s,m,n)\in X: s=0, m=1, n\in\left\{0,1,\ldots,\frac{R-b_0}{d}\right\}\right\},$$

$$X_3=\left\{(s,m,n)\in X: s=1, m=1, n\in\left\{0,1,\ldots,\frac{R-b_1}{d}\right\}\right\},$$

причем $X=X_1\cup X_2\cup X_3$.

Элементы блоков определяются по следующим формулам:

$$d_1(i,j)=\begin{cases}\lambda_U,\ i=(0,0,n),\ j=(0,0,n+1),\ n=0,\ldots,\left\lfloor\frac{R}{d}\right\rfloor-1,\\ n\mu_U,\ i=(0,0,n),\ j=(0,0,n-1),\ n=1,\ldots,\left\lfloor\frac{R}{d}\right\rfloor,\\ -\lambda_U-n\mu_U-\lambda_M,\ i=(0,0,n),\ j=(0,0,n),\ n=0,\ldots,\left\lfloor\frac{R}{d}\right\rfloor-1,\\ -\left\lfloor\frac{R}{d}\right\rfloor\mu_U-\lambda_M,\ i=\left(0,0,\left\lfloor\frac{R}{d}\right\rfloor\right),\ j=\left(0,0,\left\lfloor\frac{R}{d}\right\rfloor\right),\end{cases}$$

$$d_2(i,j)=\begin{cases}\lambda_U,\ i=(0,1,n),\ j=(0,1,n+1),\ n=0,\ldots,\left\lfloor\frac{R-b_0}{d}\right\rfloor-1,\\ n\mu_U,\ i=(0,1,n),\ j=(0,1,n-1),\ n=1,\ldots,\left\lfloor\frac{R-b_0}{d}\right\rfloor,\\ -n\mu_U-\mu_M-\theta_{in},\ i=(0,1,n),\ j=(0,1,n),\ n=0,\ldots,\left\lfloor\frac{R-b_0}{d}\right\rfloor,\end{cases}$$

$$d_3(i,j)=\begin{cases}\lambda_U,\ i=(1,1,n),\ j=(1,1,n+1),\ ,\ n=0,...,\left\lfloor\dfrac{R-b_1}{d}\right\rfloor-1,\\ n\mu_U,\ i=(1,1,n),\ j=(1,1,n-1),\ n=1,...,\left\lfloor\dfrac{R-b_1}{d}\right\rfloor,\\ -\lambda_U-\theta_{out}-\lambda_M,\ i=(1,1,n),\ j=(1,1,n),\ n=0,...,\left\lfloor\dfrac{R-b_1}{d}\right\rfloor-1,\\ -\left\lfloor\dfrac{R-b_1}{d}\right\rfloor\mu_U-\theta_{out},\ i=\left(1,1,\left\lfloor\dfrac{R-b_1}{d}\right\rfloor\right),\ j=\left(1,1,\left\lfloor\dfrac{R-b_1}{d}\right\rfloor\right),\end{cases}$$

$$u_1(i,j)=\lambda_M,\ i=(0,0,n),\ j=(0,1,n),\ n=0,...,\left\lfloor\frac{R-b_1}{d}\right\rfloor,$$

$$u_2(i,j)=\theta_{in},\ i=(0,0,n),\ j=(1,1,n),\ n=0,...,\left\lfloor\frac{R-b_1}{d}\right\rfloor,$$

$$l_1(i,j)=\mu_M,\ i=(0,1,n),\ j=(0,0,n),\ n=0,...,\left\lfloor\frac{R_1}{d}\right\rfloor,$$

$$l_2(i,j)=\theta_{out},\ i=(1,1,n),\ j=(0,1,n),\ n=0,...,\left\lfloor\frac{R-b_0}{d}\right\rfloor,$$

$$l_3(i,j)=\mu_M,\ i=(1,1,0),\ j=(0,0,0).$$

Исследуем СП $\mathbf{X}(t)$ и характеристики системы, которую он описывает, для следующего частного случая [4]. Объем ресурса точки доступа $R=694$ единицы ресурса, число единиц емкости ресурса для обслуживания многоадресных заявок в условиях прямой видимости $b_0=93$, число единиц емкости ресурса для обслуживания многоадресных заявок с наличием блокиратора одной из них $b_1=149$, число единиц емкости ресурса для обслуживания одноадресной заявки $d=156$.

В систему поступают два потока заявок, соответствующие потокам запросов пользователей на установление одноадресных и многоадресных соединений интенсивностей $\lambda_U=0.1$ и $\lambda_M=0.1$ соответственно. Интенсивность обслуживания одноадресных заявок $\mu_U=0.1$, а интенсивность обслуживания многоадресных заявок $\mu_M=0.1$. Движение блокираторов моделируется с помощью двух экспоненциально распределённых величин с параметрами $\theta_{in}=0.4$ и $\theta_{out}=0.5$, характеризующими длительности интервалов времени между соседними появлениями блокиратора и длительности интервала блокировки соответственно.

3. Пример численного анализа

Из рис. 3а следует, что при увеличении интенсивности поступления одноадресных заявок вероятность сброса B_U (вероятность прерывания одноадресного соединения) падает, т.к. заявки данного типа поступают чаще и занимают все доступные ресурсы. При этом вероятность B_M сброса многоадресных соединений растет, поскольку заявки одноадресного типа уже заняли доступные ресурсы и для обслуживания многоадресных заявок не остается нужного количество ресурсов сети. Этот же эффект можно заметить на рис. 3б, где при увеличении интенсивности поступления многоадресных заявок вероятность сброса B_M падает, т. к. заявки данного типа поступают чаще и занимают все доступные ресурсы. При этом вероятность B_U растет, поскольку заявки многоадресного типа уже заняли доступные ресурсы и для обслуживания одноадресных заявок не остается нужного количество ресурсов сети.

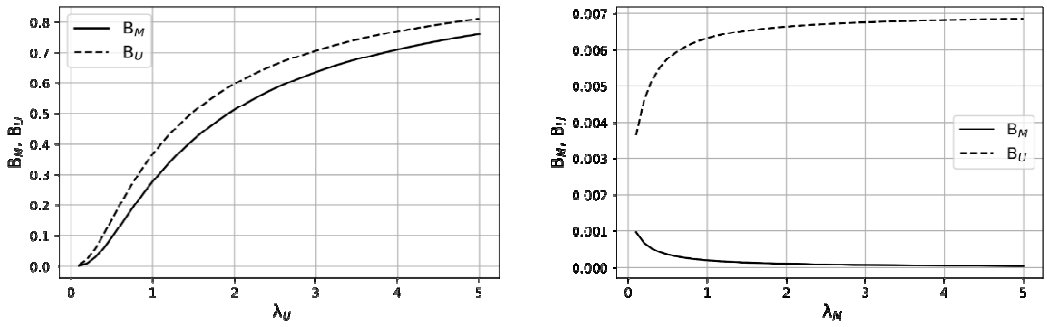


Рис. 3. Вероятности сброса в зависимости от интенсивности поступления
а) одноадресных заявок б) многоадресных заявок

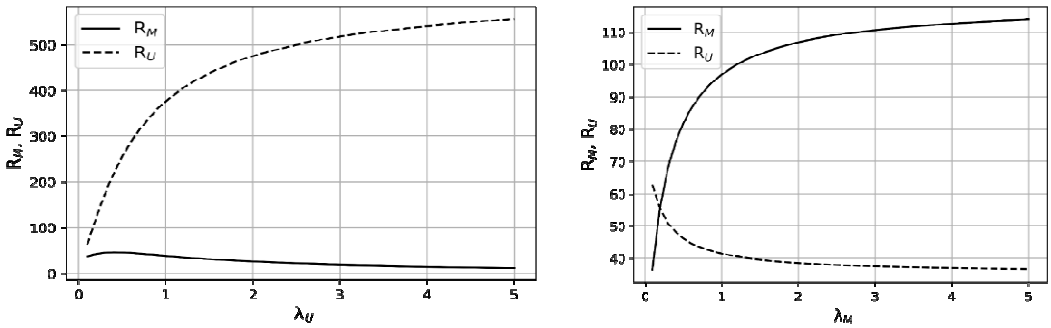


Рис. 4. Среднее количество занятого ресурса в зависимости от интенсивности поступления
а) одноадресных заявок б) многоадресных заявок

На рис. 4 приведены зависимости среднего количество занятого ресурса от интенсивностей поступления одноадресных и многоадресных заявок соответственно. Суммарный средний занятый в системе ресурс можно видеть, как сумму значений занятого ресурса для одноадресных и многоадресных заявок.

При росте интенсивности поступления одноадресных заявок среднее значение суммарного занятого ресурса стремится к $R = 694$ быстрее, чем при росте интенсивности поступления многоадресных заявок. Это объясняется тем, что согласно принципу мультивещания, многоадресным заявкам требуется не более, чем 93 (прямая видимость) или 149 (блокировка прямой видимости) единицы ресурса из 694, имеющихся в системе, а остальной ресурс может быть занят только одноадресными заявками.

Заключение

В настоящей работе рассмотрена математическая модель точки доступа в миллиметровом диапазоне длин волн, обслуживающая комбинацию одноадресных и многоадресных соединений. Для проведения численного анализа была разработана программа на языке программирования Python.

ЛИТЕРАТУРА

1. 3GPP, Release description (Release 15), 3GPP TR 21.915 V0.4.0, December 2018.

2. 3GPP, Study on channel model for frequencies from 0.5 to 100 GHz (Release 14), 3GPP TR 38.901 V14.1.1, July 2017.

3. *Samuylov A., Moltchanov D., Krupko A., Kovalchukov R., Moskaleva F., Gaidamaka Yu.* Performance Analysis of Mixture of Unicast and Multicast Sessions in 5G NR Systems // 10th International Congress on Ultra Modern Telecommunications and Control Systems and Workshops (ICUMT-2018), Moscow, Russia, 2018. – P. 287–293.

4. *Бесчастный В.А., Гайдамака Ю.В.* Модель обслуживания трафика одноадресных и многоадресных соединений высокочастотной сети 5G // Современные информационные технологии и ИТ-образование. – 2019. – № 2.

СТАЦИОНАРНОЕ АСИМПТОТИЧЕСКОЕ РАСПРЕДЕЛЕНИЕ ВЕРОЯТНОСТЕЙ СУММАРНОГО ОБЪЁМА ЗАНЯТЫХ РЕСУРСОВ В БЕСКОНЕЧНОЛИНЕЙНОЙ РЕСУРСНОЙ СМО С ВХОДЯЩИМ MMPP-ПОТОКОМ

А.А. Галилейская, Е.Ю. Лисовская
Томский государственный университет

Введение

Системы массового обслуживания, в которых каждое требование, поступающее в систему, имеет случайный объём, представляют собой модели различных технических устройств, в частности, современных информационных и вычислительных систем, и общий объём всех требований в системе ограничен, как отмечено в [1–4]. В последнее время были исследованы системы массового обслуживания с ограничением на общий объём требований при инверсии обслуживания (дисциплина LIFO) [5,6]. В этом случае можно получить алгоритмы, пригодные для численных расчётов стационарных характеристик.

Тихоненко и его коллеги внесли большой вклад в разработку методов исследования ресурсных систем массового обслуживания. В работах [7,8] авторы рассматривают два класса систем массового обслуживания:

1. Системы, в которых общий объём требований неограничен, а время обслуживания зависит от объёма требования. В системе реализован алгоритм AQM (Active Queue Management). Этот алгоритм предполагает, что каждому входящему в систему требованию может быть отказано в обслуживании, и оно может быть потеряно, даже если в памяти имеется свободное место для его размещения, с вероятностью, зависящей от объёма этого требования и общего объёма других требований, присутствующих в системе на момент его получения. Для описанной системы определяется стационарное распределение количества требований и вероятности потерь.

2. Системы с ограниченным общим объёмом, в которых объём требований и время их обслуживания являются независимыми. Для этой системы определяется стационарное распределение количества требований и вероятность потери спроса.

Методы теории массового обслуживания широко используются для описания процесса передачи информации. Кроме того, из-за неоднородности предоставляемых услуг (телефонные звонки, отправка текстовых сообщений, видео и аудио сообщений, использование Интернета) необходимо учитывать объём передаваемой информации [9–12]. В связи с этим актуальна разработка новых ресурсных моделей, сформулированных в терминах систем массового обслуживания, которые позволили бы оценить объём используемого ресурса.

1. Постановка задачи

Рассмотрим систему массового обслуживания $\mathrm{MMPP}^{(\nu)}|\mathrm{GI}|\infty$ с копированием заявки на второй фазе, на вход которой поступает MMPP-поток заявок, управляемый цепью Маркова $k(t)=1,2,...,K$ и задаваемый матрицей инфинитезимальных характеристик $\mathbf{Q}=\|q_{\nu k}\|$, $\nu,k=\overline{1,K}$ и диагональной матрицей условных интенсивностей $\mathbf{\Lambda}=\mathrm{diag}\{\lambda_1,...,\lambda_K\}$ (см. рис. 1).

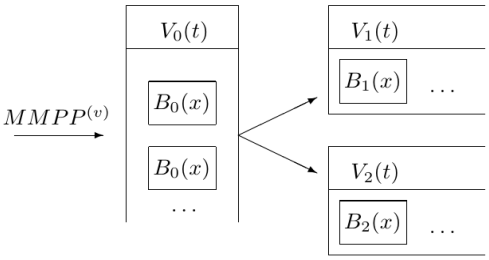


Рис. 1. Бесконечнолинейная ресурсная СМО с *n* типами неоднородных приборов и входящим MMPP-потоком

Дисциплина обслуживания определяется следующим образом. Заявка, поступившая в систему, занимает любой свободный прибор на первой фазе и требует для обслуживания случайное количество некоторого ресурса $G_0(\nu)$ и обслуживается в течение случайного времени, имеющего функцию распределения $B_0(x)$. После обслуживания на первой фазе заявка копируется на вторую и обслуживается на первом блоке в течение случайного времени с функцией распределения $B_1(x)$ и на втором блоке – с $B_2(x)$. Также заявка требует случайное количество некоторого ресурса на первом блоке с функцией распределения $G_1(\nu)$ и на втором с $G_2(\nu)$. По окончании обслуживания заявка покидает систему, освобождая все занятые приборы и ресурсы. Количество занимаемых ресурсов и время обслуживания не зависят друг от друга.

Пусть $V_i(t)$ – суммарный объём занятого ресурса в системе в момент времени *t*, $i=\overline{0,2}$. Поставим задачу нахождения характеристик многомерного случайного процесса $\{V_0(t),V_1(t),V_2(t)\}$. Отметим, что исследуемый процесс не является марковским. Для его исследования применим метод многомерного динамического просеивания [13].

Изобразим четыре параллельных оси времени (рис. 2). Ось под названием MMPP будет отображать события входящего потока, ось под номером 0 будет соответствовать первому просеянному потоку, ось под номером 1 – второму, ось под номером 2, соответственно, третьему просеянному потоку. Зафиксируем произвольный момент времени *T*.

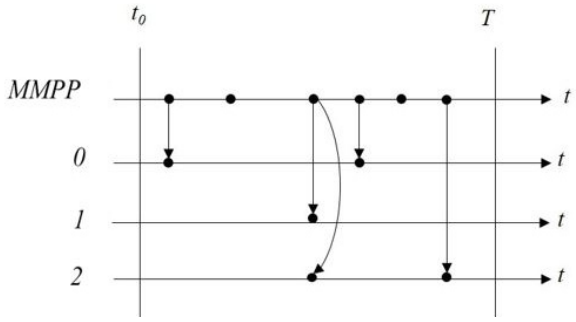


Рис. 2. Просеивание заявок входящего потока

Пусть имеется набор функций $S_0(t)$, $S_1(t)$, $S_2(t)$, $S_{12}(t)$, значения которых лежат в диапазоне $[0,1]$ и обладают свойством $S_0(t)+S_1(t)+S_2(t)+S_{12}(t)\le 1$ для любых *t*.

Событие входящего потока может просеяться только на одну из трёх осей 0, 1, 2, либо ни на одну. Вероятность того, что заявка входящего потока, поступившая в систему в момент времени $t>t_0$,

- сформирует событие потока на оси 0, т.е. к моменту времени T не закончит обслуживание на первой фазе, равна

$$S_0(t) = 1 - B_0(T-t);$$

- сформирует событие потока на оси 1, т.е. к моменту времени T закончит обслуживание на первой фазе и не закончит на первом блоке второй фазы, равна

$$S_1(t) = (B_2 * B_0)(T-t) - \int_0^{T-t} B_1(T-t-x)\,B_2(T-t-x)\,dB_0(x);$$

- сформирует событие потока на оси 2, т.е. к моменту времени T закончит обслуживание на первой фазе и не закончит на втором блоке второй фазы, равна

$$S_2(t) = (B_1 * B_0)(T-t) - \int_0^{T-t} B_1(T-t-x)\,B_2(T-t-x)\,dB_0(x);$$

- сформирует события потока на осях 1 и 2, т.е. к моменту времени T закончит обслуживание на первой фазе, и будет находиться на двух блоках второй фазы одновременно, равна

$$S_{12}(t) = B_0(T-t) - (B_1 * B_0)(T-t) - (B_2 * B_0)(T-t) + \int_0^{T-t} B_1(T-t-x)\,B_2(T-t-x)\,dB_0(x);$$

Обозначим $W_i(t)$ – суммарный объём занятого ресурса просеянными заявками на i-й оси ($i = \overline{0,2}$). Нетрудно показать, что

$$P\{V_0(T) < z_0, V_1(T) < z_1, V_2(T) < z_2\} = P\{W_0(T) < z_0, W_1(T) < z_1, W_2(T) < z_2\} \qquad (1)$$

для любых $z_0 > 0$, $z_1 > 0$, $z_2 > 0$. Будем использовать равенство (1) для исследования процесса $\mathbf{V}(t)$ с помощью процесса $\mathbf{W}(t)$.

2. Система дифференциальных уравнений Колмогорова

Добавим компоненту $k(t)$ – состояние управляющей цепи Маркова в момент времени t, к процессу $\{W_0(t), W_1(t), W_2(t)\}$, тогда полученный трёхмерный процесс будет являться марковским. Введём обозначение для его распределения вероятностей:

$$P(k, z_0, z_1, z_2, t) = P\{k(t) = k, W_0(t) < z_0, W_1(t) < z_1, W_2(t) < z_2\}.$$

Для этого распределения составим Δt-методом прямую систему дифференциальных уравнений Колмогорова. По формуле полной вероятности запишем:

$$\begin{aligned} P(k, z_0, z_1, z_2, t+\Delta t) &= P(k, z_0, z_1, z_2, t)(1-\lambda_k \Delta t)(1+q_{kk}\Delta t) + \\ &+ \lambda_k \Delta t S_0(t) \int_0^{z_0} P(k, z_0 - y, z_1, z_2, t)\,dG_0(y) + \lambda_k \Delta t S_1(t) \int_0^{z_1} P(k, z_0, z_1 - y, z_2, t)\,dG_1(y) + \\ &+ \lambda_k \Delta t S_2(t) \int_0^{z_2} P(k, z_0, z_1, z_2 - y, t)\,dG_2(y) + \\ &+ \lambda_k \Delta t \left(1 - S_0(t) - S_1(t) - S_2(t) - S_{12}(t)\right) P(k, z_0, z_1, z_2, t) + \\ &+ \lambda_k \Delta t S_{12}(t) \int_0^{z_1}\int_0^{z_2} P(k, z_0, z_1 - y_1, z_2 - y_2, t)\,dG_1(y_1)\,dG_2(y_2) + \sum_{v \neq k} q_{vk} \Delta t P(v, z_0, z_1, z_2, t) + o(\Delta t). \end{aligned} \qquad (2)$$

Из (2) получаем систему дифференциальных уравнений Колмогорова

$$\frac{\partial P(k,z_0,z_1,z_2,t)}{\partial t}=\lambda_k S_0(t)\left[\int_0^{z_0}P(k,z_0-y,z_1,z_2,t)dG_0(y)-P(k,z_0,z_1,z_2,t)\right]+$$
$$+\lambda_k S_1(t)\left[\int_0^{z_1}P(k,z_0,z_1-y,z_2,t)dG_1(y)-P(k,z_0,z_1,z_2,t)\right]+$$
$$+\lambda_k S_2(t)\left[\int_0^{z_2}P(k,z_0,z_1,z_2-y,t)dG_2(y)-P(k,z_0,z_1,z_2,t)\right]+$$
$$+\lambda_k S_{12}(t)\left[\int_0^{z_1}\int_0^{z_2}P(k,z_0,z_1-y_1,z_2-y_2,t)\,dG_1(y_1)\,dG_2(y_2)-P(k,z_0,z_1,z_2,t)\right]+$$
$$+\sum_{v\neq k}q_{vk}P(v,z_0,z_1,z_2,t),\ k=\overline{1,K},\ z_0>0,\ z_1>0,\ z_2>0.$$

Начальное условие для решения $P(k,z_0,z_1,z_2,t)$ в момент времени t_0 определим в виде

$$P(k,z_0,z_1,z_2,t_0)=\begin{cases}r(k),\ z_0=z_1=z_2=0,\\ 0,\ \text{иначе},\end{cases}$$

где $r(k)$ – стационарное распределение вероятностей состояний цепи Маркова $k(t)$.

Введём частичные характеристические функции вида:

$$h(k,v_0,v_1,v_2,t)=\int_0^{\infty}e^{jv_0z_0}\int_0^{\infty}e^{jv_1z_1}\int_0^{\infty}e^{jv_2z_2}P(k,dz_0,dz_1,dz_2,t),$$

где $j=\sqrt{-1}$ – мнимая единица. Тогда можем записать следующие дифференциальные уравнения:

$$\frac{\partial h(k,v_0,v_1,v_2,t)}{\partial t}=h(k,v_0,v_1,v_2,t)\left[\lambda_k S_0(t)\left(G_0^*(v_0)-1\right)+\lambda_k S_1(t)\left(G_1^*(v_1)-1\right)+\right.$$
$$\left.+\lambda_k S_2(t)\left(G_2^*(v_2)-1\right)+\lambda_k S_{12}(t)\left(G_1^*(v_1)G_2^*(v_2)-1\right)\right],\ k=\overline{1,K},$$
$$G_i^*(v_i)=\int_0^{\infty}e^{jv_iy}dG_i(y).$$

Перепишем эту систему в матричном виде:

$$\frac{\partial \mathbf{h}(v_0,v_1,v_2,t)}{\partial t}=\mathbf{h}(v_0,v_1,v_2,t)\left\{\mathbf{\Lambda}\left[S_0(t)\left(G_0^*(v_0)-1\right)+S_1(t)\left(G_1^*(v_1)-1\right)+\right.\right.$$
$$\left.\left.+S_2(t)\left(G_2^*(v_2)-1\right)+S_{12}(t)\left(G_1^*(v_1)G_2^*(v_2)-1\right)\right.\right. \quad (3)$$

с начальным условием

$$\mathbf{h}(v_0,v_1,v_2,t_0)=\mathbf{r}, \quad (4)$$

где

$$\mathbf{h}(v_0,v_1,v_2,t)=\left[h(1,v_0,v_1,v_2,t),...,h(K,v_0,v_1,v_2,t)\right],$$

$\mathbf{r}=\left[r(1),...,r(K)\right]$ – вектор стационарного распределения вероятностей состояний управляющей цепи Маркова $k(t)$, удовлетворяющий системе:

$$\begin{cases}\mathbf{rQ}=\mathbf{0},\\ \mathbf{re}=1,\end{cases}$$

и **e** – единичный вектор-столбец.

3. Метод асимптотического анализа

Так как прямое решение уравнения (3) не представляется возможным, для решения задачи (3)–(4) воспользуемся методом асимптотического анализа [13] в условии неограниченно растущей интенсивности входящего потока [14] и предельно частых изменений состояний цепи Маркова [15,16].

Подставим в уравнение (3) $\mathbf{\Lambda} = N\overline{\mathbf{\Lambda}}$ и $\mathbf{Q} = N\overline{\mathbf{Q}}$, где $N \to \infty$ – некоторый параметр, который используется для асимптотического анализа. Тогда можно записать:

$$\frac{1}{N}\frac{\partial \mathbf{h}(v_0, v_1, v_2, t)}{\partial t} = \mathbf{h}(v_0, v_1, v_2, t)\left\{\overline{\mathbf{\Lambda}}\left[S_0(t)\left(G_0^*(v_0)-1\right) + S_1(t)\left(G_1^*(v_1)-1\right) + S_2(t)\left(G_2^*(v_2)-1\right) + S_{12}(t)\left(G_1^*(v_1)G_2^*(v_2)-1\right)\right] + \overline{\mathbf{Q}}\right\} \tag{5}$$

с начальным условием (4).

Теорема 1. Асимптотическая характеристическая функция первого порядка трехмерного случайного процесса $\{k(t), W_0(t), W_1(t), W_2(t)\}$ имеет вид

$$\mathbf{h}(v_0, v_1, v_2, t) = \mathbf{r}\exp\left\{N\lambda\left[jv_0a_1^{(0)}\int_{t_0}^{t} S_0(\tau)d\tau + jv_1a_1^{(1)}\int_{t_0}^{t} S_1(\tau)d\tau + jv_2a_1^{(2)}\int_{t_0}^{t} S_2(\tau)d\tau + \left(jv_1a_1^{(1)} + jv_2a_1^{(2)}\right)\int_{t_0}^{t} S_{12}(\tau)d\tau\right]\right\},$$

где $\lambda = \mathbf{r}\overline{\mathbf{\Lambda}}\mathbf{e}$ – средняя интенсивность входящего потока, $a_1^{(i)}$ – математическое ожидание занимаемого ресурса на *i*-м блоке системы.

Доказательство. Обозначим

$$\frac{1}{N} = \varepsilon,\ v_0 = \varepsilon x_0,\ v_1 = \varepsilon x_1,\ v_2 = \varepsilon x_2,\ \mathbf{h}(v_0, v_1, v_2, t) = \mathbf{f}_1(x_0, x_1, x_2, t, \varepsilon). \tag{6}$$

Перепишем задачу (4)–(5) с учётом введенных обозначений:

$$\varepsilon\frac{\partial \mathbf{f}_1(x_0, x_1, x_2, t, \varepsilon)}{\partial t} = \mathbf{f}_1(x_0, x_1, x_2, t, \varepsilon)\left\{\overline{\mathbf{\Lambda}}\left[S_0(t)\left(G_0^*(\varepsilon x_0)-1\right) + S_1(t)\left(G_1^*(\varepsilon x_1)-1\right) + S_2(t)\left(G_2^*(\varepsilon x_2)-1\right) + S_{12}(t)\left(G_1^*(\varepsilon x_1)G_2^*(\varepsilon x_2)-1\right)\right] + \overline{\mathbf{Q}}\right\} \tag{7}$$

с начальным условием

$$\mathbf{f}_1(x_0, x_1, x_2, t_0, \varepsilon) = \mathbf{r}. \tag{8}$$

Найдём асимптотическое решение задачи (7)–(8) в два этапа.

Этап 1. Подставляя в (7) $\varepsilon = 0$, получим $\mathbf{f}_1(x_0, x_1, x_2, t)\overline{\mathbf{Q}} = \mathbf{0}$. Сравнивая это уравнение с первым в системе для **r**, можно сделать вывод, что его решение можно записать в виде

$$\mathbf{f}_1(x_0, x_1, x_2, t) = \mathbf{r}\Phi_1(x_0, x_1, x_2, t), \tag{9}$$

где $\Phi_1(x_0, x_1, x_2, t)$ – некоторая скалярная функция, которая удовлетворяет условию

$$\Phi_1(x_0, x_1, x_2, t_0) = 1. \tag{10}$$

Этап 2. Умножим (7) на вектор **e**, подставим (9), учитывая разложение

$$e^{j\varepsilon x} = 1 + j\varepsilon x + O(\varepsilon^2), \tag{11}$$

разделим результаты на ε и произведем асимптотический переход $\varepsilon \to 0$. Тогда, учитывая, что $\overline{\mathbf{Q}}\mathbf{e} = 0$ и $\mathbf{re} = 1$, для функции $\Phi_1(x_0, x_1, x_2, t)$ получим следующее дифференциальное уравнение:

$$\frac{\partial\Phi_1(x_0,x_1,x_2,t)}{\partial t}=\Phi_1(x_0,x_1,x_2,t)\overline{\Lambda}\Big[S_0(t)\big(G_0^*(\varepsilon x_0)-1\big)+S_1(t)\big(G_1^*(\varepsilon x_1)-1\big)+$$
$$+S_2(t)\big(G_2^*(\varepsilon x_2)-1\big)+S_{12}(t)\big(G_1^*(\varepsilon x_1)G_2^*(\varepsilon x_2)-1\big)\Big]\mathbf{e}. \qquad (12)$$

Проинтегрировав (12) от t_0 до t, учитывая начальное условие (10), получим

$$\Phi_1(x_0,x_1,x_2,t)=\exp\left\{\lambda\left[jx_0a_1^{(0)}\int_{t_0}^{t}S_0(\tau)d\tau+jx_1a_1^{(1)}\int_{t_0}^{t}S_1(\tau)d\tau+\right.\right.$$
$$\left.\left.+jx_2a_1^{(2)}\int_{t_0}^{t}S_2(\tau)d\tau+\left(jx_1a_1^{(1)}+jx_2a_1^{(2)}\right)\int_{t_0}^{t}S_{12}(\tau)d\tau\right]\right\}.$$

Подставляя это выражение в (9) и выполняя замены, обратные к (6), получим

$$\mathbf{h}(v_0,v_1,v_2,t)=\mathbf{f}_1(x_0,x_1,x_2,t,\varepsilon)\approx\mathbf{f}_1(x_0,x_1,x_2,t)=$$
$$=\mathbf{r}\exp\left\{\lambda\left[jx_0a_1^{(0)}\int_{t_0}^{t}S_0(\tau)d\tau+jx_1a_1^{(1)}\int_{t_0}^{t}S_1(\tau)d\tau+jx_2a_1^{(2)}\int_{t_0}^{t}S_2(\tau)d\tau+\left(jx_1a_1^{(1)}+jx_2a_1^{(2)}\right)\int_{t_0}^{t}S_{12}(\tau)d\tau\right]\right\}=$$
$$=\mathbf{r}\exp\left\{N\lambda\left[jv_0a_1^{(0)}\int_{t_0}^{t}S_0(\tau)d\tau+jv_1a_1^{(1)}\int_{t_0}^{t}S_1(\tau)d\tau+\right.\right.$$
$$\left.\left.+jv_2a_1^{(2)}\int_{t_0}^{t}S_2(\tau)d\tau+\left(jv_1a_1^{(1)}+jv_2a_1^{(2)}\right)\int_{t_0}^{t}S_{12}(\tau)d\tau\right]\right\}.$$

Теорема 1 доказана.

Теорема 2. Асимптотическая характеристическая функция второго порядка многомерного случайного процесса $\{k(t),W_0(t),W_1(t),W_2(t)\}$ имеет вид

$$\mathbf{h}(v_0,v_1,v_2,t)\approx\mathbf{r}\exp\left\{N\lambda\left[jv_0a_1^{(0)}\int_{t_0}^{t}S_0(\tau)d\tau+jv_1a_1^{(1)}\int_{t_0}^{t}\big(S_1(\tau)+S_{12}(\tau)\big)d\tau+\right.\right.$$
$$\left.+jv_2a_1^{(2)}\int_{t_0}^{t}\big(S_2(\tau)+S_{12}(\tau)\big)d\tau\right]+\frac{(jv_0)^2}{2}\left[N\lambda a_2^{(0)}\int_{t_0}^{t}S_0(\tau)d\tau+N\kappa\left(a_1^{(0)}\right)^2\int_{t_0}^{t}S_0^2(\tau)d\tau\right]+$$
$$+\frac{(jv_1)^2}{2}\left[N\lambda a_2^{(1)}\int_{t_0}^{t}\big(S_1(\tau)+S_{12}(\tau)\big)d\tau+N\kappa\left(a_1^{(1)}\right)^2\int_{t_0}^{t}\big(S_1(\tau)+S_{12}(\tau)\big)^2d\tau\right]+$$
$$+\frac{(jv_2)^2}{2}\left[N\lambda a_2^{(2)}\int_{t_0}^{t}\big(S_2(\tau)+S_{12}(\tau)\big)d\tau+N\kappa\left(a_1^{(2)}\right)^2\int_{t_0}^{t}\big(S_2(\tau)+S_{12}(\tau)\big)^2d\tau\right]+$$
$$+j^2v_0v_1a_1^{(0)}a_1^{(1)}N\kappa\left[\int_{t_0}^{t}S_0(\tau)\big(S_1(\tau)+S_{12}(\tau)\big)d\tau\right]+$$
$$+j^2v_0v_2a_1^{(0)}a_1^{(2)}N\kappa\left[\int_{t_0}^{t}S_0(\tau)\big(S_2(\tau)+S_{12}(\tau)\big)d\tau\right]+$$
$$\left.\left.+j^2v_1v_2a_1^{(1)}a_1^{(2)}\left[N\lambda\int_{t_0}^{t}S_{12}(\tau)d\tau+N\kappa\int_{t_0}^{t}\big(S_1(\tau)+S_{12}(\tau)\big)\big(S_2(\tau)+S_{12}(\tau)\big)d\tau\right]\right]\right\},$$

где $\kappa = 2\mathbf{g}\left(\overline{\mathbf{\Lambda}} - \lambda \mathrm{I}\right)\mathbf{e}$, $a_1^{(i)}$ и $a_2^{(i)}$ – первый и второй начальные моменты случайных величин с функцией распределения вероятностей $G_i(y)$.

Доказательство. Перейдём к построению гауссовской аппроксимации суммарных объёмов занятых ресурсов на каждом блоке системы. Представим функцию $\mathbf{h}(v_0, v_1, v_2, t)$ в виде

$$\mathbf{h}(v_0, v_1, v_2, t) = \mathbf{h}_2(v_0, v_1, v_2, t)\exp\left\{N\lambda\left[jv_0 a_1^{(0)}\int_{t_0}^{t} S_0(\tau)\,d\tau + jv_1 a_1^{(1)}\int_{t_0}^{t} S_1(\tau)\,d\tau + \right.\right.$$
$$\left.\left. + jv_2 a_1^{(2)}\int_{t_0}^{t} S_2(\tau)\,d\tau + \left(jv_1 a_1^{(1)} + jv_2 a_1^{(2)}\right)\int_{t_0}^{t} S_{12}(\tau)\,d\tau\right]\right\}, \tag{13}$$

получим уравнение относительно функции $\mathbf{h}_2(v_0, v_1, v_2, t)$:

$$\frac{1}{N}\frac{\partial \mathbf{H}_2(\mathbf{u}, \mathbf{v}, t)}{\partial t} + \mathbf{H}_2(\mathbf{u}, \mathbf{v}, t)\lambda\sum_{i=1}^{n} p_i S_i(t)\left(ju_i + jv_i a_1^{(i)}\right) =$$
$$= \mathbf{H}_2(\mathbf{u}, \mathbf{v}, t)\left[\overline{\mathbf{\Lambda}}\sum_{i=1}^{n} p_i S_i(t)\left(e^{ju_i} G_i^*(v_i) - 1\right) + \overline{\mathbf{Q}}\right]. \tag{14}$$

Выполним здесь следующие замены:

$$\frac{1}{N} = \varepsilon^2,\ v_0 = \varepsilon x_0,\ v_1 = \varepsilon x_1,\ v_2 = \varepsilon x_2,\ \mathbf{h}_2(v_0, v_1, v_2, t) = \mathbf{f}_2(x_0, x_1, x_2, t, \varepsilon). \tag{15}$$

С использованием обозначений (15) уравнение (14) перепишется в виде:

$$\varepsilon^2\frac{\partial \mathbf{f}_2(x_0, x_1, x_2, t, \varepsilon)}{\partial t} + \mathbf{f}_2(x_0, x_1, x_2, t, \varepsilon)\lambda\left[j\varepsilon x_0 S_0(t) + j\varepsilon x_1 S_1(t) + j\varepsilon x_2 S_2(t) + \right.$$
$$\left. + \left(j\varepsilon x_1 + j\varepsilon x_2\right) S_{12}(t)\right] = \mathbf{f}_2(x_0, x_1, x_2, t, \varepsilon)\left\{\overline{\mathbf{\Lambda}}\left[\left(G_0^*(\varepsilon x_0) - 1\right) S_0(t) + \right.\right. \tag{16}$$
$$\left.\left. + \left(G_1^*(\varepsilon x_1) - 1\right) S_1(t) + \left(G_2^*(\varepsilon x_2) - 1\right) S_2(t) + \left(G_1^*(\varepsilon x_1) G_2^*(\varepsilon x_2) - 1\right) S_{12}(t)\right] + \overline{\mathbf{Q}}\right\}.$$

Найдём асимптотическое, при $\varepsilon \to 0$, решение этой задачи, т.е. $\mathbf{f}_2(x_0, x_1, x_2, t) = \lim_{\varepsilon \to 0} \mathbf{f}_2(x_0, x_1, x_2, t, \varepsilon)$.

Этап 1. Выполним предельный переход при $\varepsilon \to 0$ в (16), получим

$$\mathbf{f}_2(x_0, x_1, x_2, t, \varepsilon)\overline{\mathbf{Q}} = 0.$$

Представим $\mathbf{f}_2(x_0, x_1, x_2, t)$ в виде

$$\mathbf{f}_2(x_0, x_1, x_2, t) = \mathbf{r}\Phi_2(x_0, x_1, x_2, t), \tag{17}$$

где $\Phi_2(x_0, x_1, x_2, t)$ – некоторая скалярная дифференцируемая функция, удовлетворяющая условию $\Phi_2(x_0, x_1, x_2, t_0) = 1$.

Этап 2. Решение уравнения (16) запишем в виде разложения

$$\mathbf{f}_2(x_0, x_1, x_2, t) = \Phi_2(x_0, x_1, x_2, t)\left\{\mathbf{r} + \mathbf{g}\left[j\varepsilon x_0 S_0(t) + j\varepsilon x_1 S_1(t) + j\varepsilon x_2 S_2(t) + \right.\right.$$
$$\left.\left. + \left(j\varepsilon x_1 + j\varepsilon x_2\right) S_{12}(t)\right] + O\left(\varepsilon^2\right)\right\}, \tag{18}$$

где **g** – некоторая вектор-строка. Подставим разложение (18) в (16), используя разложение (11), получим матричное уравнение для вектора **g**: $\mathbf{g}\overline{\mathbf{Q}} = \mathbf{r}\left(\lambda\mathbf{I} - \overline{\mathbf{\Lambda}}\right)$.

Этап 3. Домножим (16) на вектор **e**, используя (18) и разложение $e^{j\varepsilon x} = 1 + j\varepsilon x + \frac{(j\varepsilon x)^2}{2} + O(\varepsilon^3)$, в результате несложных преобразований получаем

$$\frac{\partial \Phi_2(x_0,x_1,x_2,t)}{\partial t} = \Phi_2(x_0,x_1,x_2,t)\left[\frac{(jx_0)^2}{2}\left(\lambda a_2^{(0)} S_0(t) + \kappa\left(a_2^{(0)}\right)^2 S_0^2(t)\right) + \right.$$

$$+\frac{(jx_1)^2}{2}\left[\lambda a_2^{(1)}\left(S_1(t)+S_{12}(t)\right) + \kappa\left(a_1^{(1)}\right)^2\left(S_1(t)+S_{12}(t)\right)^2\right] +$$

$$+\frac{(jx_2)^2}{2}\left[\lambda a_2^{(2)}\left(S_2(t)+S_{12}(t)\right) + \kappa\left(a_1^{(2)}\right)^2\left(S_2(t)+S_{12}(t)\right)^2\right] +$$

$$+j^2 x_0 x_1 a_1^{(0)} a_1^{(1)} \kappa\left[S_0(t)\left(S_1(t)+S_{12}(t)\right)\right] + j^2 x_0 x_2 a_1^{(0)} a_1^{(2)} \kappa\left[S_0(t)\left(S_2(t)+S_{12}(t)\right)\right] +$$

$$\left. +j^2 x_1 x_2 a_1^{(1)} a_1^{(2)}\left[\lambda S_{12}(t) + \kappa\left(S_{12}^2(t) + S_1(t)S_2(t)S_1(t)S_{12}(t) + S_2(t)S_{12}(t)\right)\right]\right],$$

где $\kappa = 2\mathbf{g}\left(\bar{\mathbf{\Lambda}} - \lambda \mathrm{I}\right)\mathbf{e}$.

Решение этого уравнения, с учётом начального условия, имеет вид:

$$\Phi_2(\mathbf{x},\mathbf{y},t) = \exp\left\{\frac{(jx_0)^2}{2}\left(\lambda a_2^{(0)}\int_{t_0}^{t} S_0(\tau)d\tau + \kappa\left(a_2^{(0)}\right)^2\int_{t_0}^{t} S_0^2(\tau)d\tau\right) + \right.$$

$$+\frac{(jx_1)^2}{2}\left[\lambda a_2^{(1)}\left(\int_{t_0}^{t} S_1(\tau)d\tau + \int_{t_0}^{t} S_{12}(\tau)d\tau\right) + \kappa\left(a_1^{(1)}\right)^2\left(\int_{t_0}^{t} S_1(\tau)d\tau + \int_{t_0}^{t} S_{12}(\tau)d\tau\right)^2\right] +$$

$$+\frac{(jx_2)^2}{2}\left[\lambda a_2^{(2)}\left(\int_{t_0}^{t} S_2(\tau)d\tau + \int_{t_0}^{t} S_{12}(\tau)d\tau\right) + \kappa\left(a_1^{(2)}\right)^2\left(\int_{t_0}^{t} S_2(\tau)d\tau + \int_{t_0}^{t} S_{12}(\tau)d\tau\right)^2\right] +$$

$$+j^2 x_0 x_1 a_1^{(0)} a_1^{(1)} \kappa\left[\int_{t_0}^{t} S_0(\tau)d\tau\left(\int_{t_0}^{t} S_1(\tau)d\tau + \int_{t_0}^{t} S_{12}(\tau)d\tau\right)\right] +$$

$$+j^2 x_0 x_2 a_1^{(0)} a_1^{(2)} \kappa\left[\int_{t_0}^{t} S_0(\tau)d\tau\left(\int_{t_0}^{t} S_2(\tau)d\tau + \int_{t_0}^{t} S_{12}(\tau)d\tau\right)\right] + j^2 x_1 x_2 a_1^{(1)} a_1^{(2)}\left[\lambda\int_{t_0}^{t} S_{12}(\tau)d\tau + \right.$$

$$\left. \left. +\kappa\left(\int_{t_0}^{t} S_{12}^2(\tau)d\tau + \int_{t_0}^{t} S_1(\tau)S_2(\tau)d\tau + \int_{t_0}^{t} S_1(\tau)S_{12}(\tau)d\tau + \int_{t_0}^{t} S_2(\tau)S_{12}(\tau)d\tau\right)\right],$$

подставляя которое в (17), получаем

$$\mathbf{f}_2(x_0,x_1,x_2,t) = \mathbf{r}\exp\left\{\frac{(jx_0)^2}{2}\left(\lambda a_2^{(0)}\int_{t_0}^{t} S_0(\tau)d\tau + \kappa\left(a_2^{(0)}\right)^2\int_{t_0}^{t} S_0^2(\tau)d\tau\right) + \right.$$

$$+\frac{(jx_1)^2}{2}\left[\lambda a_2^{(1)}\left(\int_{t_0}^{t} S_1(\tau)d\tau + \int_{t_0}^{t} S_{12}(\tau)d\tau\right) + \kappa\left(a_1^{(1)}\right)^2\left(\int_{t_0}^{t} S_1(\tau)d\tau + \int_{t_0}^{t} S_{12}(\tau)d\tau\right)^2\right] +$$

$$+\frac{(jx_2)^2}{2}\left[\lambda a_2^{(2)}\left(\int_{t_0}^{t} S_2(\tau)d\tau + \int_{t_0}^{t} S_{12}(\tau)d\tau\right) + \kappa\left(a_1^{(2)}\right)^2\left(\int_{t_0}^{t} S_2(\tau)d\tau + \int_{t_0}^{t} S_{12}(\tau)d\tau\right)^2\right] +$$

$$+j^2x_0x_1a_1^{(0)}a_1^{(1)}\kappa\left[\int_{t_0}^{t}S_0(\tau)d\tau\left(\int_{t_0}^{t}S_1(\tau)d\tau+\int_{t_0}^{t}S_{12}(\tau)d\tau\right)\right]+$$

$$+j^2x_0x_2a_1^{(0)}a_1^{(2)}\kappa\left[\int_{t_0}^{t}S_0(\tau)d\tau\left(\int_{t_0}^{t}S_2(\tau)d\tau+\int_{t_0}^{t}S_{12}(\tau)d\tau\right)\right]+j^2x_1x_2a_1^{(1)}a_1^{(2)}\left[\lambda\int_{t_0}^{t}S_{12}(\tau)d\tau+\right.$$

$$\left.+\kappa\left(\int_{t_0}^{t}S_{12}^2(\tau)d\tau+\int_{t_0}^{t}S_1(\tau)S_2(\tau)d\tau+\int_{t_0}^{t}S_1(\tau)S_{12}(\tau)d\tau+\int_{t_0}^{t}S_2(\tau)S_{12}(\tau)d\tau\right)\right].$$

Выполним замены, обратные к (13) и (15), запишем приближённое равенство для характеристической функции $\mathbf{h}(v_0,v_1,v_2,t)$:

$$\mathbf{h}(v_0,v_1,v_2,t)\approx\mathbf{r}\exp\left\{N\lambda\left[jv_0a_1^{(0)}\int_{t_0}^{t}S_0(\tau)d\tau+jv_1a_1^{(1)}\int_{t_0}^{t}\left(S_1(\tau)+S_{12}(\tau)\right)d\tau+\right.\right.$$

$$\left.+jv_2a_1^{(2)}\int_{t_0}^{t}\left(S_2(\tau)+S_{12}(\tau)\right)d\tau\right]+\frac{(jv_0)^2}{2}\left[N\lambda a_2^{(0)}\int_{t_0}^{t}S_0(\tau)d\tau+N\kappa\left(a_1^{(0)}\right)^2\int_{t_0}^{t}S_0^2(\tau)d\tau\right]+$$

$$+\frac{(jv_1)^2}{2}\left[N\lambda a_2^{(1)}\int_{t_0}^{t}\left(S_1(\tau)+S_{12}(\tau)\right)d\tau+N\kappa\left(a_1^{(1)}\right)^2\int_{t_0}^{t}\left(S_1(\tau)+S_{12}(\tau)\right)^2d\tau\right]+$$

$$+\frac{(jv_2)^2}{2}\left[N\lambda a_2^{(2)}\int_{t_0}^{t}\left(S_2(\tau)+S_{12}(\tau)\right)d\tau+N\kappa\left(a_1^{(2)}\right)^2\int_{t_0}^{t}\left(S_2(\tau)+S_{12}(\tau)\right)^2d\tau\right]+$$

$$+j^2v_0v_1a_1^{(0)}a_1^{(1)}N\kappa\left[\int_{t_0}^{t}S_0(\tau)\left(S_1(\tau)+S_{12}(\tau)\right)d\tau\right]+j^2v_0v_2a_1^{(0)}a_1^{(2)}N\kappa\left[\int_{t_0}^{t}S_0(\tau)\left(S_2(\tau)+S_{12}(\tau)\right)d\tau\right]+$$

$$\left.\left.+j^2v_1v_2a_1^{(1)}a_1^{(2)}\left[N\lambda\int_{t_0}^{t}S_{12}(\tau)d\tau+N\kappa\int_{t_0}^{t}\left(S_1(\tau)+S_{12}(\tau)\right)\left(S_2(\tau)+S_{12}(\tau)\right)d\tau\right]\right]\right\}. \qquad (19)$$

Теорема 2 доказана.

Следствие. Полагая в (19) $t=T$, $t_0\to-\infty$, учитывая (1), получим асимптотическую характеристическую функцию стационарного трёхмерного распределения вероятностей суммарного объёма занятых ресурсов на каждой фазе системы, которая совпадает с характеристической функцией трёхмерного гауссовского распределения с вектором математического ожидания $\mathbf{a}=N\lambda\left[a_1^{(0)}b_0\quad a_1^{(1)}b_1\quad a_1^{(2)}b_2\right]$ и ковариационной матрицей

$$\mathbf{K}=N\begin{bmatrix}\lambda a_2^{(0)}b_0+\kappa\left(a_1^{(0)}\right)^2\beta_0 & \kappa a_1^{(0)}a_1^{(1)}\beta_{01} & \kappa a_1^{(0)}a_1^{(2)}\beta_{02}\\ \kappa a_1^{(0)}a_1^{(1)}\beta_{01} & \lambda a_2^{(1)}b_1+\kappa\left(a_1^{(1)}\right)^2\beta_1 & \lambda a_1^{(1)}a_1^{(2)}b_{12}+\kappa a_1^{(1)}a_1^{(2)}\beta_{12}\\ \kappa a_1^{(0)}a_1^{(2)}\beta_{02} & \lambda a_1^{(1)}a_1^{(2)}b_{12}+\kappa a_1^{(1)}a_1^{(2)}\beta_{12} & \lambda a_2^{(2)}b_2+\kappa\left(a_1^{(2)}\right)^2\beta_2\end{bmatrix},$$

где

$$b_0=\int_0^{\infty}\left(1-B_0(\tau)\right)d\tau,\ \beta_0=\int_0^{\infty}\left(1-B_0(\tau)\right)^2d\tau,$$

$$b_1=\int_0^{\infty}\left(B_0(\tau)-\left(B_1*B_0\right)(\tau)\right)d\tau,\ \beta_1=\int_0^{\infty}\left(B_0(\tau)-\left(B_1*B_0\right)(\tau)\right)^2d\tau,$$

$$b_2 = \int_0^\infty \left(B_0(\tau) - (B_2 * B_0)(\tau)\right) d\tau,\ \beta_2 = \int_0^\infty \left(B_0(\tau) - (B_2 * B_0)(\tau)\right)^2 d\tau,$$

$$\beta_{01} = \int_0^\infty \left(1 - B_0(\tau)\right)\left(B_0(\tau) - (B_1 * B_0)(\tau)\right) d\tau,\ \beta_{02} = \int_0^\infty \left(1 - B_0(\tau)\right)\left(B_0(\tau) - (B_2 * B_0)(\tau)\right) d\tau,$$

$$b_{12} = \int_0^\infty B_0(T-t) - (B_1 * B_0)(\tau) - (B_2 * B_0)(\tau) + \int_0^\tau B_1(\tau - x) B_2(\tau - x) dB_0(x),$$

$$\beta_{12} = \int_0^\infty \left(B_0(\tau) - (B_1 * B_0)(\tau)\right)\left(B_0(\tau) - (B_2 * B_0)(\tau)\right) d\tau.$$

4. Численный анализ точности асимптотических результатов

Пусть входящий MMPP-поток задан следующими матрицами

$$\mathbf{\Lambda} = N \cdot \begin{bmatrix} 0.5 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1.5 \end{bmatrix},\ \mathbf{Q} = N \cdot \begin{bmatrix} -0.8 & 0,4 & 0.4 \\ 0.3 & -0.6 & 0.3 \\ 0.4 & 0.4 & -0.8 \end{bmatrix},$$

при $N \to \infty$ получаем асимптотическое условие растущей интенсивности входящего потока. Время обслуживания имеет гамма-распределение с параметрами $\alpha_0 = \beta_0 = 0.5$, $\alpha_1 = \beta_1 = 1.5$, $\alpha_2 = \beta_2 = 2.5$. Ресурсы имеют равномерное распределение в диапазоне [0;3] для первой фазы, [0;2] для первого блока второй фазы, [0;1] для второго блока второй фазы.

Целью является найти нижнюю границу параметра *N* применимости приближения (19). Для этого мы провели серию экспериментов, увеличивая значения *N*, и сравнили асимптотические распределения с эмпирическими, используя расстояние Колмогорова

$$\Delta = \max_x \left| G(x) - F(x) \right|,$$

где $G(x)$ – асимптотическая гауссовская функция распределения с параметрами, приведенными в табл. 1, а $F(x)$ – эмпирическая.

Таблица 1

Параметры гауссовской аппроксимации

Фаза системы	Математическое ожидание	Дисперсия
Первая	1,5*N*	3,204*N*
Первый блок второй фазы	1*N*	1,4*N*
Второй блок второй фазы	0,5*N*	0,352

В табл. 2–4 приведены значения расстояния Колмогорова между асимптотической и эмпирической функциями распределений суммарных объёмов занятых ресурсов на трех блоках системы.

Таблица 2

Расстояние Колмогорова для первой фазы системы

N	1	3	5	7	10	20	50	100
Δ	0.374	0.102	0.056	0.040	0.030	0.020	0.013	0.009

Таблица 3

Расстояние Колмогорова для первого блока второй фазы

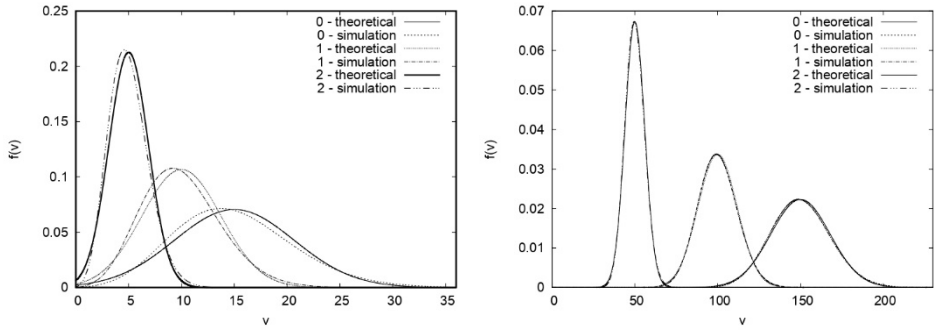
N	1	3	5	7	10	20	50	100
Δ	0.372	0.118	0.065	0.044	0.031	0.019	0.012	0.008

Таблица 4

Расстояние Колмогорова для второго блока второй фазы

N	1	3	5	7	10	20	50	100
Δ	0.367	0.115	0.062	0.042	0.028	0.018	0.011	0.008

Точность аппроксимации увеличивается с ростом интенсивности входящего потока N, а рис. 3 демонстрирует это.



а) N = 35 б) N = 200

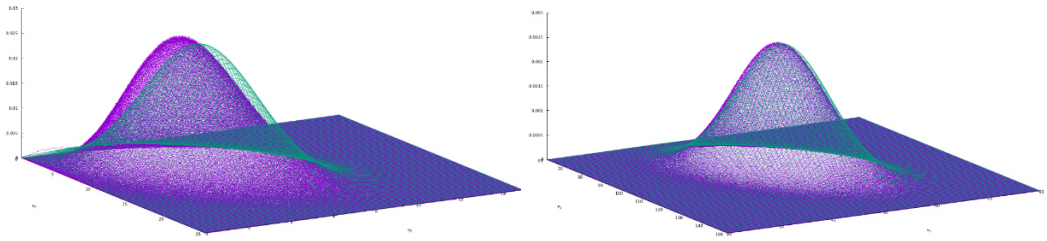
Рис. 3. Распределение суммарного объёма занятого ресурса для каждой фазы системы

Аналогичный вывод можно сделать и для совместного распределения, как показано в табл. 5 и на рис. 4.

Таблица 5

Расстояние Колмогорова для первого и второго блоков второй фазы системы

N	1	3	5	7	10	20	50	100
Δ	0.374	0.118	0.068	0.046	0.031	0.019	0.012	0.008



а) N = 10 б) N = 100

Рис. 4. Совместное распределение суммарного объёма занятого ресурса для первого и второго блоков второй фазы системы

Заключение

Была исследована бесконечнолинейная ресурсная СМО с неограниченным числом приборов, с копированием заявки на второй фазе системы и с входящим ММРР-потоком. С помощью метода асимптотического анализа показано, что совместное асимптотическое распределение вероятности суммарного объёма занятого ресурса на каждой фазе ситсемы сходится к трёхмерному гауссовскому распределению в асимптотическом условии растущей интенсивности входящего потока.

ЛИТЕРАТУРА

1. *Aleksandrov A.M., Kats B.A.* Serving of the flows of heterogeneous requirements // News of the Academy of Sciences of the USSR. Technical Cybernetics. – 1973. – № 2. – P. 47–53.

2. *Romm E.L., Skitovich V.V.* On a generalization of the erlang problem // Automation and Remote Control. – 1971. – № 32 (6) – P. 1000–1003.

3. *Tikhonenko O.M.* Characterization of the total demand volume in single-line priority queuing systems // Automation and Remote Control. – 1989. – № 2. – P. 106–115.

4. *Tikhonenko O.M.* Distribution of the total volume of messages in a single-line queuing system with group arrival // Automation and Remote Control. – 1985. – № 11. – P. 78–83.

5. *Pechinkin A.V.* System with lifo discipline and a limit on the total volume of requirements // Automation and Remote Control. – 1998. – № 4. – P. 106–116.

6. *Pechinkin A.V.* System MAP/G/1/n with LIFO discipline with interruption and restriction on the total volume of requirements // Automation and Remote Control. – 1999. – № 12. – P. 114–120.

7. *Tikhonenko O.M.* Characterization of limited memory service systems // Automation and Remote Control. – 1997. – №6. – P. 105–110.

8. *Tikhonenko O.M.* A shared processor system with limited memory, controlled by AQM // Automation and Remote Control. – 2015. – №10. – P. 90–105.

9. *Naumov V., Samouylov K., Sopin E., Andreev S.* Two approaches to analyzing dynamic cellular networks with limited resources // 2014 6th International Congress on Ultra Modern Telecommunications and Control Systems and Workshops (ICUMT), IEEE, 2014. – P. 485–488.

10. *Naumov V., Samuilov K., Samuilov A.* On the total amount of resources occupied by the serviced requests // Automation and Remote Control. – 2016. – № 8. – P. 125–135.

11. *Brown L., Gans N., Mandelbaum A., Sakov A., Shen H., Zeltyn S., Zhao L.* Statistical analysis of a telephone call center: A queueing-science perspective // Journal of the American statistical association. – 2005. – 100 (469). – P. 36–50.

12. *Down D.G., Wu R.* Multi-layered round robin routing for parallel servers // Queueing Systems. – 2006. – № 53 (4). – P. 177–188.

13. *Моисеев А.Н., Назаров А.А.* Бесконечнолинейные системы и сети массового обслуживания. – Томск: Изд-во НТЛ, 2015. – 240 с.

14. *Назаров А.А., Моисеева С.П.* Метод асимптотического анализа в теории массового обслуживания. – Томск: Изд-во НТЛ, 2006. – 112 с.

15. *Лисовская Е.Ю., Моисеева С.П.* Асимптотический анализ системы MMPP|GI|∞ с обслуживанием требований случайного объёма // Математическое и программное обеспечение информационных, технических и экономических систем: материалы IV Международной молодежной научной конференции. Томск, 20–21 мая 2016 г. – Томск: Издательский Дом Томского государственного университета, 2016. – С. 99–104.

16. *Лисовская Е.Ю., Моисеева С.П.* Исследование бесконечнолинейной системы массового обслуживания требований случайного объёма с входящим MMPP-потоком // Информационные технологии и математическое моделирование (ИТММ-2016): материалы XV Международной конференции им. А.Ф. Терпугова (12–16 сентября 2016 г.). Томск: Изд-во Том. ун-та, 2016, Ч. 1. – С. 77–82.

V. КОМПЬЮТЕРНЫЕ НАУКИ, ИНТЕЛЛЕКТУАЛЬНЫЙ АНАЛИЗ И ВИЗУАЛИЗАЦИЯ ДАННЫХ

МЕТОДЫ НЕЧЁТКОЙ ЛОГИКИ В АНАЛИЗЕ ФОРМАЛЬНЫХ ПОНЯТИЙ

В.К. Голышев
Сибирский федеральный университет

Введение

В настоящее время в интеллектуальном анализе данных актуальна задача извлечения знаний из данных в форме концептуальных структур. Для решения этой задачи широко используются методы анализа формальных понятий (АФП) [1,2]. АФП базируется на теории решёток замкнутых множеств (решёток формальных понятий или решёток Галуа). Основная математическая идея АФП – возможность построения полной решётки по любому бинарному отношению и формализация понятия в виде пары (объём и содержание). Решётка позволяет извлекать и представлять знания из данных, выявлять связи между понятиями. Её можно представить с помощью диаграммы Хассе, на которой будут видны связи, либо с помощью множества импликаций [2,3]. В данной работе рассматривается обобщение анализа формальных понятий на основе нечёткой логики [4,5]. Целью работы является разработка программы извлечения формальных нечётких понятий из нечёткого контекста.

1. Основные определения классического анализа формальных понятий

АФП работает с формальными контекстами, информация в которых представлена в объектно-признаковой форме. Формальный контекст – это тройка $K=(G,M,I)$, где G – множество объектов, M – множество признаков, а $I: G\times M\to\{0,1\}$ – отношение инцидентности, такое, что $(g,m)\in I$ тогда и только тогда, когда объект g имеет признак m (или признак m присущ объекту g). Для произвольных подмножеств объектов $A\subseteq G$ и признаков $B\subseteq M$ следующая пара операторов, называемых отображениями Галуа,

$$A'=\{m\in M:\ \forall g\in A\ (g,m)\in I\},\ B'=\{g\in G:\ \forall m\in B\ (g,m)\in I\}$$

задаёт соответствие между частично упорядоченными множествами $(2^G,\subseteq)$ и $(2^M,\subseteq)$ [6]. Здесь A' – это множество признаков, которыми обладают все объекты из A; B' – это множество объектов, которые обладают всеми признаками из B. Двойное применение отображения «'» определяет оператор замыкания «"» на $(2^G,\subseteq)$ или $(2^M,\subseteq)$ в алгебраическом смысле. Если $A=A''$ ($B=B''$), то множество объектов $A\subseteq G$ (признаков $B\subseteq M$) называется замкнутым относительно оператора «"» в контексте $K=(G,M,I)$. Формальное понятие строится на замыканиях множеств объектов и признаков. Под формальным понятием контекста K понимают пару (A,B), где $A\subseteq G$, $B\subseteq M$, причём $A=B'$, $B=A'$, где A – объём, а B – содержание понятия [1–3]. Таким образом, понятие – это пара из объёма и содержания, когда каждый объект из множества объектов, составляющих объём понятия, характеризуется всей совокупностью свойств, определяющих содержание понятия, и наоборот. Понятия иерархически упорядочены отношением «быть более общим понятием», которое задаётся следующим образом [1–3]: $(A,B)\geq(C,D)\ \Leftrightarrow\ A\supseteq C$.

Частично упорядоченное по вложению объёмов множество формальных понятий контекста *K* образует решётку формальных понятий [1–3]. Решётка понятий, построенная на формальном контексте, является инструментом представления и извлечения знаний из данных контекста. В роли знаний выступают понятия, организованные иерархично. Задача извлечения всех формальных понятий контекста имеет высокую вычислительную сложность. Это связано с тем, что число возможных понятий может быть экспоненциальным от размера контекста. Наиболее часто для решения этой задачи используются алгоритмы Гантера и Замыкай-по-одному, а также их модификации [2,7].

Пример 1. Для контекста (табл. 1, 2), описывающего сказочных персонажей, решётка состоит из 14 формальных понятий (рис. 1).

Таблица 1

Сказочные персонажи

	С	**Л**	**И**	**ОН**
АП	1	0	0	0
ДН	1	0	1	0
ИМ	1	1	1	0
СР	0	1	1	1
КБ	0	0	1	1
ЗГ	1	0	0	1
ТЗ	1	1	0	0
ЖП	0	1	0	1

Таблица 2

Расшифровка обозначений в табл. 1

G		***M***	
АП	Алеша Попович	**С**	Сила
ДН	Добрыня Никитич	**Л**	Ловкость
ИМ	Илья Муромец	**И**	Интеллект
СР	Соловей-разбойник	**ОН**	Особые навыки
КБ	Кощей Бессмертный		
ЗГ	Змей Горыныч		
ТЗ	Тугарин Змей		
ЖП	Жар-птица		

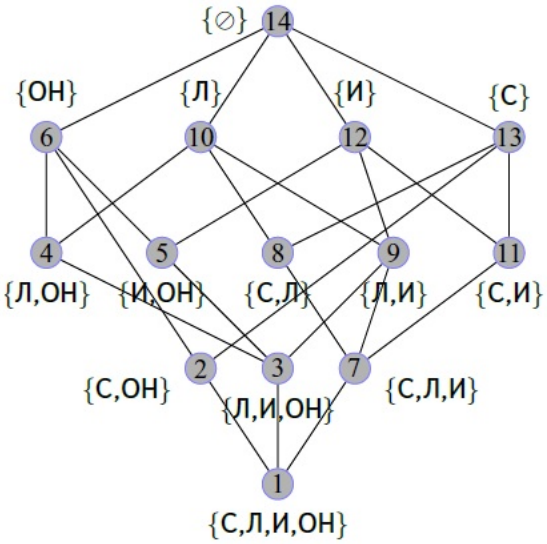


Рис. 1. Решётка из 14 формальных понятий для примера 1

2. Нечёткий анализ формальных понятий

Классический анализ формальных понятий работает с бинарными контекстами в виде (0,1)-матриц. На практике при измерении свойств объектов возникают ситуации, при которых оценка истинности утверждения «объект *g* обладает признаком *m*» является расплывчатой. Для оценивания таких суждений привлекают математический аппарат нечёткой логики [8]. Учёт факторов неопределённости и неполноты информации приводит к нечётким контекстам. Под формальным нечётким контекстом будем понимать тройку $\tilde{K} = \left(G, M, \tilde{I}\right)$, где G – множество объектов, M – множество признаков, а $\tilde{I} : G \times M \to [0,1]$ – отношение инцидентности, определяющее степень принадлежности признака *m* объекту *g*. По сути, $\tilde{I}$ задаёт нечёткое соответствие на множествах G и M. Методы построения нечётких контекстов рассмотрены в работе [8]. В настоящее время существует три подхода к обнаружению понятий в нечётком контексте.

– Метод α-сечения с построением четких понятий [8].

– Метод α-сечения с построением нечётких понятий [8–10], когда формальный контекст интерпретируется как совокупность нечётких множеств свойств, каждое из которых описывает один из объектов нечёткого формального контекста. Именно это одностороннее предпочтение объектам используется при конструировании решёток нечётких понятий, которые можно считать «скелетами» нечётких онтологий. Теоретически возможен и альтернативный взгляд, когда предпочтение отдается свойствам (отсюда ясно второе название метода – схема асимметричного порога).

– Подход, использующий оператор замыкания нечёткого множества [11].

В рамках данной работы рассматривается последний метод. Введём некоторые понятия, необходимые для дальнейшего изложения.

Полная резидуальная решётка с замыканием – это алгебра $\mathbf{L} = \langle L, \wedge, \vee, \otimes, \to, *, 0, 1\rangle$, такая, что (см. [11]):

1. $\langle L, \wedge, \vee, 0, 1\rangle$ – полная решётка L с наименьшим элементом 0 и наибольшим элементом 1.
2. $\langle L, \otimes, 1\rangle$ – коммутативный моноид.
3. $\otimes$ и $\to$ удовлетворяют свойству сопряженности: $a \otimes b \le c \Leftrightarrow a \le b \to c$ $\forall a, b, c \in L$.
4. Для всех $a, b \in L$ замыкание * удовлетворяет следующим свойствам: $1^* = 1$; $a^* \le a$; $(a \to b)^* \le a^* \to b^*$; $a^{**} = a^*$.

Элементы $a \in L$ называют степенями принадлежности, операции $\otimes$ и $\to$ есть «нечёткая конъюкция» и «нечёткая импликация». Три наиболее важные пары сопряженных операций на [0,1]:

– операции Лукасевича:

$$a \otimes b := \max(0, a + b - 1),\ a \to b := \min(1, 1 - a + b); \tag{1}$$

– операции Гёделя:

$$a \otimes b := \min(a, b),\ a \to b := \begin{cases} 1, & a \le b, \\ 0, & a > b; \end{cases} \tag{2}$$

– операции произведения:

$$a \otimes b := ab,\ a \to b := \begin{cases} 1, & a \le b, \\ b / a, & a > b. \end{cases} \tag{3}$$

В работе рассматриваются два типа операции замыкания [4,11]: идентичность и глобализация. Отличие этих операций заключается в детализации свойства $a^* \le a$: для

идентичности оно обращается в равенство $a^* = a$, $a \in L$, а для глобализации имеет следующий вид: $a^* = 1$, если $a = 1$, и $a^* = 0$ во всех других случаях.

Пусть **L** – структура степеней истинности [11]. Нечёткое множество A в универсальном множестве U – это отображение $A\colon U \to L$, где функция $A(u)$ интерпретируется как степень принадлежности элемента u нечёткому множеству A. Для $U = \{u_1, \ldots, u_n\}$ нечёткое множество A есть совокупность пар вида $\{(u_i, a_i) \mid u_i \in U, a_i \in L\}$, где $a_i = A(u_i)$ для каждого $i \in \{1, \ldots, n\}$. Далее будем использовать следующее обозначение: $A = \{a_1 / u_1, \ldots, a_n / u_n\}$.

Пусть $\mathbf{L}^U$ – совокупность всех нечётких множеств в U. Операции с нечёткими множествами определяются покомпонентно [12]. Бинарные нечёткие отношения (**L**-отношения) между G и M можно рассматривать как нечёткие множества в универсальном множестве $G \times M$.

Для A, B степень включения подмножества A в B определяется [8] как

$$Sb(A, B) \coloneqq \bigwedge_{u \in U} (A(u) \to B(u)). \tag{4}$$

Формула (4) обобщает классическое соотношение включения $\subseteq$: $Sb(A, B)$ представляет степень, по которой A является подмножеством B. В частности, мы пишем $A \subseteq B$ тогда и только тогда, когда $Sb(A, B) = 1$.

Тройка (G, M, I_L) называется формальным нечётким контекстом, если $I_L : G \times M \to L$ – это нечёткое отношение между множествами G и M и L является множеством степеней принадлежности. Элементы из G и M называются объектами и признаками соответственно. Нечёткое отношение I_L присваивает каждому объекту $g \in G$ и каждому признаку $m \in M$ степень принадлежности $I_L(g, m) \in L$ признака m объекту g. С учётом (1)–(3) для произвольных нечётких множеств $A \in \mathbf{L}^G$ и $B \in \mathbf{L}^M$ операторы нечёткого логического вывода определяются следующим образом [4,11]:

$$\begin{aligned} A^{\uparrow}(m) &\coloneqq \bigwedge_{g \in G} \left(A(g)^* \to I_L(g, m) \right), \\ B^{\downarrow}(g) &\coloneqq \bigwedge_{m \in M} \left(B(m)^* \to I_L(g, m) \right), \ g \in G, \ m \in M. \end{aligned} \tag{5}$$

Операторы $\uparrow\downarrow$ формируют нечёткие соответствия Галуа с замыканиями *. Формальное нечёткое понятие – это кортеж $(A, B) \in \mathbf{L}^G \times \mathbf{L}^M$ такой, что $A^{\uparrow} = B$ и $B^{\downarrow} = A$. Здесь A называется (нечётким) объёмом, B – (нечётким) содержанием формального понятия (A, B). Обозначим множество всех нечётких понятий данного контекста (G, M, I) через $B(G^*, M^*, I)$. Понятие (A_1, B_1) называется подпонятием (A_2, B_2), если оно более общее [1,7], т.е. имеет больший объём. Более формально, $(A_1, B_1) \le (A_2, B_2) \Leftrightarrow A_1 \subseteq A_2 (B_1 \supseteq B_2)$. Понятие (A_2, B_2) называется надпонятием (A_1, B_1). Множество всех нечётких понятий, подчиняющихся такому порядку, формирует решётку нечётких понятий [11], которая обозначается $B(G^*, M^*, I) \coloneqq (B(G^*, M^*, I), \le)$.

3. Алгоритм извлечения понятий из нечёткого контекста

В [4] предложен алгоритм извлечения понятий из нечёткого контекста (рис. 2). Данный алгоритм является обобщением классического алгоритма Гантера извлечения формальных понятий из бинарного контекста. Основным результатом данной работы

является программа, реализующая алгоритм, предложенный в [4]. Программа разработана на языке Python и использовалась для проведения вычислительных экспериментов.

Алгоритм 1. Алгоритм извлечения понятий из нечеткого контекста

Вход: $K = (G, M, I)$ - нечеткий контекст

Выход: $\mathfrak{B}(G^*, M^*, I)$ - все нечеткие понятия из контекста K

1: $B := \varnothing^{\uparrow\downarrow}$

2: $\mathfrak{B}(G^*, M^*, I) := \varnothing$

3: **До тех пор пока** $B \neq M$ **выполнять**

4: $B := B^+$ ▷ B^+ - нахождение наименьшего содержания, большего данного B

5: $A := B^{\downarrow}$ ▷ находим объем A, который в паре с содержанием B

6: ▷ является формальным нечетким понятием

7: добавить понятие (A, B) в $\mathfrak{B}(G^*, M^*, I)$

8: **Конец цикла**

Рис. 2. Алгоритм извлечения понятий из нечёткого контекста

Пример 2. Рассмотрим нечёткий формальный контекст (табл. 3), описывающий сказочных персонажей. Множество объектов *G* и множество признаков *M* приведены в табл. 2.

Таблица 3

Сказочные персонажи

	С	**Л**	**И**	**ОН**
АП	1	0.4	0.2	0
ДН	1	0.4	0.6	0
ИМ	1	0.8	0.8	0
СР	0	0.8	0.8	1
КБ	0.2	0.2	0.6	1
ЗГ	1	0.2	0	0.8
ТЗ	1	0.5	0.4	0
ЖП	0.2	0.6	0.2	1

При использовании сопряженных операций Лукасевича (1) и идентичности в качестве замыкания, получим 122 нечётких формальных понятия (табл. 4). Такое большое количество понятий связано с особенностью выбранных операций.

Таблица 4

Нечёткие формальные понятия, полученные с помощью идентичности

№	***A* (объём)**	***B* (содержание)**
1	$\{0.2/_{КБ}, 0.2/_{ЖП}\}$	$\{С, Л, И, ОН\}$
2	$\{0.2/_{КБ}, 0.5/_{ЗГ}, 0.2/_{ЖП}\}$	$\{С, 0.7/_{Л}, 0.5/_{И}, ОН\}$
3	$\{0.2/_{КБ}, 0.8/_{ЗГ}, 0.2/_{ЖП}\}$	$\{С, 0.4/_{Л}, 0.2/_{И}, ОН\}$
.	.	.
.	.	.

№	*A* (объём)	*B* (содержание)
.	.	.
121	$\left\{АП, ДН, ИМ, СР, КБ, {}^{0.8}\!/_{ЗГ}, ТЗ, ЖП\right\}$	$\left\{{}^{0.2}\!/_{Л}, {}^{0.2}\!/_{И}\right\}$
122	$\{АП, ДН, ИМ, СР, КБ, ЗГ, ТЗ, ЖП\}$	$\left\{{}^{0.2}\!/_{Л}\right\}$

Рассмотрим, например, второе формальное понятие из табл. 4 и покажем, как работают операторы нечёткого логического вывода (5) с замыканиями. Здесь объём $A = \left\{{}^{0.2}\!/_{КБ}, {}^{0.5}\!/_{ЗГ}, {}^{0.2}\!/_{ЖП}\right\}$. Найдём содержание $B = A^{\uparrow}$ почленно с помощью формул (1), (5):

$$A^{\uparrow}(C) = \left[A^{*}(АП) \to I_L(АП, C)\right] \wedge \left[A^{*}(ДН) \to I_L(ДН, C)\right] \wedge$$
$$\wedge \left[A^{*}(ИМ) \to I_L(ИМ, C)\right] \wedge \left[A^{*}(СР) \to I_L(СР, C)\right] \wedge$$
$$\wedge \left[A^{*}(КБ) \to I_L(КБ, C)\right] \wedge \left[A^{*}(ЗГ) \to I_L(ЗГ, C)\right] \wedge$$
$$\wedge \left[A^{*}(ТЗ) \to I_L(ТЗ, C)\right] \wedge \left[A^{*}(ЖП) \to I_L(ЖП, C)\right] =$$
$$= [0 \to 1] \wedge [0 \to 1] \wedge [0 \to 1] \wedge [0 \to 0] \wedge [0.2 \to 0.2] \wedge [0.5 \to 1] \wedge [0 \to 1] \wedge [0.2 \to 0.2] \overset{(1)}{=}$$
$$= \min(1, 1 - 0 + 1) \wedge \min(1, 1 - 0 + 1) \wedge \min(1, 1 - 0 + 1) \wedge \min(1, 1 - 0 + 0) \wedge$$
$$\wedge \min(1, 1 - 0.2 + 0.2) \wedge \min(1, 1 - 0.5 + 1) \wedge \min(1, 1 - 0 + 1) \wedge \min(1, 1 - 0.2 + 0.2) =$$
$$= 1 \wedge 1 \wedge 1 \wedge 1 \wedge 1 \wedge 1 \wedge 1 \wedge 1 = 1;$$

$$A^{\uparrow}(Л) = \left[A^{*}(АП) \to I_L(АП, Л)\right] \wedge \left[A^{*}(ДН) \to I_L(ДН, Л)\right] \wedge$$
$$\wedge \left[A^{*}(ИМ) \to I_L(ИМ, Л)\right] \wedge \left[A^{*}(СР) \to I_L(СР, Л)\right] \wedge \left[A^{*}(КБ) \to I_L(КБ, Л)\right] \wedge$$
$$\wedge \left[A^{*}(ЗГ) \to I_L(ЗГ, Л)\right] \wedge \left[A^{*}(ТЗ) \to I_L(ТЗ, Л)\right] \wedge \left[A^{*}(ЖП) \to I_L(ЖП, Л)\right] =$$
$$= [0 \to 0.4] \wedge [0 \to 0.4] \wedge [0 \to 0.8] \wedge [0 \to 0.8] \wedge$$
$$\wedge [0.2 \to 0.2] \wedge [0.5 \to 0.2] \wedge [0 \to 0.5] \wedge [0.2 \to 0.6] \overset{(1)}{=}$$
$$= \min(1, 1 - 0 + 0.4) \wedge \min(1, 1 - 0 + 0.4) \wedge \min(1, 1 - 0 + 0.8) \wedge \min(1, 1 - 0 + 0.8) \wedge$$
$$\wedge \min(1, 1 - 0.2 + 0.2) \wedge \min(1, 1 - 0.5 + 0.2) \wedge \min(1, 1 - 0 + 0.5) \wedge \min(1, 1 - 0.2 + 0.6) =$$
$$= 1 \wedge 1 \wedge 1 \wedge 1 \wedge 0.7 \wedge 1 \wedge 1 \wedge 1 = 0.7.$$

Аналогично для признаков *И* и *ОН*:

$$A^{\uparrow}(И) = [0 \to 0.2] \wedge [0 \to 0.6] \wedge [0 \to 0.8] \wedge [0 \to 0.8] \wedge$$
$$\wedge [0.2 \to 0.6] \wedge [0.5 \to 0] \wedge [0 \to 0.4] \wedge [0.2 \to 0.2] = 1 \wedge 1 \wedge 1 \wedge 1 \wedge 1 \wedge 0.5 \wedge 1 \wedge 1 = 0.5;$$

$$A^{\uparrow}(ОН) = [0 \to 0] \wedge [0 \to 0] \wedge [0 \to 0] \wedge [0 \to 1] \wedge$$
$$\wedge [0.2 \to 1] \wedge [0.5 \to 0.8] \wedge [0 \to 0] \wedge [0.2 \to 1] = 1 \wedge 1 \wedge 1 \wedge 1 \wedge 1 \wedge 1 \wedge 1 \wedge 1 = 1.$$

В итоге имеем $B = A^{\uparrow} = \left\{C, {}^{0.7}\!/_{Л}, {}^{0.5}\!/_{И}, ОН\right\}$. Точно таким же образом можем посчитать $B^{\uparrow}$ и увидим, что $A^{\uparrow} = B$ и $B^{\downarrow} = A$, т.е. (A, B) – нечёткое формальное понятие.

При использовании глобализации вместо идентичности решётка будет содержать всего лишь 4 понятия (табл. 5, рис. 3).

Таблица 5

Нечёткие формальные понятия, полученные с помощью глобализации

№	*A* (объём)	*B* (содержание)
1	$\{0.2/КБ, 0.2/ЖП\}$	$\{С, Л, И, ОН\}$
2	$\{КБ, 0.8/ЗГ, ЖП\}$	$\{0.2/С, 0.2/Л, 0.2/И, ОН\}$
3	$\{АП, ДН, ИМ, 0.2/КБ, ЗГ, ТЗ, 0.2/ЖП\}$	$\{С, 0.2/Л\}$
4	$\{АП, ДН, ИМ, СР, КБ, ЗГ, ТЗ, ЖП\}$	$\{0.2/Л\}$

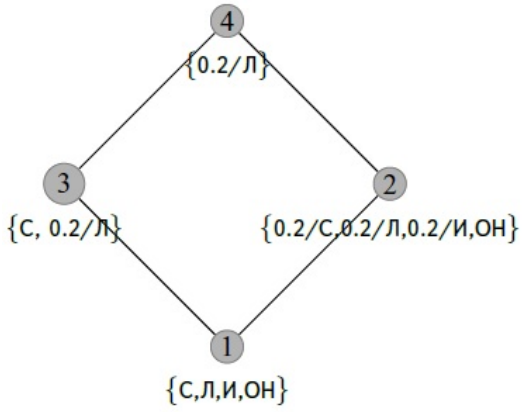


Рис. 3. Решётка из 4 понятий, полученных с помощью глобализации

Заключение

Основным результатом работы является программа, реализующая обобщенный алгоритм Гантера. Поставленная цель была достигнута. Однако из примера видно, что выбор оператора замыкания нечёткого множества и операций нечёткого логического вывода при построении решётки нечётких формальных понятий существенно влияют на вычислительную сложность метода, и с помощью замыканий можно контролировать размер получаемой решётки.

Таким образом, проблема разработки эффективных алгоритмов построения решётки нечётких формальных понятий является актуальной и требует дальнейших исследований.

ЛИТЕРАТУРА

1. *Игнатов Д.И.* Анализ формальных понятий: от теории к практике // В кн.: Анализ изображений, сетей и текстов. Доклады всероссийской научной конференции АИСТ'12. Модели, алгоритмы и инструменты анализа данных; результаты и возможности для анализа изображений, сетей и текстов. Екатеринбург, 16–18 марта 2012 года / Науч. ред.: Д.И. Игнатов, Р.Э. Яворский. Вып. 1. — М.: Национальный открытый университет «ИНТУИТ», 2012. – С. 3–15.

2. *Кузнецов С.О.* Автоматическое обучение на основе анализа формальных понятий // Автоматика и телемеханика. – 2001. – № 10. – С. 3–27.

3. *Ganter B., Wille R.* Formal Concept Analyses: Mathematical Foundations // Springer Science and Business Media. – 2012. – P. 314.

4. *Bělohlávek R., Funioková T.T., Vychodil V.* Galois connections with hedges // Eleventh International Fuzzy Systems Association World Congress. – 2005. – P. 1250–1255.

5. *Bělohlávek R.* Algorithms for fuzzy concepts lattices // Proc. Fourth Int. Conf. on Recent Advances in Soft Computing. – 2002. – P. 200–205.

6. *Биркгоф Г.* Теория решёток. – М.: Наука, 1984. – 568 с.

7. *Незнанов А.А., Строк Ф.В.* Оценка и сравнение производительности алгоритмов анализа формальных понятий // Двенадцатая национальная конференция по искусственному интеллекту с международным участием (20–24 сентября 2010 г., Тверь, Россия): Труды конф. Т. 2. – М.: Физматлит, 2010. – С. 82–89.

8. *Офицеров В.П., Смирнов С.В.* Нечёткий анализ формальных понятий при разработке онтологий // Онтология проектирования. – 2017. – № 4 (26). С. 487–495.

9. *Yang K.M., Kim E.H., Hwang S.H., Choi S.H.* Fuzzy Concept Mining based on Formal Concept Analysis // Int. J. of Computers. – 2008. Iss. 3. – Vol. 2. – P. 279–290.

10. *Tho Q.T., Hui S.C., Fong A.C.M., Cao T.H.* Automatic Fuzzy Ontology Generation for the Semantic Web // IEEE Trans. on Knowledge and Data Engineering. – 2006. № 18(6). – P. 842–856.

11. *Glodeanu C.V.* Attribute Exploration in a Fuzzy Setting // CEUR Workshop Proceedings. – 2012. – P. 114–128.

12. *Пегат А.* Нечёткое моделирование и управление [Электронный ресурс] / А. Пегат; пер. с англ. — 3-е изд. (эл.). — Электрон. текстовые дан. (1 файл pdf : 801 с.). — М.: БИНОМ. Лаборатория знаний, 2015. — (Адаптивные и интеллектуальные системы).

ИССЛЕДОВАНИЕ 2-СБАЛАНСИРОВАННЫХ ЗНАКОВЫХ ГРАФОВ

Э. И. Ибрагимова

Сибирский федеральный университет

Введение

Графы, в которых каждое ребро имеет знак «+» или «–», называются знаковыми. Знаковый граф является сбалансированным, если каждый его простой цикл положительный, где знак цикла определяется как произведение знаков, входящих в него рёбер. Всюду в дальнейшем под сбалансированностью будем иметь в виду 2-сбалансированность.

Данные графы возникли при изучении проблем социальной психологии [1] и нашли применение в анализе данных при бинарной классификации, в топологической теории графов, в моделировании социальных сетей, в портфельном анализе, в управлении рисками и в генных регуляторных сетях, которые являются основанием для нейронных сетей. В [2] рассматривается использование знаковых графов в теории групп, где граф напряжения критической группы двойных покрытий представляется знаковым графом, что облегчает изучение двойных покрытий. В [3] описывается применение знаковых графов в медицине при изучении онкологических заболеваний. В этом случае с помощью знаковых графов моделируются процессы распространения раковых клеток внутри организма и наследственность онкологических заболеваний. Также знаковые графы используются при вычислении энергии основного состояния в неферромагнитной модели Изинга (модель Изинга – математическая модель статистической физики, предназначенная для описания намагничивания материала): для этого нужно найти наибольшее сбалансированное подмножество рёбер в знаковом графе Σ, которым описывается модель. В связи с широким применением знаковых графов актуальны исследования, направленные на разработку алгоритмов решения задач, связанных со знаковыми графами.

Целью данной работы является разработка алгоритмов распознавания сбалансированности и поиска максимального сбалансированного подграфа. Для достижения поставленной цели поставлены следующие задачи:

– Выявить свойства знакового графа.
– Разработать алгоритм распознавания сбалансированности знакового графа.
– Сформулировать задачу поиска максимального сбалансированного подграфа.
– Построить модель целочисленного программирования для задачи поиска наибольшего сбалансированного множества рёбер знакового графа.
– Разработать программу вычисления точного решения этой задачи с помощью пакета программ CPLEX.
– Выполнить вычислительные экспериментов, целью которых является проверка адекватности разработанных алгоритмов.

1. Сбалансированный знаковый граф и его свойства

Знаковым графом называется пара $\Sigma = (G, \sigma)$, где $G = (V, E)$ – неориентированный граф, на рёбрах которого задана функция знака σ: $E \to \{+, -\}$, $n = |V| \geq 2$, $m = |E| \geq 1$. Знак цикла определяется как произведение знаков рёбер, образующих этот цикл. Цикл графа является положительным, если он содержит четное число отрицательны рёбер, и отрицательным в противном случае.

Знаковый граф $\Sigma = (G, \sigma)$ называется сбалансированным, если любой его простой цикл положительный. На рис. 1 приведены примеры а) несбалансированного и б) сблансированного графов.

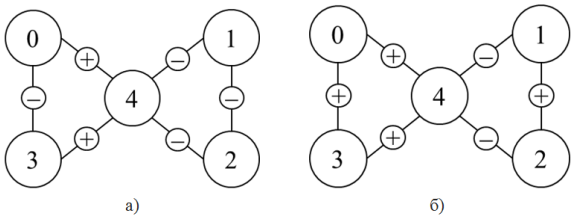


Рис. 1. Несбалансированный и сбалансированный графы

Необходимое и достаточное условие сбалансированности устанавливает следующая теорема.

Теорема 1 (Картрайт–Харари, [1]). Знаковый граф $\Sigma = (G, \sigma)$ сбалансирован тогда и только тогда, когда множество его вершин V можно разбить на две доли A и B, одна из которых может быть пустой, таким образом, что $A \cup B = V$, $A \cap B = \varnothing$ и любое ребро, соединяющее вершины из одной доли, имеет знак «+», а ребро, соединяющее вершины из разных долей, имеет знак «–».

Из данной теоремы вытекают важные свойства знаковых графов.

1. Если в графе $\Sigma = (G, \sigma)$ все рёбра положительные, то граф Σ всегда сбалансированный.
2. Если граф $\Sigma = (G, \sigma)$ является двудольным и все его рёбра одного знака, то граф Σ сбалансированный.
3. Каждая из долей A и B сбалансированного графа $\Sigma = (G, \sigma)$ порождает подграф, содержащий только положительные рёбра, а значит, являющийся сбалансированным.
4. Если ни одна из долей A и B сбалансированного графа $\Sigma = (G, \sigma)$ не является пустой, то множество отрицательных рёбер образует разрез (не обязательно минимальный) этого графа.
5. Всякая положительная клика графа сбалансированного графа $\Sigma = (G, \sigma)$ – клика, содержащая только положительные рёбра, –всегда принадлежит одной из его долей A или B.
6. Всякий ациклический знаковый граф $\Sigma = (G, \sigma)$ всегда сбалансированный.

Указанные свойства 1–6 могут быть использованы при решении базовых задач, связанных со знаковыми этими графами.

Также из теоремы 1 следуют условия согласованности рёбер: любое ребро, соединяющее вершины из одного подмножества должно быть положительно, а ребро, соединяющее вершины из разных подмножеств – отрицательно. На рис. 2 слева показаны случаи, когда условия согласованности выполняются, а справа – не выполняются.

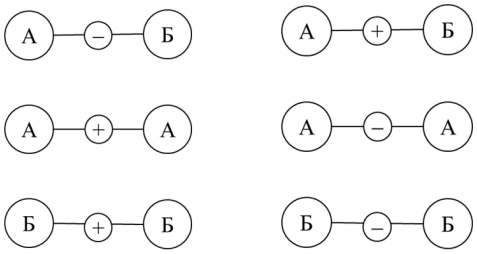


Рис. 2. Условия согласованности

2. Задача 1. Распознавание сбалансированности знакового графа

Пусть задан знаковый граф $\Sigma = (G, \sigma)$, где $G = (V, E)$ –простой неориентированный граф, $n = |V| \geq 2$, $m = |E| \geq 1$. Необходимо проверить, является ли граф Σ сбалансированным.

Для решения данной задачи был разработан алгоритм Signed_Balance, который представлен на рис. 3.

Алгоритм Signed_Balance

Вход: количества вершин и рёбер графа, список рёбер с указанием знака для каждого ребра.

1: ***begin***
2: $q \leftarrow \varnothing$
3: $q \leftarrow s$
4: ***while*** $q \neq \varnothing$
5: $s \leftarrow q^*$
6: $Z[s] \leftarrow Z[y] \cdot V[y][s]$
7: **delete**(q*)
8: $q \leftarrow N[s]$
9: ***end while***
10: **Matching**(Z)

Выход: ДА / НЕТ (исходный граф сбалансирован / несбалансирован)

Рис. 3. Алгоритм распознавания сбалансированности знакового графа

При описании алгоритма Signed_Balance используются следующие обозначения:

s – текущая вершина с окрестностью $N[s]$;

y – вершина, предшествующая s;

$Z[s]$, $Z[y]$ – метки вершин s и y соответственно, определяемые значениями +1 или –1;

$V[y][s]$ – знак ребра между s и y;

q – структура данных, задающая очередь просмотра вершин графа;

q* – последняя вершина из q.

Процедура Matching(Z) алгоритма Signed_Balance выполняет проверку условий согласованности рёбер. В данном случае, ребро $e = \{s, y\}$ считается согласованным, если знак этого ребра равен произведению знаков меток его концевых вершин s, y.

На выходе алгоритм Signed_Balance выдает сведения о том, сбалансирован исходный граф или нет, и, в случае сбалансированности, вариант разделения множества вершин графа на подмножества A и B. Разработанный алгоритм основан на классическом алгоритме поиска в глубину в графе [4], поэтому вычислительная слож-

ность алгоритма Signed_Balance для знакового графа с n вершинами и m ребрами составляет $O(n+m)$. Алгоритм Signed_Balance на практике работает быстрее, если в нем использованы свойства 1, 2, 6 знаковых графов. Вычислительная сложность применя этих свойств сопоставимо с $O(n+m)$. Это означает, что задача распознавания сбалансированности знакового графа является полиномиально разрешимой. Вычислительные эксперименты, проведенные с целью проверки правильности алгоритма Signed_Balance, приведены в табл. 1 раздела «Вычислительные эксперименты».

3. Задача 2. MAX-SET-BALANCED

Сформулируем задачу поиска в знаковом графе наибольшего по мощности сбалансированного множества ребер (MAX-SET-BALANCED, или коротко MSB).

Пусть задан знаковый граф $\Sigma=(G,\sigma)$, где $G=(V,E)$ – простой неориентированный граф, $n=|V|\geq 2$, $m=|E|\geq 1$. Требуется найти наибольшее по мощности множество ребер $E'\subseteq E$, которое сбалансировано в $\Sigma=(G,\sigma)$.

Множество ребер $E'\subseteq E$ считается сбалансированным в $\Sigma=(G,\sigma)$, если порожденный подграф $\Sigma(E')$ сбалансирован. Задача MSB является NP-трудной. Действительно, в частном случае, когда в графе $\Sigma=(G,\sigma)$, все рёбра отрицательные, задача MSB сводится к поиску в G двудольного подграфа с наибольшим числом рёбер. Это эквивалентно решению известной задачи комбинаторной оптимизации – задачи о наибольшем разрезе графа (MAX-CUT), для которой доказана NP-трудность [5]. Задачу MSB можно сформулировать как задачу целочисленного программирования (ЦП) и решать с помощью с помощью пакета программ CPLEX [6].

Обозначим через E^+ множество положительных, а через E^- множество отрицательных рёбер знакового графа $\Sigma=(G,\sigma)$, полагая, что $E^+\cup E^-=E$, $E^+\cap E^-=\varnothing$. Припишем каждой вершине $i\in V$ метку $l(i)=+1$ или $l(i)=-1$. Назовём полученный помеченный граф состоянием L знакового графа $\Sigma=(G,\sigma)$. Пусть L – некоторое произвольное состояние знакового знакового графа $\Sigma=(G,\sigma)$. Определим $n=|V|$ булевых переменных:

$$x_i=\begin{cases}1, \text{если вершина } i \text{ имеет метку } l(i)=+1,\\ 0, \text{если вершина } i \text{ имеет метку } l(i)=-1.\end{cases}$$

Введём еще $m=|E|$ булевых переменных:

$$x_i=\begin{cases}1, \text{если ребро } (i,j) \text{ согласованно относительно } L,\\ 0, \text{если ребро } (i,j) \text{ не согласованно относительно } L.\end{cases}$$

Тогда MSB в терминах ЦП имеет вид:

$$\sum_{(i,j)\in E} y_{ij}\to \max \tag{1}$$

при условиях:

$$2y_{ij}-(2x_i-1)(2x_j-1)=1,\ (i,j)\in E^+, \tag{2}$$

$$2y_{ij}+(2x_i-1)(2x_j-1)=1,\ (i,j)\in E^-, \tag{3}$$

$$x_i\in\{0,1\},\ i\in V, \tag{4}$$

$$y_{ij}\in\{0,1\},\ (i,j)\in E. \tag{5}$$

Заметим, что целевая функция (1) линейная, а ограничения (2)–(3) квадратичные. В задаче (1)–(5) поиск оптимального решения осуществляется в пространстве 2^{n+m} решений. Целевая функция (1) означает максимизацию количества согласованных рёбер,

входящих в знаковый граф $\Sigma = (G, \sigma)$. Ограничения (2) и (3) обеспечивают условие согласованности положительных и отрицательных ребер соотвественно. Решение задачи ЦП (1)–(5) можно находить с помощью пакета программ CPLEX, а также возможно применение приближенных методов решения, включая эвристики. На основе CPLEX проведены численные эксперименты, цель которых заключалась в проверке адекватности математической модели рассматриваемой задачи (1)–(5) на тестовых графах. Результаты экспериментов приведены в табл. 1 раздела «Вычислительные эксперименты». Вычислительные эксперименты проводились на персональном компьютере со следующими характеристиками: IntelCore i7-3630QM CPU @ 2.40 GHz, 8Гб оперативной памяти, 64-разрядная операционная система Windows 8.1.

Эксперименты показали, что нахождение точного решения задачи (1)–(5) на рассматриваемом компьютере с помощью демоверсии пакета CPLEX возможно только для графов размерности не более 15.

4. Вычислительные эксперименты

В данном разделе приведены результаты проверки правильности алгоритма Signed_Balance и адекватности модели ЦП задачи MSB.

Таблица 1

Результаты вычислительных экспериментов

Знаковый граф $\Sigma = (G, \sigma)$, количество его вершин и рёбер	Структура графа G	Особенности расстановки знаков рёбер	Результат работы алгоритма Signed_Balance	Результат работы CPLEX	Время работы CPLEX, мс
G_1, $n = 5$, $m = 4$	дерево	все рёбра одного знака	сбалансирован	все рёбра дерева	10
G_2, $n = 5$, $m = 4$	дерево	знаки рёбер произвольные	сбалансирован	все рёбра дерева	12
G_3, $n = 4$, $m = 4$	двудольный	все рёбра одного знака	сбалансирован	все ребра графа	10
G_4, $n = 4$, $m = 4$	двудольный	знаки рёбер произвольные	несбалансирован	3 ребра графа	39
G_5, $n = 6$, $m = 8$	произвольный	знаки рёбер произвольные	несбалансирован	7 рёбер графа	34

Проведённые вычислительные эксперименты показывают, что результаты работы алгоритма Signed_Balance и программы CPLEX, с помощью которой проверялась модель ЦП задачи MSB, не противоречат друг другу. Модель и алгоритм являются правильными.

Заключение

В работе показано, что распознавание сбаласированности знакового графа является полиномиально разрешимой задачей, а нахождение наибольшего сбалансированного подмножества рёбер знакового графа – оптимизационной NP-трудной задачей. Был разработан алгоритм распознавания сбалансированности знакового графа и модель целочисленного программирования задачи MSB. В дальнейшем планируется продолжить исследования в направлении создания алгоритмов приближенного решения задачи MSB и рассмотрение более общего случая *k*-сбалансированности.

ЛИТЕРАТУРА

1. *Harary F.* Structural Balance: A Generalization of Heider's Theory // Psychological Review. – 1956. – Vol. 63 (5). – P. 277–293.

2. *Reiner V., Tseng D.* Critical groups of covering, voltage, and signed graphs. URL: https://arxiv.org/pdf/1301.2977v2.pdf (Дата доступа 22.04.2019)

3. *Pal A., Banerjee D.* ApplicationOf Signed Graph In Decision Making // International Journal Of Computers & Technology. – 2015. – Vol. 14 (6). – P. 5824–5833.

4. *Емеличев В.А.* Лекции по теории графов. – М.: Наука, 1990.
5. *Гэри М., Джонсон Д.* Вычислительные машины и труднорешаемые задачи: монография. – М.: Мир, 1982.
6. CPLEX. URL: http://www.ilg.com/products/cplex/ (Дата доступа 22.04.2019)

СТРУКТУРА НЕЗАВИСИМОСТИ СЛУЧАЙНЫХ СОБЫТИЙ И МАТРОИДЫ

А.Р. Идрисова
Сибирский Федеральный Университет

Введение

Целью работы является описание алгоритма построения конфигураций с заданной точностью. В данной работе описаны примеры конфигураций со структурой зависимости определяемой матроидом.

1. Основные определения

Определение: Матроид – это упорядоченная пара (E,J), где E – непустое конечное множество, J – совокупность подмножества E. При этом должны выполняться аксиомы:

1. $\theta \in J$.
2. Если $A \in J$ и $B \in B \subset A$, то $B \in J$.
3. Если $A,B \in J$ и $|A| > |B|$, то $\exists x \in A$ такой, что $B \cup \{x\} \in J$ [1].

Множество E называется **основным** множеством матроида, а множества из J называются **независимыми** множествами.

Определение: Базис – максимальное по включению независимое множество [2].

Определение: Цикл – минимальное по включению зависимое множество.

Определение: Универсальный матроид $U_{k,n}$ – множество E имеет мощность n, независимыми множествами являются подмножества мощностью не больше k [3].

Определение: Два события $A,B \in F$ независимы если $P(A \cap B) = P(A) \cdot P(B)$.

Определение: События $A_1,...,A_n$ называются независимыми в совокупности, если для любого $1 \le k \le n$ и любого набора различных меж собой индексов $1 \le i_1,...,i_n \le n$ имеет место равенство: $P(A_{i_1} \cap \ldots \cap A_{i_k}) = P(A_{i_1}) \cdot \ldots \cdot P(A_{i_k})$.

Определение: Назовем события $A_1,...,A_n$ независимыми ровно по k ($k < n$), если любые $k+1$ событий зависимы.

2. Примеры матроидов

Пример 1.

Пусть E – конечная система векторов некоторого линейного пространства над полем F, а J состоит из E, а также пустого множества. Тогда, как известно из линейной алгебры, аксиомы 2 и 3 будут выполнены. Аксиома 1 выполнена по определению. Матроид (E,J) – векторный матроид. [2]

Пример 2.

Рассмотрим матрицу $A = \begin{bmatrix} 1 & 0 & 0 & 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 1 & 1 & 1 & 0 \\ 0 & 0 & 1 & 0 & 1 & 1 & 0 \end{bmatrix}$. Пусть $E = \{1,2,3,4,5,6,7\}$ – номера столбцов матрицы A, а I – набор таких подмножеств $J \subset E$, для которого множество соответствующих столбцов матрицы A линейно независимо над полем $\mathbb{R}$. Тогда J

состоит из подмножеств $E \setminus \{7\}$, содержащих не более трёх элементов, кроме $\{1,2,4\}$, $\{2,3,5\}$, $\{2,3,6\}$ и любого подмножества, содержащего $\{5,6\}$. Пара (E,J) является примером матроида. E – основное множество, J – множество независимых подмножеств [3].

Пример 4. Матроид Фано.

Рис. 1 показывает 3-базовый матроид, называемый матроидом Фано, пример который появился в 1935 г. в статье Уитни.

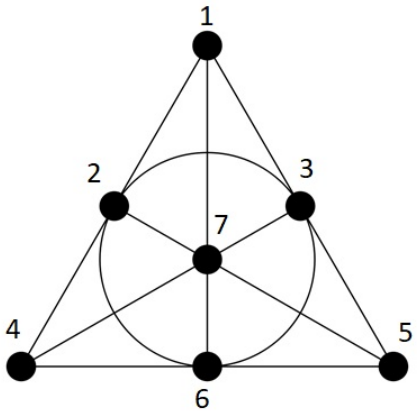


Рис. 1

Здесь точки – это элементы основного множества, а кривые «протянуты» через каждый трехэлементный цикл.

Таким образом, независимыми множествами будут множества не лежащие на одной кривой, а это множества по одному, по двум и по трем элементам, кроме: $\{124\}, \{135\}, \{176\}, \{347\}, \{257\}, \{456\}, \{236\}$.

Пример 5. Тетраэдр Бернштейна.

Рассмотрим правильный тетраэдр, три грани которого окрашены соответственно в красный, синий, зеленый цвета, а четвёртая грань содержит все три цвета. A – выпадение грани, содержащей красный цвет, B – выпадение грани, содержащей синий цвет, C – выпадение грани, содержащей зеленый цвет.

Так как каждый цвет есть на двух гранях из четырёх, вероятность каждого из этих событий равна $P(A) = P(B) = P(C) = \frac{1}{2}$. Так как одна грань содержит все три цвета, а остальные – по одному, то вероятность пересечения любых двух событий равна $P(A \cap B) = P(A \cap C) = P(B \cap C) = \frac{1}{4}$; $P(A) \cdot P(B) = P(A) \cdot P(C) = P(B) \cdot P(C) = \frac{1}{4}$, т.е. все события попарно независимы. Вероятность пересечения всех трёх: $P(A \cap B \cap C) = \frac{1}{4}$; $P(A) \cdot P(B) \cdot P(C) = \frac{1}{2} \cdot \frac{1}{2} \cdot \frac{1}{2} = \frac{1}{8}$. События не являются независимыми в совокупности, т.к. $P(A \cap B \cap C) \neq P(A) \cdot P(B) \cdot P(C)$.

Пусть $E = \{A, B, C\}$ – множество рассматриваемых событий, а $J = \{\theta, \{A\}, \{B\}, \{C\}, \{AB\}, \{AC\}, \{BC\}\}$ – множества независимых событий. J состоит из всех подмножеств E, содержащих не более двух элементов, что соответствует универсальному матроиду $U_{2,3}$.

3. Конфигурация со структурой зависимости, определяемой матроидом

Определение: Конфигурация со структурой зависимости, определяемой матроидом $M=(E,J)$ – описание условий случайного эксперимента, в результате которого наступают события $A=\{A_1,...,A_n\}$, такие, что существует биекция между A и E, которая переводит любое множество независимых в совокупности событий $A_{i_1},...,A_{i_p}$ в элемент множества J, и наоборот.

Примеры.

• Монета: $E=\{\mathrm{O},\mathrm{P}\}$ $J=\{\theta,\{\mathrm{O}\},\{\mathrm{P}\}\}$ $U_{1,2}$

• Игральная кость: $E=\{1,2,3,4,5,6\}$ $J=\{\theta,\{1\},\{2\},\{3\},\{4\},\{5\},\{6\}\}$ $U_{1,6}$

• Тетраэдр Бернштейна

В работе удалось получить способ построения конфигураций, структура зависимости для которых описывается универсальным матроидом $U_{k,n}$.

Пример конфигурации для $U_{2,4}$.

Рассмотрим правильный додекаэдр, у которого грани закрашены в разные цвета. R – выпадение грани, содержащей красный цвет, B – выпадение грани, содержащей синий цвет, Y – выпадение грани, содержащей желтый цвет, G – выпадение грани, содержащей зеленый цвет. $E=\{R,G,B,Y\}$ $J=\{\theta,\{R\},\{G\},\{Y\},\{B\},\{RG\},\{RB\},\{RY\},\{GB\},\{GY\},\{BY\}\}$ $U_{2,4}$ будет додекаэдр с такими гранями: две грани не закрашены: θ, шесть граней окрашены в два цвета: RG,RB,RY,GB,GY,BY, четыре грани окрашены в три цвета: RGB,RGY,RBY,GBY.

Вероятность выпадения любого одно случайного цвета: $1/2$. Вероятность выпадения любых двух случайных цветов: $1/4$. Вероятность выпадения любых трех случайных цветов: $1/12$. Вероятность выпадения четырех цветов: 0.

Как мы видим, два события попарно независимы, а более двух событий уже зависимы, следовательно, получена конфигурация $U_{2,4}$.

Пример конфигурации для матроида Фано.

Пусть $E=\{1,2,3,4,5,6,7\}$. Тогда из определения J, будет состоять из множеств по одному, по двум и по трём элементам, кроме: $\{124\}$, $\{135\}$, $\{176\}$, $\{347\}$, $\{257\}$, $\{456\}$, $\{236\}$.

Нетрудно убедиться, что конфигурацией для матроида Фано будет октаэдр с гранями: $\{124\}$, $\{135\}$, $\{176\}$, $\{347\}$, $\{257\}$, $\{456\}$, $\{236\}$, $\{1234567\}$.

Проверим зависимость:

$$P(A_i)=\frac{1}{2},\ P(A_iA_j)=\frac{1}{4},\ P(A_iA_jA_k)=\begin{cases}\frac{1}{8}\text{ для всех остальных наборов } i,j,k,\\ \frac{1}{4}\text{ для } i,j,k\text{, лежащих на одной кривой.}\end{cases}$$

Заключение

В ходе работы мы научились строить конфигурации для $U_{k,n}$ при $k\le n$. Также была написана программа для построения конфигураций для $U_{k,n}$. Не всегда конфигурация будет реализована с помощью правильного многогранника. Конфигурация не единственна.

ЛИТЕРАТУРА

1. *Асанов М.О., Баранский В.А., Расин В.В.* Дискретная математика: графы , матроиды, алгоритмы. – Ижевск: НИЦ "РХД", 2001. –288 с.

2. *Эвнин А.Ю.* Элементарное введение в матроиды // Матем. обр. – 2005. – № 2 (33). – С. 2–33.

3. *Oxley J.* What is a matroid? // Oxford University Press, 1992.

ЧИСЛЕННЫЙ АНАЛИЗ БЫСТРОДЕЙСТВИЯ ТРАНСПОРТНОГО ПРОТОКОЛА В РЕЖИМЕ ГРУППОВОГО ОТКАЗА С ПРЯМОЙ КОРРЕКЦИЕЙ ОШИБОК

П.Х. Карим, С.П. Сущенко, А.Л. Чаллуб

Томский государственный университет

Введение

Пропускная способность транспортного соединения является крайне важной характеристикой компьютерных сетей. Данный показатель определяет качество сетевых сервисов для абонентов и определяется значениями протокольных параметров (размер окна, длительность тайм-аута), характеристиками тракта передачи данных (длительность круговой задержки, достоверность передачи данных в различных направлениях транспортного соединения) [1]. В настоящее время получают широкое распространение технологии прямой коррекции ошибок [2–4], применяемые как дополнительный сервис в транспортных протоколах наряду с методом решающей обратной связи для снижения объема повторно передаваемого трафика. Исследование транспортного соединения и анализ его потенциальных возможностей выполнялся в [2–16]. Но аналитические результаты получены только для однозвенного тракта передачи данных [5–7], либо при существенных ограничениях на параметры протокола [8–14]. Примером одной из модификаций транспортного протокола является подключаемый к UDP протоколу дополнительный протокольный механизм под названием QUIC (Quick UDP Internet Connections) [15,16]. Это новое дополнение протокола пока еще не является стандартом и в настоящее время эксперименты с ним и исследование его эффективности продолжаются. Протокол QUIC развернут как дополнительный сервис в кампании Google. Следует отметить, что протокол QUIC имеет ряд недостатков, к числу которых относится использование ограниченного множества значений параметров механизма прямой коррекции ошибок. Исследователи протокола QUIC [13] применяли операцию поразрядного паритета и анализировали его только в натурных экспериментах (тестирование на оборудовании) [16]. Как правило исследования преимуществ метода прямой коррекции ошибок проводятся на качественном уровне и для некоторых частных случаев численно. Известные исследования не определяют области в пространстве протокольных параметров и характеристик транспортного соединения, в которых применение метода опережающего исправления ошибок дает положительные результаты. Скрытные возможности транспортного протокола с применением метода прямой коррекции ошибок не изучены полностью. Отсутствуют аналитические зависимости комплексного влияния протокольных параметров, характеристик тракта передачи данных и параметров метода коррекции, на быстродействие транспортного соединения. Не исследовано влияние соотношений между длительностью круговой задержки и протокольными параметрами на пропускную способность тракта передачи данных управляемого транспортным протоколом.

В работе предложена математическая модель процесса передачи данных с прямой коррекцией ошибок в фазе информационного переноса в виде цепи Маркова с дискретным временем [12]. Проведены числительные анализы пропускной способности транспортного канала с применением механизма примой коррекции ошибок и показано превосходство новой модели по сравнению классической моделью. Доказано, что в опре-

деленных условиях применения метода примой коррекции ошибок увеличивает пропускную способность транспортного канала.

1. Модель транспортного соединения

Рассмотрим процесс переноса данных между абонентами транспортного протокола, основанного на алгоритме с решающей обратной связью. Примером семейства таких надежных протоколов является доминирующий в современных компьютерных сетях протокол TCP [1]. Полагаем, что взаимодействующие абоненты имеют неограниченный поток данных для передачи, а обмен выполняется сегментами данных транспортного протокола одинаковой длины. Считаем, что участки переприема вдоль тракта передачи данных имеют одинаковое быстродействие в обоих направлениях, а длительность цикла передачи сегмента в отдельном звене составляет *t*. В общем случае длина пути от источника до адресата, переносящего информационный поток, и длина обратного пути, по которому передаются подтверждения на принятые сегменты могут быть различными. Полагаем, что длина тракта передачи данных, выраженная в количестве участков переприема, в прямом направлении равна $D_{\text{п}} \geq 1$. Обратный тракт, по которому доставляются подтверждения отправителю о корректности приема последовательности блоков сегментов, имеет длину $D_{\text{о}} \geq 1$. Заданы вероятности искажения сегмента в каналах связи для прямого $R_{\text{п}}(d)$, $d = \overline{1, D_{\text{п}}}$ и обратного – $R_{\text{о}}(d)$, $d = \overline{1, D}_{\text{о}}$ направлений передачи каждого участка переприема. Тогда достоверность передачи сегментов вдоль тракта от источника до адресата и обратно составят $F_{\text{п}} = \prod_{d=1}^{D_{\text{п}}} \left(1 - R_{\text{п}}(d)\right)$ и $F_{\text{о}} = \prod_{d=1}^{D_{\text{о}}} \left(1 - R_{\text{о}}(d)\right)$ соответственно. Считаем, что потерь сегментов из-за отсутствия буферной памяти в узлах тракта не происходит. Полагаем, что передача данных отправителем реализуется блоками, содержащими *B* сегментов, из которых $1 \leq A \leq B$ – являются информационными, а $B - A$ – дополнительными (избыточными) той же длины. Полагаем, что все сегменты имеют контрольные суммы, позволяющие обнаружить ошибки в каждом из них. Потеря (искажение) до $B - A$ произвольных сегментов в блоке позволяет восстановить все сегменты блока (например, передачей дублей при $A = 1$, $B \geq 2$, оправкой избыточного сегмента с поразрядной четностью всех информационных сегментов по технологии RAID–массивов четвертого уровня [11] при $A > 1$, $B = A + 1$ и др.). Управление потоком данных реализуется механизмом скользящего окна [1] с протокольным параметром ширины окна $\omega \geq 1$, выраженном в количестве блоков. Полагаем, что подтверждения о корректности полученных адресатом блоков сегментов переносятся в каждом сегменте встречного потока. При невозможности прямого восстановления переданных сегментов блока (искажении более $B - A$ сегментов в блоке) весь блок передается повторно. Тогда, процесс информационного переноса в виртуальном соединении, управляемом транспортным протоколом, может быть описан Марковским процессом с дискретным временем (с длительностью такта *t*) в силу того, что время между получениями подтверждений имеет геометрическое распределение с параметром $F_{\text{о}}$. Данная модель является обобщением формализаций процесса передачи данных, предложенных в [3–6], на случай транспортного соединения произвольной длины и механизма прямой коррекции ошибок. Область возможных состояний цепи Маркова определяется длительностью тайм-аута ожидания подтверждения *S*, выраженной в количестве циклов продолжительности *t*. Размер тайм-аута связан с длиной тракта, шириной окна и размером блока неравенствами $S \geq \omega B + 1$, $S \geq D_п + D_о + B - 1$. Очевидно, что сумму длин прямого и обратного трактов можно интерпретировать как круговую задержку одиночного сегмента $D = D_{\text{п}} + D_{\text{о}}$ в детерминированном транспортном соединении, выраженную в длительностях *t*. Круговая задержка для блока

сегментов составит $D+B-1$. Состояниям цепи Маркова $i=\overline{0,\omega B}$ соответствует размер очереди переданных, но не подтвержденных сегментов в источнике потока, а состояниям $i=\overline{\omega B+1,S-1}$ – время, в течение которого отправитель не активен и ожидает получение подтверждения о корректности приема переданной последовательности из ω блоков сегментов. Из нулевого состояния в $D+B-2$ источник продвигается с каждым тактом t с вероятностью детерминированного события. В состояниях $i \geq D+B-2$ после истечения очередного дискретного цикла t к отправителю начинают прибывать подтверждения и, в зависимости от результатов доставки блоков сегментов с учетом технологии прямой коррекции ошибок, отправитель передает новые блоки сегментов (при положительном подтверждении), либо повторно – искаженные (не поддающиеся прямому восстановлению). Завершение цикла пребывания в состоянии $D+B-2$ соответствует времени доведения первого блока сегментов до адресата и получения на него подтверждения. Дальнейший рост номера состояния происходит с вероятностью искажения подтверждения $1-F_{\text{o}}$ в обратном тракте. Получение подтверждения в состояниях $i \geq D+B-2$ вызывает переход в $D-1$ состояние при $\omega \geq K+2$ только в случае успешной передаче доставленных адресату блоков, в противном случае следует переход в 0 состояние. Здесь $K=\left\lfloor \dfrac{D-2}{B} \right\rfloor$, где $\lfloor ... \rfloor$ – означает «целая часть» дроби [12]. В силу того, что в состояниях $i \geq \omega B$ источник приостанавливает отправку блоков сегментов, получение подтверждений при $\omega \geq K+2$ в состояниях $i=\overline{(\omega+k)B-1,(\omega+k+1)B-2}$, $k=\overline{1,K}$ приводит к переходу в состояния $D-kB-1$, $k=\overline{1,K}$, только при успешной доставке данных (иначе – в 0 состояние). В состояниях $i=\overline{(\omega+K+1)B-1,S-2}$ выполняется переход в нулевое состояние, поскольку размер очереди переданных, но не подтвержденных информационных сегментов при этом обнуляется. В состоянии $S-1$ истекает тайм-аут ожидания подтверждения от получателя о корректности принятых блоков сегментов и происходит безусловный переход в нулевое состояние [12].

2. Вероятности состояний для транспортного протокола с механизмом прямой коррекции ошибок

Переходные вероятности π_{ij} из исходного состояния i в результирующее j цепи Маркова, описывающей процесс передачи информационного потока с технологией прямой коррекцией ошибок в режиме группового отказа для $\omega \geq K+2$, $S \geq D+B(\omega+1)-2$ имеют вид:

$$\pi_{ij}=\begin{cases}1,\ i=\overline{0,D+B-3},\ j=i+1;\\ 1-F_o,\ i=\overline{D+B-2,S-2},\ j=i+1;\\ F_o\psi^k,\ i=\overline{D+Bk-2,D+(k+1)B-3},\ k=\overline{1,G},\ j=D-1;\\ F_o(1-\psi^k),\ i=\overline{D+Bk-2,D+(k+1)B-3},\ k=\overline{1,G},\ j=0;\\ F_o\psi^G,\ i=\overline{D+B(G+1)-2,B(\omega+1)-2},\ j=D-1;\\ F_o(1-\psi^G),\ i=\overline{D+B(G+1)-2,B(\omega+1)-2},\ j=0;\\ F_o\psi^{G+k},\ i=\overline{B(\omega+k)-1,B(\omega+k+1)-2},\ k=\overline{1,K},\ j=D-Bk-1;\\ F_o(1-\psi^{G+k}),\ i=\overline{B(\omega+k)-1,B(\omega+k+1)-2},\ k=\overline{1,K},\ j=0;\\ F_o,\ i=\overline{B(\omega+K+1)-1,S-2},\ j=0;\\ 1,\ j=0,\ i=S-1.\end{cases}$$

Здесь $G=\left\lfloor\frac{B(\omega+1)-2-(D+B-2)+1}{B}\right\rfloor=\omega-\left\lfloor\frac{D-1}{B}\right\rfloor$ – расстояние между моментами начала прекращения активности отправителя $B(\omega+1)-2$ (завершения оправки ω блоков) и начала поступления ему квитанций $D+B-2$, выраженное в размерах B, $\psi=\sum_{i=A}^{B}C_B^iF_{\text{п}}^i(1-F_{\text{п}})^{B-i}$. Решая систему уравнений равновесия, находим вероятности состояний цепи Маркова и далее получаем показатель пропускной способности транспортного соединения с применением механизма прямой коррекции ошибок:

$$Z(D,\omega,S,A,B,F_{\text{п}},F_{\text{о}})=\frac{P_{\text{о}}A\psi\left(1-\overline{F}_{\text{о}}^B\right)\left(1-\overline{F}_0^B\psi\right)}{BДF_{\text{о}}(1-\psi)}\left\{\left[1-\left(\overline{F}_{\text{о}}^B\psi\right)^{\omega}\right]\left(\frac{1-\psi}{1-\overline{F}_{\text{о}}^B\psi}\right)-\overline{F}_{\text{о}}^{S-D-B+2}\left(1-\psi^{\omega}\right)\right\},\quad(1)$$

где

$$P_{\text{о}}=ДF_{\text{о}}\left(1-\overline{F}_{\text{о}}^B\psi\right)\Big/\left\{ДF_{\text{о}}\left(D-BK-1\right)\left(1-\overline{F}_{\text{о}}^B\psi\right)+\left(1-\overline{F}_{\text{о}}^B\psi\right)^2\left[1+F_{\text{о}}\left(B-1\right)-\overline{F}_{\text{о}}^{S-D-B+2}\right]+\right.$$

$$\left.+BF_{\text{о}}\left[K\left(1-\psi\right)\left(1-\overline{F}_{\text{о}}^B\psi\right)+\psi^{G+1}\left[1-\left(\overline{F}_{\text{о}}^B\psi\right)^K\right]\left(\overline{F}_{\text{о}}^{B\omega-D+1}\left(1-\overline{F}_{\text{о}}^B\psi\right)-\left(1-\psi\right)\overline{F}_{\text{о}}^{G+1}\right)\right]\right\},$$

$$Д=\left(1-\overline{F}_{\text{о}}^B\psi\right)\left[\psi^{G+1}\overline{F}_{\text{о}}^{B\omega-D+1}\left(\overline{F}_{\text{о}}^B\psi\right)^K\right]+\left(1-\psi\right)\left[1-\left(\overline{F}_{\text{о}}^B\psi\right)^{K+G+1}\right],\ \overline{F}=1-F_{\text{о}}.$$

Очевидно, что поиск в многомерном пространстве признаков $(D,\omega,S,A,B,F_{\text{п}},F_{\text{о}})$ областей, обеспечивающих превосходство механизма прямой коррекции ошибок перед классической протокольной процедурой с решающей обратной связью по критерию пропускной способности, является сложной задачей. Решением данной проблемы может быть редуцирование размерности признакового пространства. Эффективными вариантами редуцирования размерности признакового пространства являются случаи абсолютно надежного обратного тракта передачи данных ($F_{\text{о}}=1$), неограниченной ширины окна $\omega\to\infty$ (а следовательно и длительности тайм-аута $S\to\infty$), а также случай однородного тракта передачи данных в прямом и обратном направлениях ($F_{\text{п}}=F_{\text{о}}=F$).

3. Численный анализ пропускной способности транспортного соединения

Рассмотрим случай абсолютно надежного обратного тракта $F_{\text{о}} = 1$. Тогда пропускная способность транспортного соединения принимает вид:

$$Z\left(D, \omega, S, A, B, F_{\text{п}}, 1\right) = \frac{A.\psi}{B(D-1)(1-\psi) + B^2}.$$

Чтобы определить области, где механизм с применением технологии прямой коррекции ошибок, дает положительные результаты, сравним с классическим вариантом протокола [8]. В классическом варианте пропускная способность имеет вид [8]:

$$Z_{\text{кл}}\left(D, \omega A, S, F_{\text{п}}, F_{\text{о}} = 1\right) = \frac{F_{\text{п}}}{\left(D-1\right)\left(1-F_{\text{п}}\right)+1}.$$

Определим выигрыш от применения технологии прямой коррекции ошибок перед классическим транспортным протоколом $Z = Z\left(D, \omega, S, A, B, F_{\text{п}}, F_{\text{о}}\right) - Z_{\text{кл}}\left(D, A\omega, S, F_{\text{п}}, F_{\text{о}}\right)$. Обозначая $F_{\text{п}} = F$, при $F_{\text{о}} = 1$ выигрыш принимает вид

$$Z = \frac{A.\psi}{B(D-1)(1-\psi) + B^2} - \frac{F_{\text{п}}}{(D-1)\left(1-F_{\text{п}}\right)+1}.$$

Численные исследования выигрыша при наборе параметров механизма прямой коррекции ошибок $A = 2$, $B = A + 1$ показывают (см. рис. 1 слева), что область положительных значений выигрыша расширяется от $F \in (0.781, 0.91)$, до $F \in (0.76, 0.93)$ с ростом D от 56, до 66. Кроме того, экстремальное значение выигрыша растет с увеличением D (см. рис. 1) и достигается практически в одном и том же пределе $F \in (0.796, 0.89)$. На рис. 1 справа проиллюстрированы сходные зависимости для набора параметров $A = 2$, $B = 2A$. Слева экстремальное значение выигрыша: при $Z\left(D = 56, F = 0.872\right) = 0.007883031$, $Z\left(D = 60, F = 0.88\right) = 0.010861443$, $Z\left(D = 66, F = 0.89\right) = 0.01501331$. Справа экстремальное значение выигрыша: при $Z\left(D = 56, F = 0.796\right) = 0.02185867$, $Z\left(D = 60, F = 0.805\right) = 0.024329853$, $Z\left(D = 66, F = 0.816\right) = 0.027705338$.

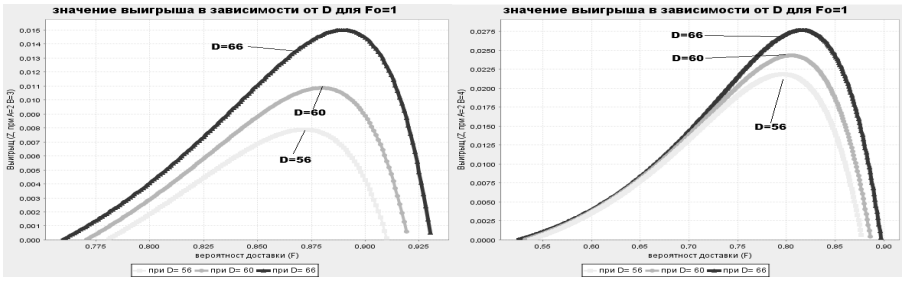


Рис. 1. Зависимость выигрыша Z от достоверности передачи данных в прямом тракте F при $F_{\text{о}} = 1$

На рис. 2 слева для $A = 3$, $B = A + 2$ видно, что область положительных значений выигрыша расширяется от $F \in (0.758, 0.842)$ до $F \in (0.725, 0.879)$ с ростом D от 56, до 66, так же, экстремальное значение выигрыша растет с увеличением D (см. рис. 2), оно находится в пределе $F \in (0.746, 0.83)$. На рис. 2 справа показаны сходные зависимости для набора параметров $A = 3$, $B = 2A$. Слева экстремальное значение выигрыша: при $Z\left(D = 56, F = 0.809\right) = 0.002493844$, $Z\left(D = 60, F = 0.818\right) = 0.005147922$, $Z\left(D = 66, F = 0.83\right) = 0.008802511$. Справа экстремальное значение выигрыша: при

$Z\left(D=56, F=0.746\right)=0.008818722$, $Z\left(D=60, F=0.755\right)=0.010963882$, $Z\left(D=66, F=0.767\right)=0.013878739$; при $A=3$, $B=A+1$ для круговой задержки (56,60,66) положительные результаты не наблюдаются.

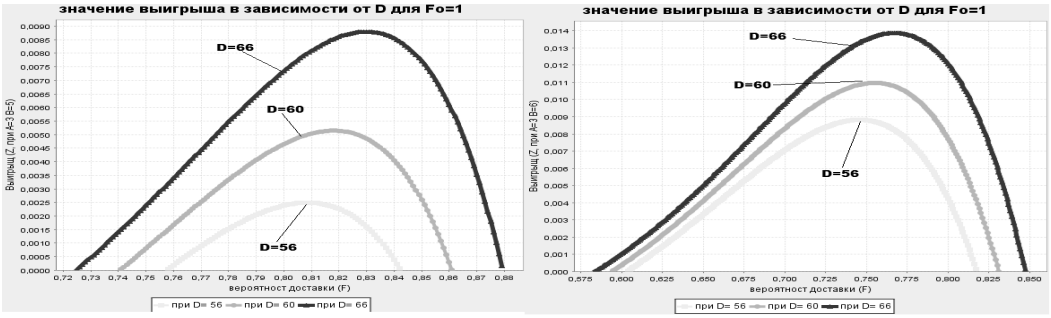


Рис. 2. Зависимость выигрыша *Z* от достоверности передачи данных в прямом тракте *F* при $F_o=1$

Рис. 3 слева для $A=4$, $B=A+3$ показывает положительные результаты выигрыша в пределах от $F\in(0.752,0.786)$ до $F\in(0.717,0.826)$ с ростом *D* от 60, до 66. При длине пути $D=56$, а также для параметров $A=4$, $B=A+1$ и $A=4$, $B=A+2$ нет положительных результатов выигрыша. Экстремальное значение выигрыша растет при увеличении значения круговой задержки *D* (см. рис. 3) в $F\in(0.708,0.783)$. На рис. 3 справа представлены сходные зависимости для набора параметров $A=3$, $B=2A$. Слева экстремальное значение выигрыша: при $Z\left(D=60, F=0.771\right)=3.82149E-4$, $Z\left(D=66, F=0.783\right)=0.003547793$. Справа экстремальное значение выигрыша: при $Z\left(D=56, F=0.708\right)=0.002154542$, $Z\left(D=60, F=0.718\right)=0.00408336$, $Z\left(D=66, F=0.731\right)=0.006695497$.

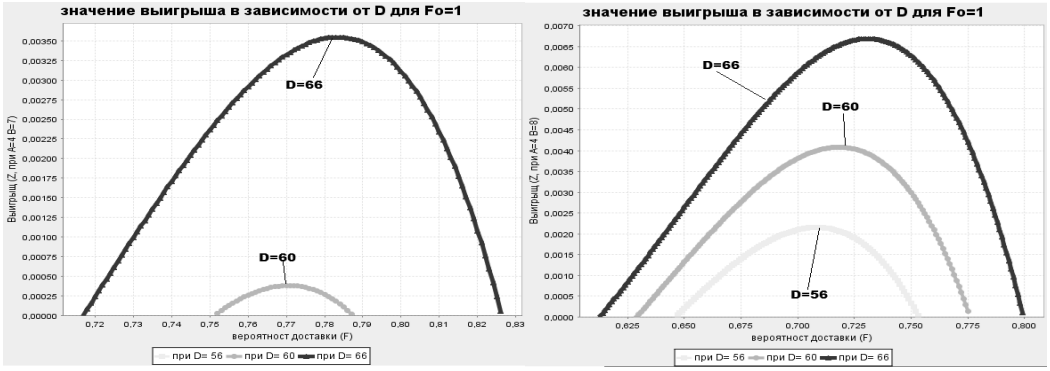


Рис. 3. Зависимость выигрыша *Z* от достоверности передачи данных в прямом тракте *F* при $F_o=1$

На рис. 4 слева для $A=5$, $B=2A$ видно, что положительные значений выигрыша получаются только для длины $D=66$ в пределах $F\in(0.639,0.749)$, в связи с тем, что длина пути короткая по сравнению с длиной отправляемых данных (*B*). Экстремальное значение выигрыша: при $Z\left(D=66, F=0.703\right)=0.00228161$.

Рис. 4. Зависимость выигрыша Z от достоверности передачи данных в прямом тракте F при $F_{\text{о}} = 1$

Проанализируем выигрыш пропускной способности канала в условиях неограниченной ширины окна ($\omega \to \infty$) и стохастической однородности прямого и обратного трактов передачи данных ($F_{\text{п}} = F_{\text{о}} = F$). Пропускная способность для протоколов с применением механизма прямой коррекции ошибок и без его использования согласно (1) определится соответственно следующим образом:

$$Z\left(D,\infty,\infty,A,B,F,F\right) = \frac{A\psi\left(1-(1-F)^{B}\right)}{B\left\{(1-\psi)F\left(D-1\right)+\left(1-(1-F)^{B}\psi\right)\left(1+F\left(B-1\right)\right)\right\}},$$

$$Z_{\text{кл}}\left(D,\infty,\infty,F,F\right) = \frac{F^{2}}{1+F\left(D-2\right)\left(1-F\right)}.$$

Положительные значения выигрыша, проиллюстрированные на рис. 5 слева для $A = 2$, $B = A+1$, показывают, что область выигрыша расширяется от $F \in (0.391, 0.923)$ до $F \in (0.383, 0.937)$ с ростом D от 56 до 66. На рис. 5 справа проиллюстрированы сходные зависимости для набора параметров $A = 2$, $B = 2A$. Положительные результаты находятся в пределах $F \in (0.136, 0.906)$. Кроме того, экстремальное значение выигрыша растет с увеличением D (см. рис. 5). Экстремальное значение выигрыша: при $Z\left(D = 56, F = 0.866\right) = 0.02172235$, $Z\left(D = 60, F = 0.874\right) = 0.02404891$, $Z\left(D = 66, F = 0.885\right) = 0.02734784$. При $2A$ экстремальное значение выигрыша: $Z\left(D = 55, F = 0.785\right) = 0.0398083$, $Z\left(D = 60, F = 0.794\right) = 0.04161215$, $Z\left(D = 65, F = 0.806\right) = 0.0441076$.

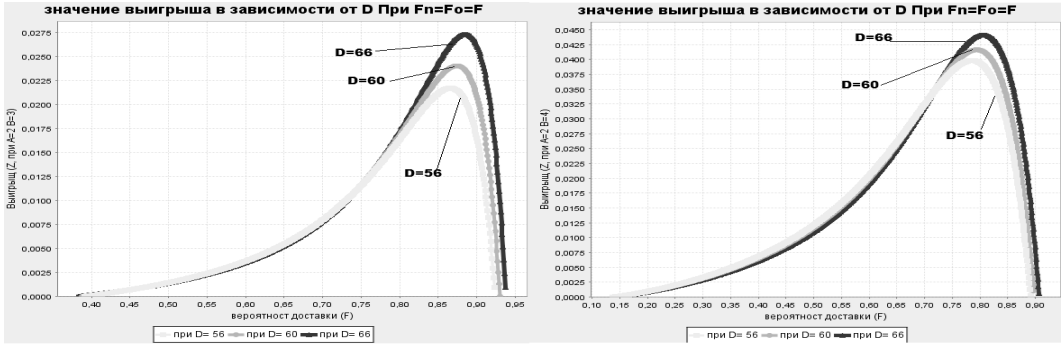


Рис. 5. Зависимость выигрыша ΔZ от достоверности передачи данных в прямом тракте F при $F_{\text{п}} = F_{\text{о}} = F$

На рис. 6 слева. для $A = 3$, $B = A+2$ видно, что область положительных значений выигрыша расширяется от $F \in (0.493, 0.876)$ до $F \in (0.479, 0.896)$ с ростом D от 56 до 66. На рис. 6 справа проиллюстрированы сходные зависимости для набора параметров

$A=2$, $B=2A$. Положительные результаты находятся в пределах $F\in(0.276,0.867)$. Кроме того, экстремальное значение выигрыша растет с увеличением D (см. рис. 6). При $A=3$, $B=A+1$ положительный результат пропускной способности не наблюдается. Слева экстремальное значение выигрыша: при $A=3$, $B=A+2$, $Z\left(D=55,F=0.803\right)=0.01737325$, $Z\left(D=60,F=0.812\right)=0.01951008$, $Z\left(D=65,F=0.824\right)=0.02247506$. Справа при $A=3$, $B=2A$ экстремальное значение выигрыша: $Z\left(D=55,F=0.736\right)=0.02604962$, $Z\left(D=60,F=0.746\right)=0.02767973$, $Z\left(D=65,F=0.758\right)=0.02990602$.

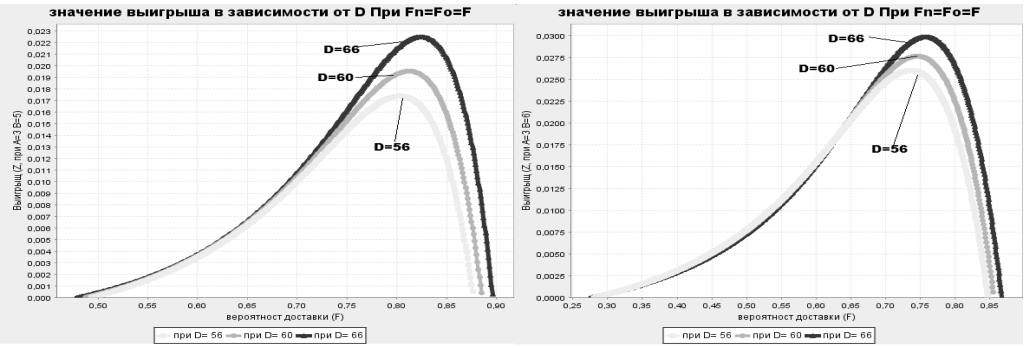


Рис. 6. Зависимость выигрыша ΔZ от достоверности передачи данных в прямом тракте F при $F_{\text{п}}=F_{\text{о}}=F$

По полученным результатам, представленным на рис. 7 А для $A=4$, $B=A+2$, можно видеть, что область положительных значений выигрыша расширяется от $F\in(0.808,0.814)$ до $F\in(0.762,0.861)$ с ростом D от 60 до 66, при $D=56$ положительных результатов нет.

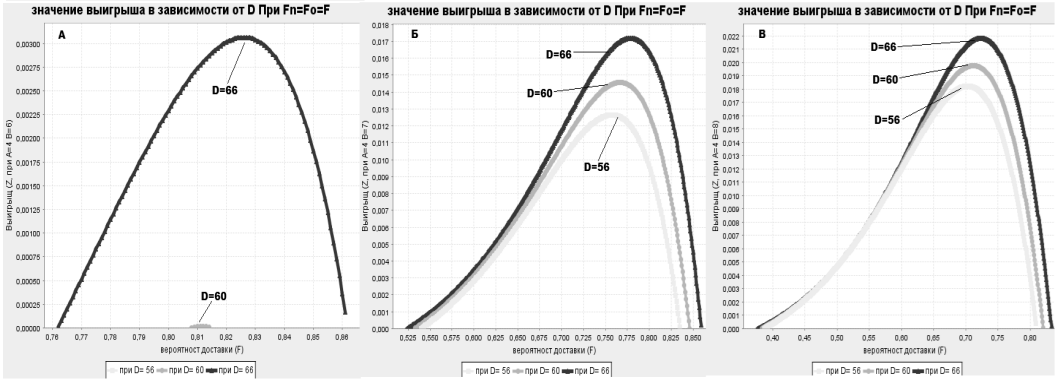


Рис. 7 А, Б, В. Зависимость выигрыша ΔZ от достоверности передачи данных в прямом тракте F при $F_{\text{п}}=F_{\text{о}}=F$

На рис. 7 Б для $A=4$, $B=A+3$ предел положительных значений выигрыша расширяется от $F\in(0.54,0.835)$ до $F\in(0.524,0.859)$ с ростом D от 56 до 66. На рис. 7 В $A=4$, $B=2A$ $F\in(0.387,0.809)$ до $F\in(0.377,0.833)$ для D от 56 до 66. При $A=4$, $B=A+1$ положительный результат пропускной способности не наблюдается. Экстремальное значение выигрыша при $A=4$, $B=A+2$ $Z\left(D=60,F=0.811\right)=1.512E-5$, $Z\left(D=66,F=0.826\right)=0.00306393$. Экстремальное значение выигрыша при $A=4$, $B=A+3$, $Z\left(D=56,F=0.757\right)=0.01264888$, $Z\left(D=60,F=0.767\right)=0.01456973$,

$Z\left(D=66, F=0.779\right)=0.01719778$. При $A=4$, $B=2A$ экстремальное значение выигрыша $Z\left(D=56, F=0.703\right)=0.01822156$, $Z\left(D=60, F=0.712\right)=0.01975585$, $Z\left(D=66, F=0.722\right)=0.02182873$.

Область положительных значений выигрыша на рис. 8 А, Б, и В расширяется в пределах $F\in(0.749, 0.779)$ до $F\in(0.697, 0.838)$, $F\in(0.563, 0.798)$ до $F\in(0.545, 0.826)$ и $F\in(0.45, 0.777)$ до $F\in(0.438, 0.803)$ для $A=5$, $B=A+3$, $A=5$, $B=A+4$ и $A=5$, $B=2A$ соответственно с ростом D от 56 до 66. При $A=5$, $B=A+1$, $B=A+2$ положительный результат пропускной способности не наблюдается. Экстремальное значение выигрыша: при $A=5$, $B=A+3$, $Z\left(D=56, F=0.765\right)=2.4066E-4$, $Z\left(D=60, F=0.776\right)=0.00231028$, $Z\left(D=66, F=0.789\right]=0.00516253$. Экстремальное значение выигрыша: при $A=5$, $B=A+4$, $Z\left(D=56, F=0.723\right)=0.00889926$, $Z\left(D=60, F=0.732\right)=0.01067038$, $Z\left(D=66, F=0.744\right)=0.01307139$. При $A=5$, $B=2A$ экстремальное значение выигрыша: $Z\left(D=56, F=0.678\right]=0.01300135$, $Z\left(D=60, F=0.688\right)=0.01447173$, $Z\left(D=66, F=0.698\right)=0.01645226$.

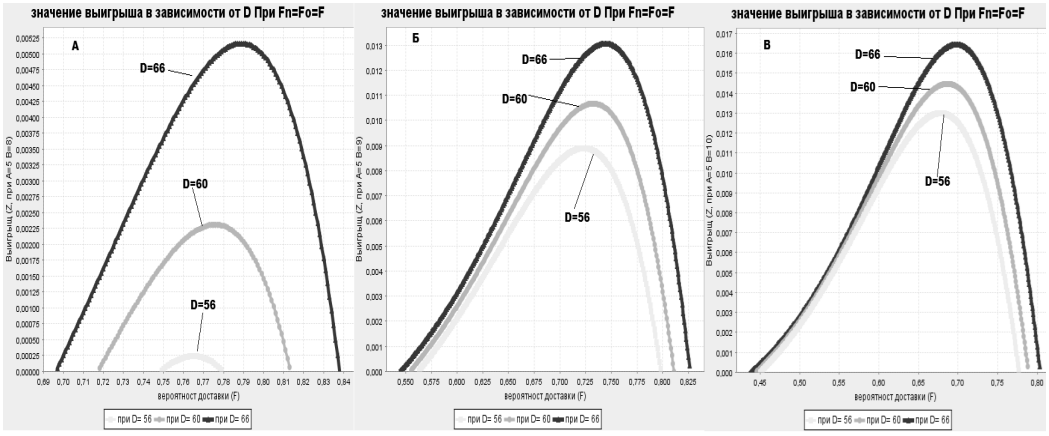


Рис. 8 А, Б, В. Зависимость выигрыша ΔZ от достоверности передачи данных в прямом тракте F при $F_{\text{п}}=F_{\text{о}}=F$

Заключение

В работе предложена модель процесса переноса сегментов данных в транспортном соединении, управляемом надежным транспортным протоколом с механизмом прямой коррекции ошибок и подтверждением данных, принятых получателем, в режиме группового повтора. Математическая модель основана на описании очереди переданных, но не подтвержденных сегментов данных цепью Маркова с конечным числом состояний и дискретным временем [12]. Представлен численный анализ пропускной способности транспортного соединения. Численные исследования выполнены для абсолютно надежного обратного тракта и для неограниченного размера окна. При данных условиях показано что, выигрыш пропускной способности увеличивается с ростом круговой задержки. Показано, что для группового режима повтора применение механизма прямой коррекции ошибок целесообразно на транспортных соединениях с большой круговой задержкой.

ЛИТЕРАТУРА

1. *Fall K., Stevens R.* TCP/IP Illustrated, Volume 1: The Protocols (2nd Edition) // Addison-Wesley Professional Computing Series. – 2012. – 1017 p.

2. *Lundqvist H., Karlsson G.* TCP with end-to-end FEC // International Zurich Seminar on Communications. – 2004. – P. 152–156.

3. *Barakat Ch., Altman E.* Bandwidth tradeoff between TCP and link-level FEC // Computer Networks. – 2002. – No. 39. – P. 133–150.

4. *Shalin R., Kesavaraja D.* Multimedia Data Transmission through TCP/IP using Hash Based FEC with AUTO-XOR Scheme // ICTACT Journal on Communication Technology. – 2012. – Vol. 03. – Issue 03. – P. 604–609.

5. *Boguslavsky L.B., Gelenbe E.* Analytical models transmission link control procedures for data computer networks with packet // Automation and Remote Control. – 1980. – No 7. – P. 181–192.

6. *Gelenbe E., Labetoulle J., Pujolle G.* Performance Evaluation of the HDLC Protocol // Computer Networks. – 1978. – Vol. 2. – No. 4/5. – P. 409–415.

7. *Kokshenev V.V., Sushchenko S.P.* Analysis of the asynchronous performance management procedures link transmission data // Computational technologies. – 2008. – Vol. 15. – Special issue No. 5. – P. 61–65.

8. *Kokshenev V.V., Mikheev P.A., Sushchnenko S.P.* Comparative Analysis of the Performance of Selective and Group Repeat Transmission Models in a Transport Protocol // Automation and Remote Control. – 2017. – Vol. 78. – No 2. – P. 65–81.

9. *Kokshenev V.V., Mikheev P.A., Sushchenko S.P.* Analysis of selective denial mode transport Protocol loaded in the transmit path, data // Vestnik TSU. Series control, computer facilities and computer science. – 2013. – No. 3(24). – P. 78–94.

10. *Kokshenev V.V., Sushchenko S.P.* Modeling sessions with Markov's chains // Theory of probability, random processes, mathematical statistics and applications: proceedings of the International scientific conference dedicated to the 80th anniversary of Professor G.A. Medvedev. Minsk on 23–26 February 2015. – Minsk, RIVS, 2015. – P. 311–316.

11. *Mikheev P.A., Sushchenko S.P., Tkachev R.V.* Estivation of High-Speed Performance of the Transport Protocol with the Mechanism of Forward Error Correction // Communications in Computer and Information Science. – 2017. – Vol. 700. – P. 259–268.

12. *Flach T., Dukkipati N., Terzis A., Raghavan B.,. Cheng Y., Cardwell N., Jain A., Hao Sh.,. Katz–Bassett E., and Govindan R.* Reducing Web Latency: the Virtue of Gentle Aggression // ACM SIGCOMM. – 2013.

13. *Forouzan B.* Data communications and networking. – Fourth edition, 2007.

14. *Herrero R.* Modeling and comparative analysis of Forward Error Correction in the context of multipath redundancy // Telecommunication Systems. Modelling, Analysis, Designand Management. – 2017. – Vol. 65. – No 4. –P. 783–794.

15. https://www.chromium.org/quic

16. *Langley A., Riddoch A., Wilk A., Vicente A., Krasic Ch., Zhang D., Yang F., Kouranov F., Swett I., Iyengar J., Bailey J., Dorfman J., Roskind J., Kulik J., Westin P., Tenneti R., Shade R., Hamilton R., Vasiliev V., Chang W.-T., Shi Z.* The QUIC Transport Protocol: Design and Internet-Scale Deployment // SIGCOMM '17, August, 2017, Los Angeles, CA, USA. – P. 183−196.

ИССЛЕДОВАНИЕ МЕТОДОВ КЛАСТЕРИЗАЦИИ ДЛЯ АНАЛИЗА ЗАДАЧ В БОЛЬШИХ ВЫЧИСЛИТЕЛЬНЫХ СЕТЯХ

А.И. Шкабара

Томский политехнический университет

Введение

В современном мире одним из наиболее актуальных видов физических исследований являются эксперименты по физике высоких энергий, вносящие неоспоримый вклад в фундаментальную науку. В результате таких исследований образуются огромное число данных, которые фиксируются детекторами ускорителя заряженных частиц. Обработка такого объема данных требует больших вычислительных мощностей, поэтому и создаются распределенные системы обработки данных, в которые входят большое количество суперкомпьютеров. Однако использование такого оборудования обходится дорого, поэтому необходима организация рационального планирования обработки данных, во избежание простоя оборудования и неравномерного распределения задач по обработки, среди суперкомпьютеров системы обработки данных. Очевидно, что определение времени обработки данных являются ключевой задачей для организации системы планирования. Решением такой задачи является система, способная предсказывать время обработки данных.

Данное исследование является частью работы по предсказыванию длительности выполнения заданий в большой вычислительной сети Большого Адронного Коллайдера. Было высказано предположение, что есть закономерности, которые влияют на вы-

полнение задач и время их окончания. Предполагается, что предварительный кластерный анализ этих задач поможет предсказывать длительность обработки точнее.

Цель данной работы – поиск закономерностей, влияющих на предсказание длительности выполнения задач в цепочках с помощью предварительной кластеризации.

1. Предсказание длительности выполнения задач по данным с LHC

Современные грид-инфраструктуры обеспечивают интеграцию аппаратурных и программных ресурсов, находящихся в разных организациях в масштабах стран, регионов, континентов в единую вычислительную среду, позволяющую решать задачи по обработке сверхбольших объемов данных, чего в настоящее время невозможно достичь в локальных вычислительных центрах.

Наиболее выделяющиеся результаты по работе с распределенными вычислениями продемонстрированы в проекте WLCG (Worldwide LHC Computing Grid или Всемирный грид для Большого адронного коллайдера), который базируется в ЦЕРН при обработке данных с экспериментов на LHC (Large Hadron Collider). Суть данной модели распределенных вычислений состоит в том, что вся информация с детекторов коллайдера после первичной обработки должна направляться для дальнейшей обработки и анализа в региональные центры разных уровней, где происходит дальнейшая фильтрация, обработка и сохранение необходимых данных [1].

Столкновения частиц в экспериментах описываются определенным набором данных (dataset), содержащим информацию о каждом столкновении. В один набор входит информация о нескольких часах работы коллайдера – один цикл. В эти данные входят дата, время работы коллайдера, настройки, с которыми проходили данные эксперименты, а также описания независимых событий – столкновений, собранных с различных датчиков детектора.

Физики присылают запрос на работу с определенным набором данных, они описывают действия, необходимые для получения результатов. Часто эти действия содержат в себе несколько задач (Task), которые могут выполняться друг за другом или параллельно, а также содержать в себе работу над несколькими наборами данных одновременно.

Задача разбивается на задания (Job), каждое задание состоит из нескольких событий (Event). Каждое задание содержит столько событий, сколько один узел может обработать одновременно. Событие описывает одно столкновение, т.к. все данные из набора независимы друг от друга.

Сначала, задача делит набор данных на задания. Эти задания распределяются между узлами Распределенной Сети и начинают выполнение. Как только все события в задании обработаны, эти данные готовы перейти к следующей задаче, не дожидаясь остальных. Следующая задача может перераспределить количество необходимых событий внутри задания. И как только достаточное количество событий готово, чтобы укомплектовать новое задание, задача начинает свое выполнение.

Отсюда следует, что задачи могут выполняться не последовательно, а с наложением, которое зависит от заданий и времени их выполнения или параллельно. К примеру, сравнение или слияние двух наборов данных должны иметь два готовых обработанных набора, соответственно очередность выполнения напрямую зависит от типа поставленной задачи. То, насколько задачи перекрывают друг друга зависит от параметров цепочки задач, которые мы можем дать модели машинного обучения.

Как правило, задачи, которые физики ставят для обработки, являются стандартными. По текущей цепочке и ее параметрам, а также по имеющимся уже примерам цепочек задач с определенным временем выполнения мы можем предсказать время работы текущей цепочки. В результате предыдущих исследований был разработан алгоритм, предсказывающий длительность заданий по параметрам, описывающим столкновения, а также длительность выполнения цепочек задач.

Только когда все 100% заданий обработаны, задачу можно считать успешно выполненной. Часто задача состоит из более чем 1000 заданий и если какие-то события вызывают ошибки, то одно необработанное задание приводит к тому, что весь таск считается необработанным. Также, важно уметь предсказывать данные аномалии и устранять баги. Решением данной проблемы может послужить предварительная кластеризация данных журналов заданий.

Кластеризация в анализе данных приобретает ценность тогда, когда она выступает одним из этапов анализа данных, построения законченного аналитического решения. Часто легче выделить группы схожих объектов, изучить их особенности и построить для каждой группы отдельную модель, чем создавать одну общую модель на всех данных [2].

2. Обзор методов кластерного анализа в литературе

В настоящее время вопросам кластеризации посвящено большое множество исследований, т.к. данный метод анализа данных подходит для любых сфер и типов задач. Многие исследования сравнивают имеющиеся методы для конкретного набора данных. Например, в [3] исследовались финансовые данные и в этом случае метод иерархических деревьев DBHT и k-medoids показывают лучшие характеристики, но последний больше зависит от шума чем DBHT.

Показатели точности прогнозирования в [4] показывают, что кластерная регрессия дает чрезвычайно точные прогнозы, но несколько нестабильные кластеры, в то время как k-means дает более стабильные кластеры, но очень плохие прогнозы в отдельных кластерах на наборе данных энергопотребления.

В [5] выполнено сравнение 18 методов кластеризации для анализа многомерных данных массы и потока цитометрии. В данном случае наиболее оптимальным оказался метод FlowSOM с выбором числа кластеров вручную.

Значительные улучшения в кластеризации можно наблюдать, если использовать подходящий метод снижения размерности. В [6] рассмотрен принцип предварительной кластеризации базы данных отпечатков пальцев для более скорого поиска совпадений в базе в следующем порядке: На первом этапе происходит подготовка базы изображений отпечатков для дальнейшей обработки. Каждое изображение заменяется на числовой вектор признаков, с которым удобнее работать. Далее база из векторов подвергается кластеризации, затем нейронная сеть глубокого обучения учится сопоставлять отпечатки пальца нужным кластерам. В процессе кластеризации использовались методы снижения размерности PCA и поле направлений, и методы кластеризации K-means и агломеративный иерархический. Для данных изображений наиболее точные результаты дал метод главных компонент вместе с алгоритмом K-means.

Большое влияние на точность оказывает тип признаков. Категориальные признаки могут значительно улучшить алгоритм или, по крайней мере снизить негативное влияние на результат за счёт удаления категориального столбца. Так, например, в [6] предложен пространственный структурно основанный метод для кластеризации категориальных признаков, что дает выигрыш в точности кластеризации на разных наборах данных. Также в [7] был представлен алгоритм k-prototypes для кластеризации больших наборов реальных данных. Сохраняя эффективность алгоритма k-средних, и убирая его ограничения на только числовые данные было показано, что он эффективен для кластеризации больших наборов данных со смешанными числовыми и категориальными значениями.

Практически для всех данных процесс кластеризации сводится к следующим универсальным шагам: предварительная обработка, снижение размерности, кластеризация, визуализация, оценка результатов. Но, как показывает обзор, для каждого набора данных и цели анализа необходим свой уникальный подход с подобранными параметрами

алгоритмов кластеризации, а также грамотная предварительная обработка и очистка данных.

3. Подготовка данных

Применение кластерного анализа в общем виде сводится к следующим этапам:

1. Отбор выборки объектов для кластеризации.
2. Определение множества переменных, по которым будут оцениваться объекты в выборке. При необходимости – нормализация значений переменных.
3. Вычисление значений меры сходства между объектами.
4. Применение метода кластерного анализа для создания групп сходных объектов (кластеров).
5. Представление результатов анализа.

После получения и анализа результатов возможна корректировка выбранной метрики и метода кластеризации до получения оптимального результата.

Для исследования был взят набор данных с информацией о заданиях, поставленных на выполнение в WLCG с 04.06.2017 по 17.06.2017. В таблице описываются основные характеристики заданий, такие как время выполнения, даты начала и конца задачи, количество ядер, вычислительный узел и т.д. Размер исходной таблицы 10517 × 46.

В данном наборе данных во многих строках содержатся пустые ячейки, при этом, некоторые столбцы более чем на 40% состоят из пустых значений. Предположение, что пустые строки систематически зависят от типа вычислительной машины «computingsite» оказалось неверно. В результате было принято решение удалить такие столбцы. Также были сброшены столбцы, не несущие никакой значимой информации для анализа. Это имя задания и его статус.

Каждое задание имеет начальный и конечный приоритет. Причем значение этого приоритета меняется с ходом выполнения задания. Начальный приоритет «assignedpriority» содержит значения [320, 901], низкий и высокий соответственно. Чем дольше выполняется задача, тем выше становится ее приоритет. Конечный приоритет «currentpriority» уже содержит следующие значения [320, 322, 321, 5000, 323, 324, 325, 338, 336, 340, 331, 329, 332, 342, 337, 333, 326, 344, 328, 343, 339, 4990, 330, 334, 335, 341, 327, 901, 902, 903, 904, 900]. Из этого можно сделать вывод, что приоритет может повлиять на модель обучения. В качестве нового столбца была выбрана разница приоритетов.

Было высказано предположение, что если задача поставлена в выходные, то она срочная и у нее присутствует большая вероятность сбоев. Также существует предположение, что члены физической группы, которые ставят задачи рано в начале дня и поздно в конце имеют большую вероятность ошибки. Время начала и время конца выполнения задания в таблице не несут большой информативной ценности для анализа. Гораздо больше смысла будет нести информация о длительности выполнения всего задания (эта информация уже содержится в таблице), а также о времени от начала дня, когда выполняется это задание, ведь тогда можно будет проследить какие-либо закономерности в обработке заданий. Также на скорость выполнения может повлиять день недели. В результате были созданы столбцы с количеством секунд с начала дня для начала и окончания выполнения задания и днем недели начала и окончания выполнения задания.

Столбец «computingsite» содержит текстовую информацию о вычислительном узле, поэтому он был переведен в категориальный вид. Т.к. алгоритмы кластеризации опираются на расстояние между признаками, то обработка категориальных признаков может ввести его в заблуждение и негативно повлиять на результат. Поэтому все категориальные признаки, такие как день недели начала и конца обработки и вычислительный узел задания, были трансформированы в отдельные столбцы.

В этап подготовки данных также вошли заполнение некоторых отсутствующих значений, приведение таблицы к одному типу переменных, проверка таблицы на мультиколлинеарность, а также нормализация всей таблицы. В результате предварительной обработки данных получилась таблица размерностью 10517 × 60.

4. Снижение размерности пространства

На практике при решении задач анализа данных и машинного обучения информация об объектах представлена в виде сложного многомерного массива данных. Таким образом, возникает необходимость извлечения из входных многомерных данных набора признаков, которые будут наиболее информативны с точки зрения дальнейшего решения задачи. Любые многомерные данные всегда можно представить в виде вектора чисел. Такие вектора обычно имеют большую длину в виду многомерности пространства, а некоторые признаки, которые содержатся в данном векторе, не всегда информативны. Поэтому решается задача понижения размерности пространства признаков, которое описывает входные данные, с целью получения относительно компактного множества информативных признаков [8]. Для точной визуализации необходимо уменьшить размерность пространства данных с 60 столбцов до двухмерного пространства, доступного для визуализации точек и кластеров.

PCA (Principal Component Analysis) – метод главных компонент. В результате преобразования пространства признаков было выделено от 2 до 5 главных компонент. На рис. 1 приведен график распределения точек для 2 главных компонент. На графике видно явное разделение на несколько кластеров. Особенно это видно на горизонтальной оси. Признаки, которые больше влияют на компоненты это приоритет, 3 вычислительный узел, а также пятница, как день недели начала и конца выполнения задания.

T-SNE (T-distributed Stochastic Neighbor Embedding) – это второй выбранный способ снижения размерности. Этот алгоритм проецирует каждый объект высокой размерности в заданную размерность таким образом, что похожие объекты проецируются близко расположенными точками, а непохожие точки проецируются расположенными далеко друг от друга [8]. На рис. 2 представлена визуализация данного метода для полученного набора данных. Сравнив графики можно увидеть, что метод главных компонент гораздо лучше выделяет крупные кластеры, чем T-SNE.

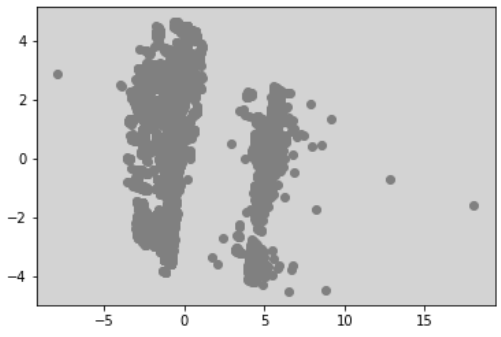


Рис. 1. Визуализация данных с применением PCA

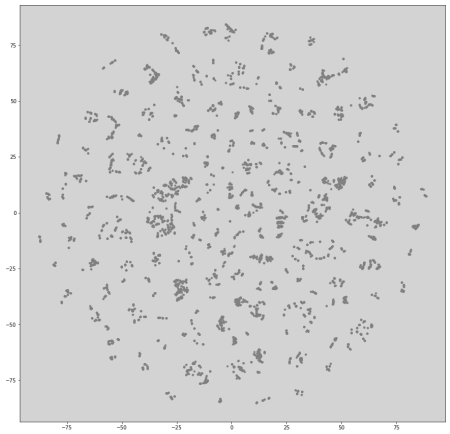


Рис. 2. Визуализация данных с применением T-SNE

5. Кластерный анализ

K-means – самый быстродействующий и часто самый эффективный за счет своей простоты алгоритм. Это итеративный алгоритм, который ищет центры кластеров и точ-

ки, которые находятся ближе всего к ним. Алгоритм k-means обычно находит набор стабильных кластеров за несколько десятков итераций. На рис. 3, 4 можно заметить, что данный алгоритм не лучшим образом подходит для геометрии именно этих данных. Видно, что для данного метода деление очень условно, особенно на графике с T-SNE.

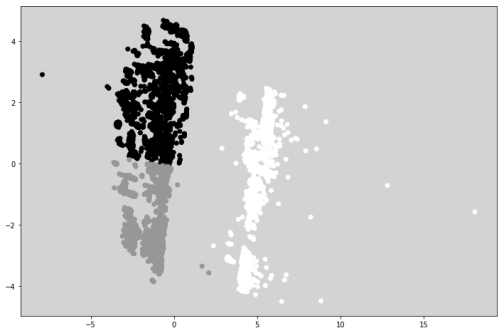


Рис. 3. Визуализация K-means для PCA

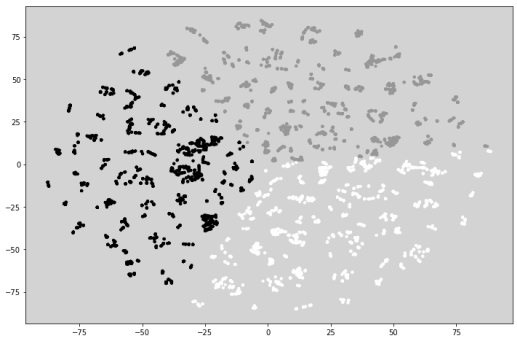


Рис. 4. Визуализация K-means для T-SNE

Иерархическая кластеризация (дендограмма)

В данной кластеризации строится иерархическое дерево, которое объединяет точки по схожим признакам, в итоге образуя кластеры, которые представлены на рис. 5, 6.

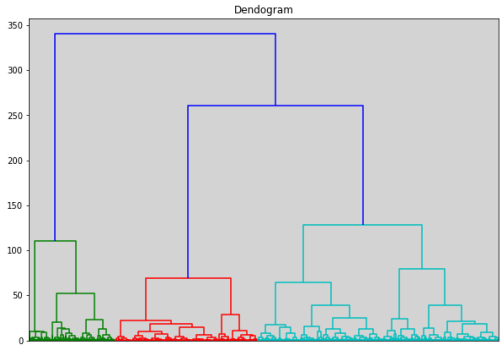


Рис. 5. Дендограмма для PCA

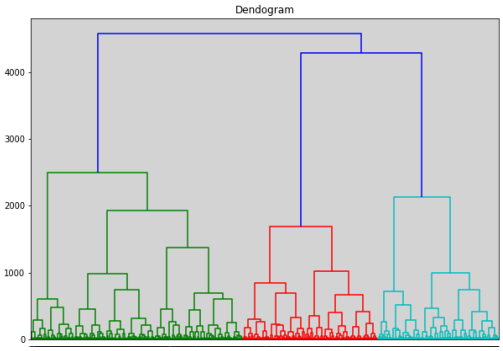


Рис. 6. Дендограмма для T-SNE

Чтобы визуализировать данное разделение, необходимо сделать срез данной дендрограммы относительно определенного количества кластеров.

Для PCA данный алгоритм показывает более хорошие результаты по сравнению с k-means. Для T-SNE кластеры были выделены немного лучше, но все равно видно, что реальная численность кластеров гораздо больше, что бесполезно для исследования, т.к. нас интересуют не более 10 кластеров

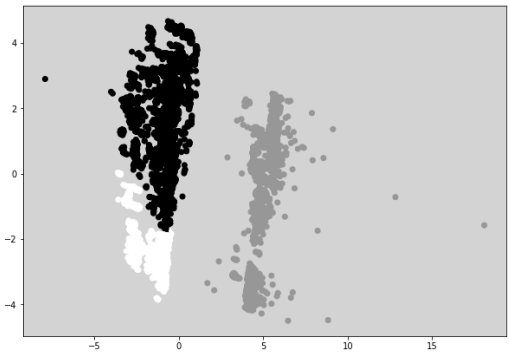


Рис. 7. Иерархическая кластеризация для PCA

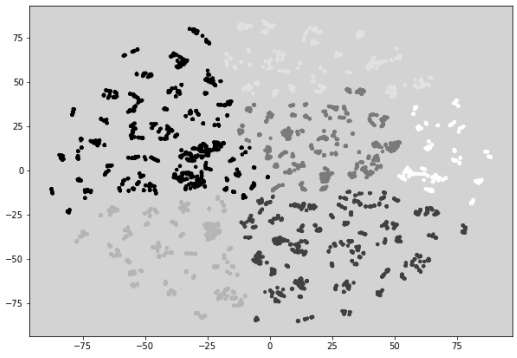


Рис. 8. Иерархическая кластеризация для T-SNE

DBSCAN (Density-Based Spatial Clustering of Applications with Noise, плотностной алгоритм пространственной кластеризации с присутствием шума) – популярный алгоритм кластеризации, используемый в анализе данных в качестве одной из замен метода k-средних. Если текущий объект – окруженная точка, то все объекты, достижимые по плотности от текущего объекта, соединяем в новый кластер. В противном случае, если объект не является окруженной точкой и не достижим по плотности ни от какого объекта, то текущий объект – выброс [8].

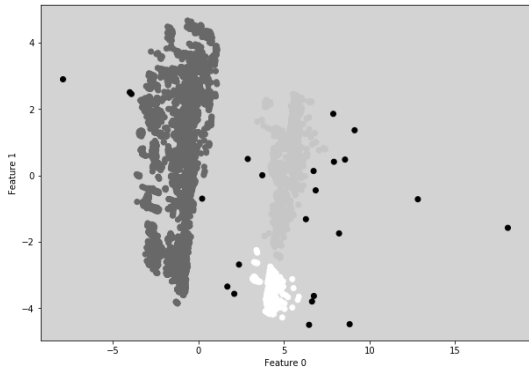


Рис. 9. DBSCAN для PCA

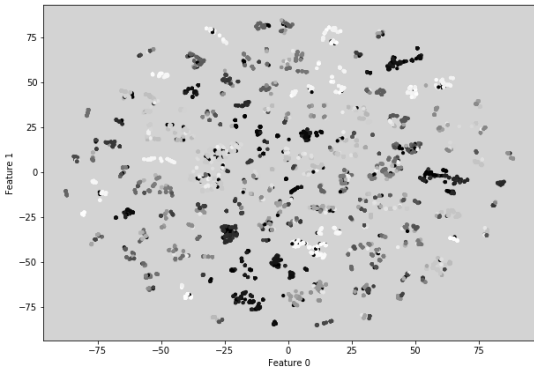


Рис. 10. DBSCAN для T-SNE

На рис. 9 данный метод показывает наиболее оптимальный результат для PCA, он более всего совпадает с визуальным делением данных, а также очень точно обнаруживает шум. Но и для T-SNE на рис. 10 он работает достаточно точно, т.к. выделяет очень большое число кластеров.

Заключение

В настоящей работе рассмотрена была проведена подготовка данных, уменьшение размерности и анализ подходов кластерного анализа. Наилучший результат для данного набора данных показывает связка PCA+DBSCAN.

Также подтвердилось предположение о том, что день недели влияет на время выполнения задания. Удалось обнаружить признаки, влияние которых на итоговую длительность выполнения заданий больше.

В планах использовать результат данного метода, чтобы на его основании ввести дополнительный столбец с результатами кластеризации и обучать следующие таблицы уже на модели для обучения с учителем, планируя, что это может помочь сделать дальнейшие предсказания более точными. Если результаты анализа будут коррелировать с

длительностью выполнения задач, то данный подход повысит качество, если нет, то будут исследованы другие подходы

ЛИТЕРАТУРА

1. *Кореньков В.В.* Распределенная система для обработки, хранения и анализа экспериментальных данных Большого адронного коллайдера // Современные информационные технологии и ИТ-образование. – 2012. – №. 8. С. 8–20.

2. *Замятин А.В.* Введение в интеллектуальный анализ данных: учебное пособие. – Томск: Издательский дом Томского государственного университета, 2016. – 120 с.

3. *Musmeci N., Aste T., Di Matteo T.* Relation between financial market structure and the real economy: comparison between clustering methods // PLoS ONE. – 2015. – 10 (3): e0116201.

4. *Hsu D.* Comparison of integrated clustering methods for accurate and stable prediction of building energy consumption data //Applied energy. – 2015. – Т. 160. – С. 153–163.

5. *Weber L.M., Robinson M.D.* Comparison of clustering methods for high‐dimensional single‐cell flow and mass cytometry data //Cytometry Part A. – 2016. – Т. 89. – №. 12. – С. 1084–1096.

6. *Qian Y. et al.* Space structure and clustering of categorical data // IEEE transactions on neural networks and learning systems. – 2015. – Т. 27. – №. 10. – С. 2047–2059.

7. *Huang Z.* Clustering large data sets with mixed numeric and categorical values // Proceedings of the 1st pacific-asia conference on knowledge discovery and data mining,(PAKDD). – 1997. – С. 21–34.

8. *Береснев Д.В., Шараев Е.В.* Анализ методов понижения размерности пространства // Комп'ютерні та інформаційні системи і технології. – 2019.

РАЗРАБОТКА И ОПТИМИЗАЦИЯ МОДЕЛИ БИЗНЕС-ПРОЦЕССА ПОДДЕРЖКИ СИСТЕМЫ ЭЛЕКТРОННОГО ДОКУМЕНТООБОРОТА

М.М. Зимина, А.И. Кущазли, И.А. Гудкова
Российский университет дружбы народов

Введение

В настоящее время всё больше и больше компаний переходит на системы электронного документооборота (СЭД). Так как в каждой организации происходит работа с документами разного типа и разной важности, СЭД является одним из методов повышения эффективности внутренней работы компании. Да и в целом, информационные технологии развиваются и привлекают немалое внимание к себе и становятся значимыми как в нашей жизни, так и в работе организации.

Современные системы электронного документооборота работают по законодательству, соблюдая все правила и существующие законы, именно это даёт право работать с юридически значимыми документами. Также возможен обмен этими документами за пределами организации, т.е. с контрагентами, использование электронной подписи, правильное хранение документов и т.д.

В данной работе с помощью нотации для моделирования бизнес-процессов BPMN проводится разработка верхнеуровневой системы электронного документооборота, разбираются основные функции системы поддержки работы СЭД. Также рассматривается жизненный цикл автоматизированных информационных систем.

1. Постановка задачи

Рассмотрим нотацию для моделирования бизнес-процессов BPMN. Она используется для создания стандартного набора условных обозначений, понятных всем бизнес-пользователям. Одной из важнейших особенностей этой нотации является допустимость установления единичного (однозначного) соответствия между взаимодействующими элементами графической нотации и исполняемого языка описания бизнес-процессов.

Документ является основным ресурсом, в котором содержится та или иная информация, которая обеспечивает информационную поддержку в качестве принятия управленческих решений. Также документ является способом оформления хозяйственных

операций в организации. Сопровождение ведения бизнес-процессов происходит за счёт документации. Документооборот – процесс движения документов, который отражает факт хозяйственной деятельности с момента создания документа или получения до завершения исполнения или отправки.

Электронный документооборот (ЭДО) – это способ организации работы с документами, при котором основная масса документов используется в электронном виде и хранится централизованно. СЭД – это компьютерная программа (программное обеспечение, система), которая позволяет организовать работу с электронными документами (создание, хранение, изменение, поиск), а также взаимодействие между сотрудниками (передачу документов, выдачу заданий, отправку уведомлений и т.п. [1,2]).

Следовательно, складывая всё воедино, можно сделать вывод, что любой процесс в СЭД можно описать как самостоятельный бизнес-процесс. Как было сказано ранее, в данной работе проводится разработка верхнеуровневых бизнес-процессов. Верхнеуровневая модель процессов, или же модель процессов «верхнего уровня», строится для того, чтобы целостно понимать структуру бизнеса заказчика и не упустить иерархичности и взаимосвязи процессов.

Модель детального бизнес-процесса представляет собой логику реализации рассматриваемого этапа жизненного цикла процесса. Модель отражает набор действий, которые должны быть выполнены по заданным правилам для получения заданного результата. Условия выполнения действий (бизнес-функций), а также их исходы отражаются посредством событий. Иными словами, модель детального бизнес-процесса представляет собой последовательности событий и бизнес-функций, отражающих правила достижения заданного результата.

2. Системы электронного документооборота

Как было сказано ранее, СЭД – это компьютерная программа, позволяющая организовать работу с любым видом электронного документа. Немаловажный факт, что СЭД имеет в себе много подразделений, ориентированных на определенные задачи организации.

Систему электронного документооборота нередко называют EDMS (Electronic Document Management Systems), т.е. система управления электронными документами. На рис. 1 рассмотрены виды информационных систем управления документационным обеспечением [2].

Рис. 1. Виды информационных систем управления документационным обеспечением

Внедрение информационных систем управления документационным обеспечением (ИСУД) предполагает осуществление автоматизации управления интеллектуальными

активами и бизнес-процессами предприятия, что определяет успешность его деятельности.

У каждой ИСУД существует свой определённый функционал, под который она разработана. К примеру, информационные системы, ориентированные на бизнес-процессы предназначены для специфических вертикальных и горизонтальных приложений, в то время как ИСУД управления потоками работ применяются в целях ускорения бизнес-процессов, увеличения эффективности и степени контролируемости процессов в организации.

3. Жизненный цикл автоматизированных информационных систем

Проектирование информационной системы – процесс, который описывает создание и сопровождение системы в виде жизненного цикла (ЖЦ), т.е. в виде какой-то последовательности стадий и выполняемых на них процессов действий. На каждом этапе определяются состав и последовательность выполняемых работ и действий для получения результата, отдельные методы, и средства, необходимые для выполнения достижения этих целей, роли и обязанности участников. Такое описание жизненного цикла дает возможность правильной и эффективной организации процесса коллективной работы с последующим обеспечением управления над этим процессом [3].

Жизненный цикл начинается с момента создания и заканчивается выходом из эксплуатации. Также в модели отражаются моменты, при которых возникла первая потребность в информационной системе и как эта потребность заканчивается.

Действия, задачи и процессы в модели жизненного цикла представляют собой структуру. Они осуществляются в ходе разработки, когда система начинает функционировать и уже с последующим сопровождением продукта, вплоть до выхода из эксплуатации при завершении потребности в использовании этой системы.

На рис. 2 рассмотрен бизнес-процесс проектирования информационной системы в нотации BPMN.

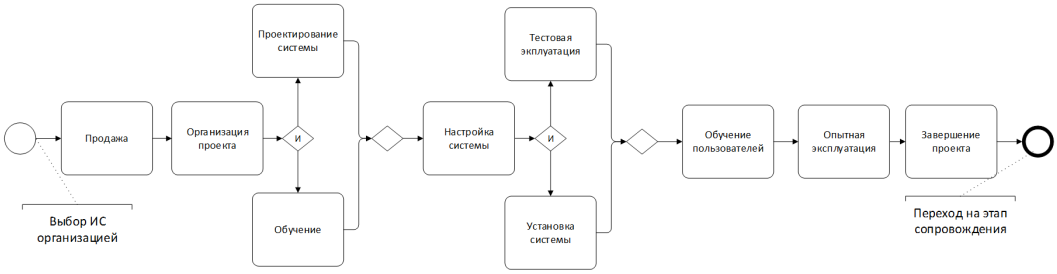


Рис. 2. Проектирование информационной системы

Как видно из рис. 2, инициация процесса происходит с выбора ИС организацией, т.е. перед решением о внедрении СЭД компания должна определиться с тем, что ей надо.

В данном случае говорится о внедрении СЭД. Весь жизненный цикл состоит из десяти этапов, из которых два могут происходить параллельно друг другу. Шлюз «и» предполагает параллельное исполнение событий (этапов), а «пустой» шлюз говорит о слиянии действий (этапов). Важно понимать, что в данном случае другое действие не начнётся, пока два параллельных события не будут завершены.

4. Участники бизнес-процесса внедрения СЭД

Процесс внедрения СЭД является очень трудоёмким и ресурсозатратным. Этот процесс называется «проектом». Одним из основных участников проекта является заказчик – лицо, которое становится владельцем результата проекта. Он участвует в со-

здании плана, выделяет средства на осуществление проекта, ведёт контроль над деятельностью проектной команды.

Поставщик – лицо, ответственное за обеспечение проекта ресурсами. Ответственное за выполнение работ лицо именуется подрядчиком. Таким образом, подрядчиком выступает организация, персонал которой является проектной командой.

В процессе разработки и эксплуатации системы участвует определённый круг лиц (представители заказчика и разработчика), заинтересованных в успешной реализации проекта. В этом процессе между ними распределяются роли, за каждой из которых закрепляется определенный набор функций (обязанностей).

Необходимо отметить, что проект является сложной системой, в которой взаимодействует и взаимозависит большое количество процессов и участников.

На рис. 3 рассмотри состав участников проекта.

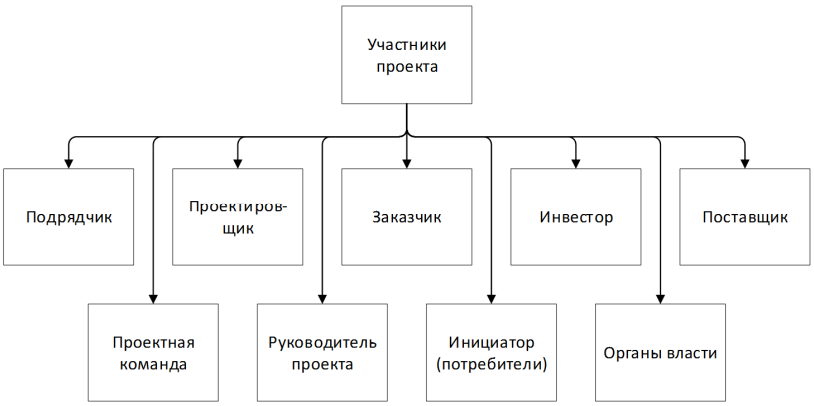


Рис. 3. Участники бизнес-процесса внедрения СЭД

Важно понимать, что управление проектами – это искусство, включающее ряд процессов управления: планирования, контроля, организации, руководства, координации ресурсов и т.п., направленное на достижение целей проекта и удовлетворения его требований. Процесс управления проектами неразрывно связан с жизненным циклом проекта, соответственно и каждый процесс в управлении про разному значим и играет разную роль в каждой из фаз проекта. Некоторые процессы могут быть делегированы членам команды проекта, но ответственность однозначно лежит на руководителе проекта. Всё это также актуально для управления программами и портфелями проектов.

5. Построение модели бизнес-процесса внедрения СЭД

Этапы внедрения, описанные выше, продемонстрированы в верхнеуровневой модели бизнес-процесса в нотации BPMN.

Более подробно остановимся на этапе проектирования системы, т.к. именно на этом этапе происходит обследование бизнес-процессов организации. Анализ и сбор требований для дальнейшей работы заключается в анкетировании сотрудников, работу которых будет затрагивать информационная система. На данном этапе рассматривается порядок разработки проектных решений, описывающих вариант функционирования системы у заказчика.

Методология реинжиниринга бизнес-процессов, предложенная в данной работе, является лишь частью проекта по реинжинирингу и встраивается в методологию управления проектами в организации для формирования тиражируемой модели проекта по реинжинирингу бизнес-процессов [3,4].

Методология основывается на методологии PMI, соответственно, фазы проекта делятся на инициацию, где принимается решение о реализации проекта, планирование, мониторинг и контроль, реализацию и завершение проекта. Организационная методо-

логия – совокупность нормативных документов банка, регламентирующих проектную деятельность и шаблоны проектной документации. Анализ и проектирование будущей системы изображены на рис. 4.

Как видно из рис. 4, инициация бизнес-процесса начинается от руководителя проекта от исполнителя с необходимости провести исследование процессов, т.к. будущая система должна быть «наложена» на описанные бизнес-процессы организации-заказчика. Далее необходимо на уровне руководителей проекта провести согласование необходимых работ и плана исследования. Шлюз с маркером «и» говорит о параллельном исполнении событий. Важно помнить, что действие (активность), завершающее параллельное исполнение, не наступит, пока все действия в цепи не прейдут к своему завершению. После согласований руководителей основной пласт работы ложится на плечи аналитика от исполнителя. Ему необходимо провести анализ проходящих в организации бизнес-процессов. Но перед этим аналитик должен собрать и проанализировать данные и требования по проекту

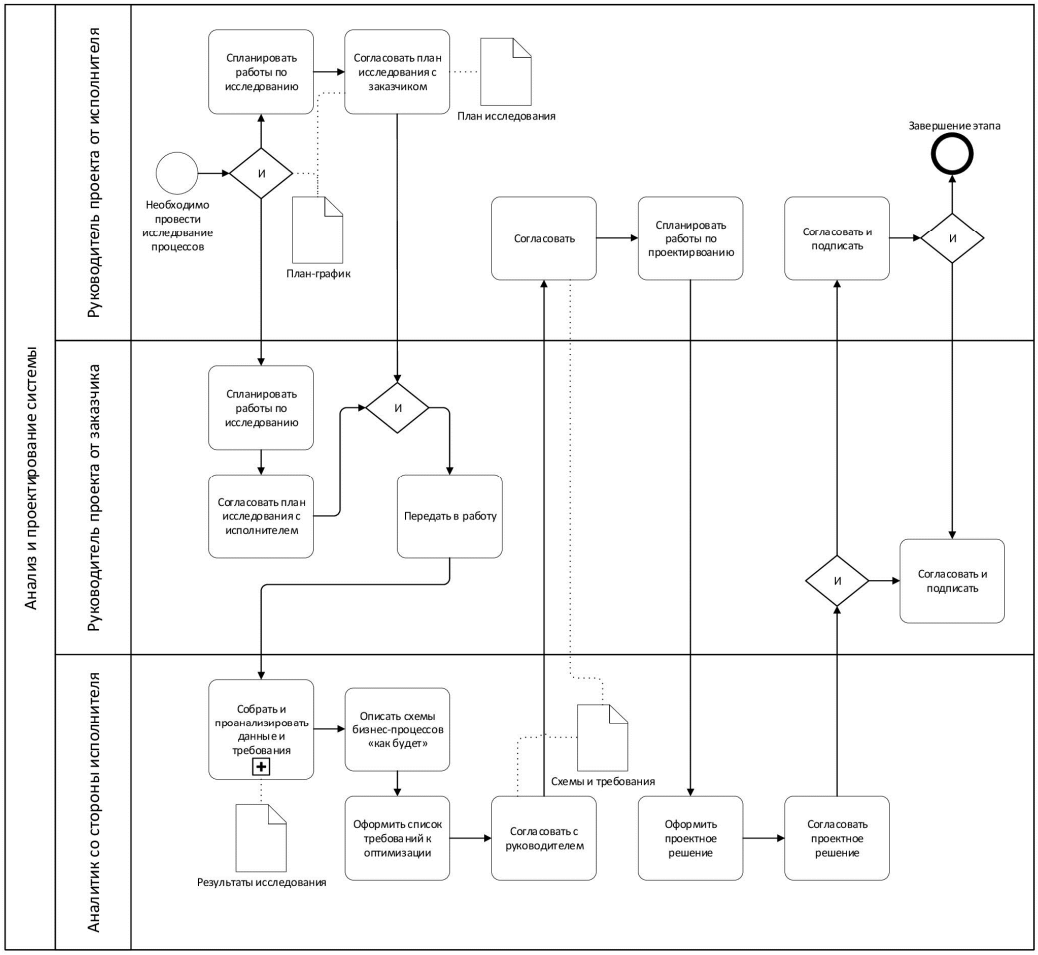


Рис. 4. Анализ и проектирование системы

Далее происходит согласование на уровне руководителей двух сторон, после подписания этап приходит к своему завершению.

Рассмотрим фазу реализации проекта (рис. 5). Ввиду того, что проекты по реинжинирингу бизнес-процессов сходны по реализации, имеет смысл создать полноценную модель проекта по реинжинирингу бизнес-процессов.

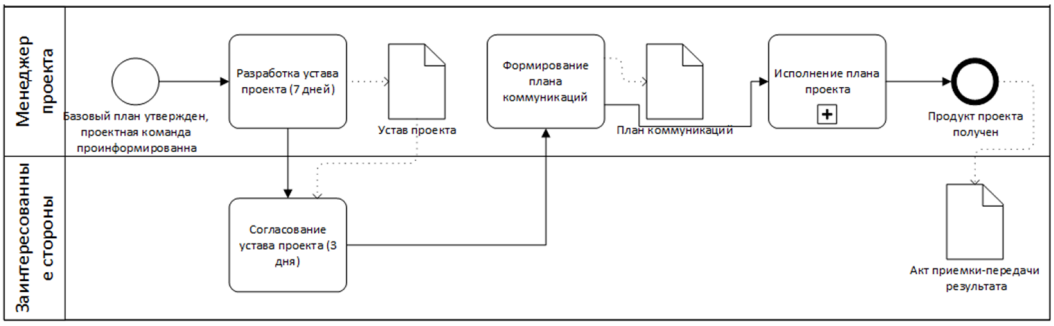


Рис. 5. Фаза реализации проекта

Методология реинжиниринга, представляя собой последовательность действий по совершенствованию бизнес-процессов, и реализуемая проектной командой, является планом по исполнению проекта на фазе реализации. Оценка уровня зрелости играет важную роль на стадии инициации и стадии реализации.

Далее рассмотрим управление проектом (рис. 6), а именно – верхний уровень процесса реализации проекта. Стадия инициации различается в зависимости от уровня зрелости управления процессами (рис. 7). Учитывая знания о системе управления процессами в банке, единственным источником такого рода инициатив может выступать только отдел бизнес-процессов и технологий. В редких случаях инициатива может исходить от правления организации.

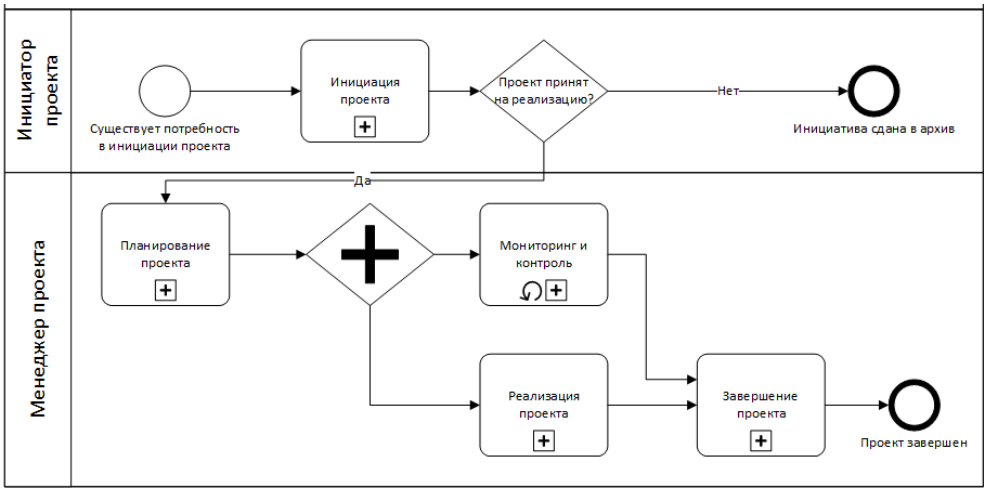


Рис. 6. Верхний уровень процесса реализации проекта.

К инициативе можно будет вернуться в момент, когда проект получит более глубокое обоснование, либо когда возникнет более очевидная необходимость в реализации проекта.

Планирование проекта представляет собой создание детального плана проекта. Ввиду отсутствия децентрализации властных полномочий, данный план одобряет председатель правления банка. Мониторинг и контроль проводится параллельно реализации проекта.

По окончании всех мероприятий процесс повторяется до тех пор, пока не будет создан продукт. Мониторинг и контроль осуществляется менеджером и администратором проекта, в то время как в реализации основной действующей силой является проектная команда. На завершающей фазе продукт передается владельцу и архивируется вся информация по проекту.

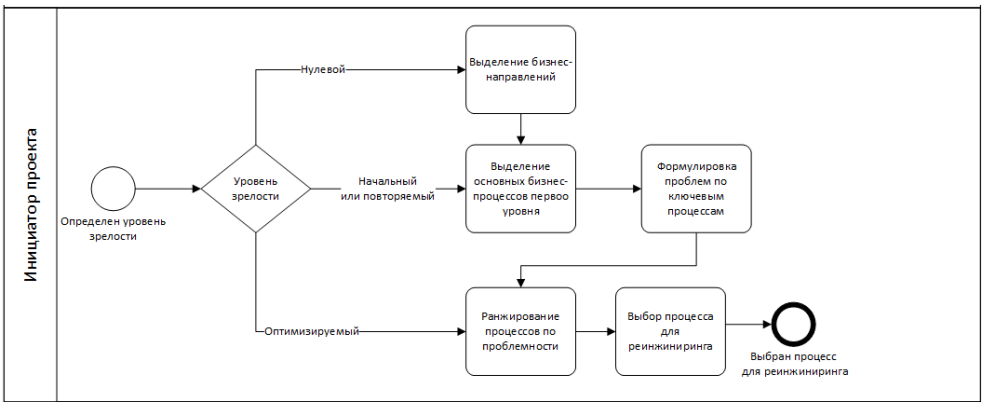


Рис. 7. Приоритезация процессов.

Перед представлением инициативы, отдел должен проанализировать зрелость управления процессами в организации (рис. 7), чтобы сформулировать примерный план действий. Проект по реинжинирингу бизнес-процессов также предполагает совершенствование системы управления процессами.

Соответственно, чем ниже уровень зрелости, тем больше внимание уделяется доведению системы до уровня, когда реинжиниринг является осуществимым. На низших уровнях зрелости необходимо проводить базовые мероприятия по формированию системы бизнес-процессов.

Необходимо выделить основные бизнес-направления и продукты и определить основные бизнес-процессы верхнего уровня. Это позволяет оценить проблемы процессов и, основываясь на степени проблемности, определить приоритетность реинжиниринга того или иного бизнес-процесса.

На рис. 7 видно, что на нулевом и начальном уровнях зрелости, совершенствование системы управление процессами занимает большую часть проекта. Оценка уровня зрелости необходима перед инициацией каждого подобного проекта, т.к. система развивается, а отдел бизнес-процессов работает над созданием методологии параллельно с реализацией проекта.

После проведения работ по приоритезации процессов и выделения процесса для реинжиниринга, отчет по окончанию данной деятельности, как и обоснование приоритетности реинжиниринга конкретного процесса, являются основой для презентации инициативы перед проектным комитетом.

6. Основные функции системы поддержки работы СЭД

Как и любая информационная система после внедрения в организацию требует развитие и сопровождение. Случаются ситуации, когда при использовании системы, происходят ошибки. Такие ошибки могут быть разного рода и их возникновение зависит от различных факторов, таких как неправильное использование системы, сбой системы и т.д. Для того, чтобы система активно функционировала и для оперативного решения проблем, для многих информационных систем существует служба поддержки, куда автоматизированные пользователи могут обратиться.

Далее рассмотрим классификации ошибок и скорость их решения.

Таблица 1

Степень воздействия

Степень воздействия	Описание
Высокая	Серьезные сбои, нарушающие использование всей системы в целом или отдельных ее служб. Сбой может охватывать работу, как отдельных ключевых пользователей, так и целые группы.
Средняя	Сбой не останавливает работу системы в целом, но часть функций приложения работает не-

Степень воздействия	Описание
	корректно. Нарушена функциональность важных бизнес-процессов заказчика.
Низкая	Сбой, влияющий на удобство работы с системой, снижающее эффективность ее использования

Степень воздействия (табл. 1) – показатель, характеризующий меру отклонения от нормального уровня использования программных продуктов, количество и статус пользователей и бизнес-процессов, подвергшихся воздействию в связи с инцидентом.

Подобные ситуации называются «инцидент». Инцидент – любое событие (сбой), не являющееся частью нормального функционирования системы или произведенных настроек, которое не позволяет использовать программный продукт или накладывает ограничения на его использование.

Таблица 2

Срочность разрешения

Срочность разрешения	Описание
Высокая	Решение требуется в минимальные сроки (в течение 36 часов).
Средняя	Решение потребуется в ближайшее время (в течение 1-2 недели).
Низкая	Решение потребуется в будущем (в течение 1-3 месяцев).

Срочность разрешения (табл. 2) – показатель характеризует приемлемую задержку разрешения инцидента с точки зрения наступления его воздействия на текущие задачи пользователей и бизнес-процессы.

Таблица 3

Приоритет инцидента

Приоритет инцидента		Степень воздействия		
		Высокая	Средняя	Низкая
Срочность	Высокая	Критический	Высокий	Средний
	Средняя	Высокий	Средний	Низкий
	Низкая	Средний	Низкий	Планируемый

Приоритет инцидента (табл. 3) – показатель, который определяется в соответствии с таблицей на основании степени воздействия и срочности.

7. Построение модели бизнес-процесса поддержки работы СЭД

Техническая поддержка заказных информационных систем (ЗИС) принципиально отличается от технической поддержки лицензионных программных продуктов, т.к. четко формализуемые проблемы, возникающие при сопровождении ЗИС, встречаются гораздо реже, чем при технической поддержке лицензионного ПО, а диапазон таких проблем гораздо шире.

Бывают случаи, когда некоторая функциональность системы, прописанная на этапе написания технического задания с требованиями, может устареть или потерять актуальность. Также при эксплуатации системы могут возникнуть новые требования к системе, отсюда следует, что по мере роста бизнеса заказчика происходит непрерывная доработка информационной системы, и, следовательно, приходится решать подчас нестандартные задачи в рамках развития и сопровождения системы. Рассмотрим пример типовой поддержки системы.

1-я линия поддержки (работа консультантов):

1. Прием обращений от конечных пользователей по телефону и эл. почте.
2. Регистрация, классификация и расстановка приоритетов запросов на основании принятых обращений.
3. Поиск известного решения в базе данных.
4. Определение и назначение исполнителей запроса.

5. Ведение базы знаний по инцидентам и типовым вопросам и ответам.

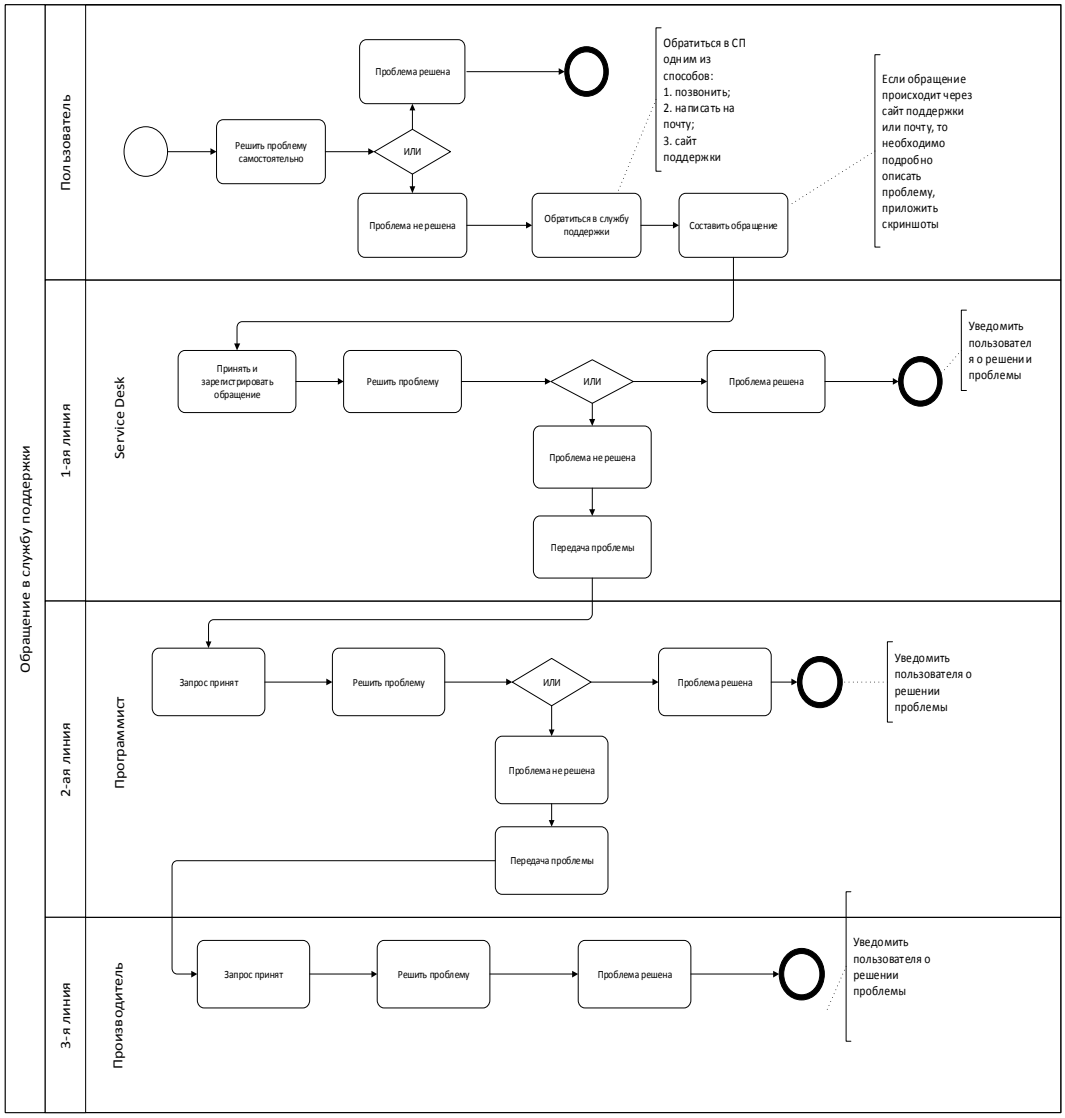
2-я линия поддержки (работа специалистов по предметной области):

1. Интервьюирование инициатора запроса с целью более четкого понимания внештатной ситуации и предпосылок к ее возникновению.
2. Восстановление работоспособности системы с использованием удаленного доступа, при помощи консультации по телефону или электронной почте.
3. Консультирование пользователей с целью устранения внештатной ситуации и недопущения повторения подобных ситуаций в будущем (в случае, если причиной внештатной ситуации послужили некорректные, с точки зрения системы, действия пользователей).
4. Выработка списка мероприятий для устранения послед. внештатной ситуации.
5. Оказание помощи пользователям при необходимости реализации нетиповых операций и бизнес-процессов с целью получения заданного результата.

3-я линия поддержки (работа разработчиков информационной системы):

1. Интервьюирование инициатора запроса с целью более четкого понимания требований (возникшей ошибки).
2. Подготовка краткого описания сути предлагаемых изменений и его согласование с инициатором запроса.
3. При необходимости проведения серьезной модернизации, составление технических заданий и согласование стоимости доработки системы.
4. Реализация изменений (устранение ошибки).
5. Тестирование.

Такой пример поддержки является стандартным для любой заказной информационной системы. Ниже представлен верхнеуровневый бизнес-процесс поддержки заказной информационной системы (рис. 8).



Заключение

В настоящей работе рассмотрена схема с помощью нотации для моделирования бизнес-процессов BPMN верхнеуровневой системы электронного документооборота. Рассмотрены основные функции системы поддержки работы СЭД. Также рассмотрен жизненный цикл автоматизированных информационных систем.

ЛИТЕРАТУРА

1. *Жеребенкова А.В.* Документооборот на предприятии. – М.: Вершина; Издание 2-е, перераб. и доп., 2011. – 384 с.

2. *Шеер А.-В.* Бизнес-процессы. Основные понятия. Теория. Методы. / Пер. с. англ. Н.А. Михайлова. Науч. ред. М.С. Каменнова, А.И. Громов. – М: Весть – Метатехнология. 2 изд. 1999г. – 173с.

3. *Чукарин А.В., Самуйлов К.Е., Яркина Н.В.* Бизнес-процессы и информационные технологии в управлении современной инфокоммуникационной компанией. – М.: Альпина Паблишер, 2016. – 512 с.

4. *Shin N, Jemella D.* Business Process Reengineering and Performance Improvement. The Case of Chase Manhattan Bank // Business Process Management Journal. – 2002. – Vol. 8. – No. 4. – P. 351–363.

РАЗРАБОТКА ПРОГРАММНОГО ОБЕСПЕЧЕНИЯ ДЛЯ ОЦЕНКИ ЭФФЕКТИВНОСТИ ИНВЕСТИРОВАНИЯ

Н.М. Килин, Л.Д. Шапиро
Томский государственный университет

Введение

Инвестирование, как одно из важнейших составляющих экономики, также нуждается в автоматизации и компьютеризации. Почти любое предприятие в определенный момент встает перед выбором: куда вложить средства, как это сделать с наибольшей для себя выгодой? Существующие методы оценки инвестирования позволяют сравнивать инвестиционные проекты, учесть множество факторов: тип инвестируемых средств, срок проекта, постоянно нарастающую инфляцию, изменение ценности денег во времени и т.д.

Целью данной работы является разработка программного приложения, которое позволит автоматизировать синхронизацию финансовых потоков предприятия, расчеты инвестиционных проектов, а также, используя технологии баз данных, хранить результаты для последующего сравнения и анализа.

1. Описание

Разрабатываемая программа должна реализовывать следующие функции:

- вводить финансовые данные, предоставляемые инвестором;
- на основе введенных данных рассчитать показатели для проекта финансирования;
- сравнивать несколько проектов финансирования в зависимости от выбранных критериев и на основе заложенных экспертных мнений выбирать наилучший для данной ситуации;
- оценить финансовый риск проекта;
- формировать отчетные документы.

Масштаб инвестиционной деятельности в экономике служит признаком ее качественного роста. Выделяют 3 группы источников финансирования инвестиций.

1. Собственные источники финансирования.
2. Приравненные к собственным.
3. Заёмные средства.

Заемные средства значительно влияют на ожидаемый “выхлоп” от проекта. Именно по этой причине, большинство финансовых проектов предполагают определенный процент доходности на начальных этапах. Норма доходности, помимо ожидаемой прибыли, должна покрывать потери вследствие естественного процесса инфляции, а также начисляемые по заемным средствам проценты. Обычно высокая норма доходности следует из высоких проектных рисков.

Риски увеличиваются пропорционально сроку проекта, влиянию внутренних и внешних факторов, которые могут значительно увеличить его стоимость.

Для того чтобы уменьшить влияния экономический неопределенности, снизить степень риска, суммы возникающих потерь, и принять обоснованное решение, применяется процедура синхронизации финансовых потоков затрат и результатов с применением ставок дисконтирования. Синхронизация – приведение данных по инвестициям и результатам инвестиций к одной базе во времени. Причина синхронизации финансовых потоков кроется во влиянии объективных факторов. Основным внешним фактором является инфляция [1].

Существуют следующие методы количественной оценки рисков:

1. “Спираль” рисков проекта.
2. Сценарный анализ
3. Анализ чувствительности.

В ходе анализа чувствительности происходит последовательное изменение переменных, которые по результатам качественного анализа оказывают влияние на проект. Ими могут быть показатели цены реализации продукции, объемов продаж, постоянных или переменных издержек, сроков реализации проекта, ставки дисконтирования и др.

Существуют разные варианты синхронизации капиталовложений и их доходности относительно времени. Рассмотрим последовательность приведения данных инвестиционного проекта к одной базе (к 0 или T) на временной оси. Обозначим: t_i – текущий момент времени, $t_i = 0,1,..,T-1,T$, где 0 – момент начала проекта, T – момент его завершения.

Для синхронизации инвестиций базой сравнения принимается 0-момент. В этом случае все данные моментов T, $T-1$, $T-2$ и т.д. пересчитываются по ценности денег 0-момента. Эта система пересчета данных получила называние "дисконтирование". Основным показателем в этих методах является параметр r – норма дисконта – постоянная на весь период проекта. Это есть минимальная удовлетворяющая инвестора доходность за один год на сумму вложенного капитала. В случае, когда период составляет месяц или квартал, норма дисконта рассчитывается как r/m, где m – число месяцев или кварталов года. Общепринятые периоды – месяц, квартал, полгода, год имеют соответственно значения $m = 12, 4, 2, 1$. Норма дисконта должна обеспечивать доходность, превышающую альтернативные безрисковые варианты вложений, например, % по банковским депозитам. С помощью нормы дисконта формируется коэффициент дисконтирования – d_t, относительно которого впоследствии осуществляется пересчет данных за весь период проекта [2].

Распределение капиталовложений и доходов во времени также влияет на полученные в итоге дисконтированные суммы, которые пересчитываются во временных промежутках с учетом Inf – инфляционной ставки. Пересчитанная ставка дисконтирования r_{inf} учитывает одновременно инфляцию и процент по заемным средствам, которые инвестор заинтересован отдать как можно быстрее.

Формула для ставки дисконтирования с учетом инфляции имеет вид:

$$r_{\text{inf},I} = r + \text{Inf} + (r \times \text{Inf}),$$

где r – исходная ставка дисконтирования, а Inf – норма инфляции (предполагаемая величина инфляции за период T). Параметры r, Inf являются внешними факторами риска, длина периода T – внутренним фактором риска, так как задается самим инвестором. Таким образом, сформированная ставка $r_{\text{inf},I}$ является количественной мерой учета рисков инвестирования. Далее она используется в процедурах синхронизации как основной корректирующий элемент потоков инвестиций и их доходности.

Существуют три метода для синхронизации капиталовложений и доходов, рассмотрим два основных: метод простых процентов и метод сложных процентов.

2. Метод простых процентов

Метод применяется только тогда, когда затраты и доход не распределены во времени, а вложены единовременно. Коэффициент дисконтирования образуется линейно, в связи с временными моментами.

Если базой сравнения выбран 0-момент, коэффициент дисконтирования определяется по формуле $d_{ti} = (1 + r_{\text{inf},I} \times |t|)^{-1}$. Дисконтированная сумма капитальных вложений рассчитывается как $\text{КВ}_d = \sum_{t=0}^{T} \text{КВ}_{ti} \times d_{ti}$, дисконтированная сумма ожидаемого дохода $D = D_0 \times d_{ti}$.

3. Метод сложных процентов

Метод применяется в случае, если капитальные вложения и доходы распределены во времени неравными долями.

Для случая, когда базой для сравнения является 0 – момент, коэффициент дисконтирования определяется по формуле $d_{ti} = (1 + r_{\text{inf},I})^{-t}$, дисконтированные суммы капиталовложений и доходов определяются аналогично методу простых процентов.

4. Интегральные показатели оценки эффективности инвестиций

В разработанном программном модуле для анализа финансовых потоков используются интегральные показатели эффективности инвестиций: дисконтированный срок окупаемости (DPP – Discounted Payback Period), индекс прибыльности (PI – Profitability Index), внутренняя норма рентабельности (IRR – Internal Rate of Return), дисконтированная прибыль (NPV – Net Present Value). В совокупности они образуют финансовую модель проекта. На основе этих показателей производится финансовый анализ инвестиционного проекта и его рисков.

Входными данными для расчета являются:

- суммы инвестиций и дохода, КВ, D;
- срок проекта, T;
- норма дисконта, r;
- норма инфляции, Inf;
- база расчета (0 или T)
- процент по заемным средствам, I;
- метод расчета (сложные или простые проценты)

Сущность экономического анализа строго зависит от предметной области инвестирования.

Экономическое значение интегральных показателей:

- NPV, дисконтированная прибыль (чистый приведенный доход) – есть разница между суммой дисконтированного дохода и суммой дисконтированных затрат. Рассчитывается по формуле [2]: $\text{NPV} = \sum_{t=0}^{T} D_{ti} \times (1 + r_{\text{inf},I})^{-t} - \sum_{t=0}^{T} \text{КВ}_{ti} \times (1 + r_{\text{inf},I})^{-t}$.
- IRR, это процентная ставка, при которой чистая приведённая стоимость (чистый дисконтированный доход – NPV) равна 0. NPV рассчитывается на основании потока платежей, дисконтированного к сегодняшнему дню. Рассчитывается по формуле [2]: $\sum_{t=0}^{T} D_{ti} \times (1 + \text{IRR})^{-t} = \sum_{t=0}^{T} \text{КВ}_{ti} \times (1 + \text{IRR})^{-t}$.
- DPP, дисконтированный срок окупаемости – период (измеряемый в месяцах, кварталах или годах) в момент окончания, которого, поступивший дисконтированный доход покрыл всю сумму дисконтированных затрат. Рассчитывается по формуле: $\text{DPP} = T_I + (\text{КВ}_d - D_d) \div D$.
- PI, индекс прибыльности – отношение дисконтированного дохода к сумме дисконтированных затрат по проекту. Рассчитывается по формуле: $PI = D_d \div KB_d$.

Наилучшим вариантом проекта считается тот, для которого все интегральные показатели одновременно отвечают следующим условиям [3]:

- дисконтированный срок окупаемости $\text{DPP} = \min$;
- индекс прибыльности $\text{PI} \geq 1$;
- внутренняя норма рентабельности $\text{IRR} > r$;
- дисконтированная прибыль $\text{NVP} > 0$ и стремится к max.

Таким образом, рассчитанные интегральные показатели, учитывая суммарное влияние внутренних и внешних рисков инвестиций, позволяют сравнивать между собой

варианты инвестиционных проектов и выбрать наиболее эффективный с позиции инвестора.

5. Программная реализация

На рис. 1 представлена упрощенная диаграмма классов разработанного приложения. Класс Main отвечает за всю вычислительную работу, в нем запрограммированы все необходимые формулы для расчета интегральных показателей. Метод Calculate() последовательно производит процедуры дисконтирования (с помощью методов класса DiscountRate), а затем вычисляет все интегральные показатели. Класс DiscountRate вынесен отдельно, так как включает в себя дисконтирование средств относительно разных временных баз и методов.

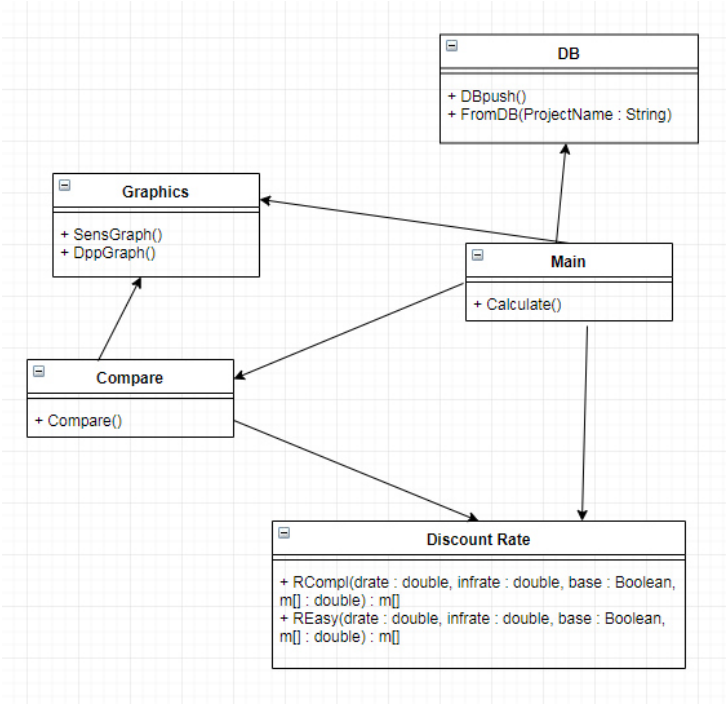


Рис. 1. Упрощенная диаграмма классов

Класс Graphics отвечает за графическую составляющую приложения: рисует графики для анализа чувствительности – метод SensGraph(), а также для нахождения точки безубыточности – метод DppGraph().

Класс DB отвечает за взаимодействие с базой данных. Он предоставляет точку доступа к БД без прямого использования языка SQL пользователем. Класс транслирует пользовательские запросы из UI в язык СУБД и отправляет их на сервер.

Приложение позволяет выбрать метод, базу расчета и ввести данные по заемным средствам, если они имеют место быть. После нажатия на кнопку “Рассчитать” (см. рис. 2), программа последовательно сделает следующие вещи: сначала вычислит дисконтированные суммы, затем рассчитает всевозможные интегральные показатели. После этого пользователь может провести графический анализ чувствительности интегральных показателей по выбранным параметрам или определить точку безубыточности инвестиционного проекта, перейдя на соответствующую вкладку, либо закончить работу с данными и сохранить нужные результаты в БД.

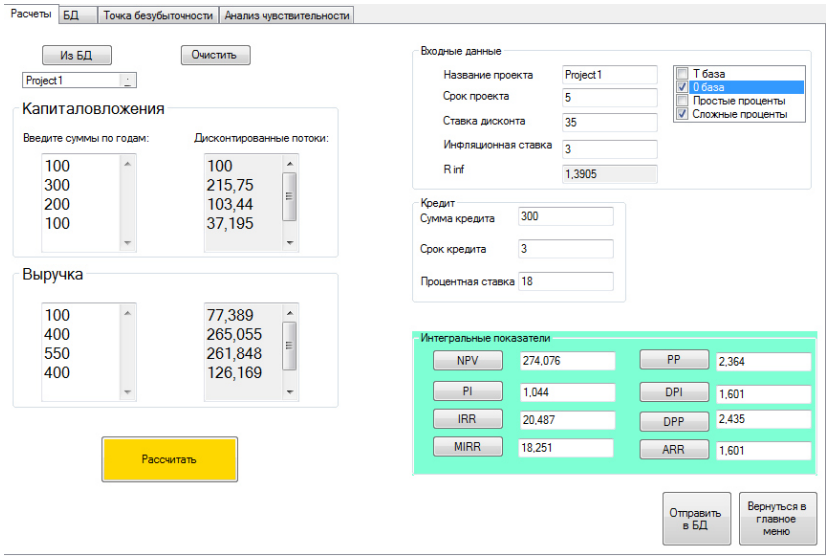


Рис. 2. Окно расчета интегральных показателей

Сравнение схем инвестирования происходит аналогично предыдущей процедуре, рис. 3. После введения входных данных, производится расчет всех показателей, теперь уже для двух проектов. На основе сравнения расчетных данных программа делает вывод о том, какой проект и с какой точки зрения лучше или выгоднее, и стоит ли его одобрить для практической реализации.

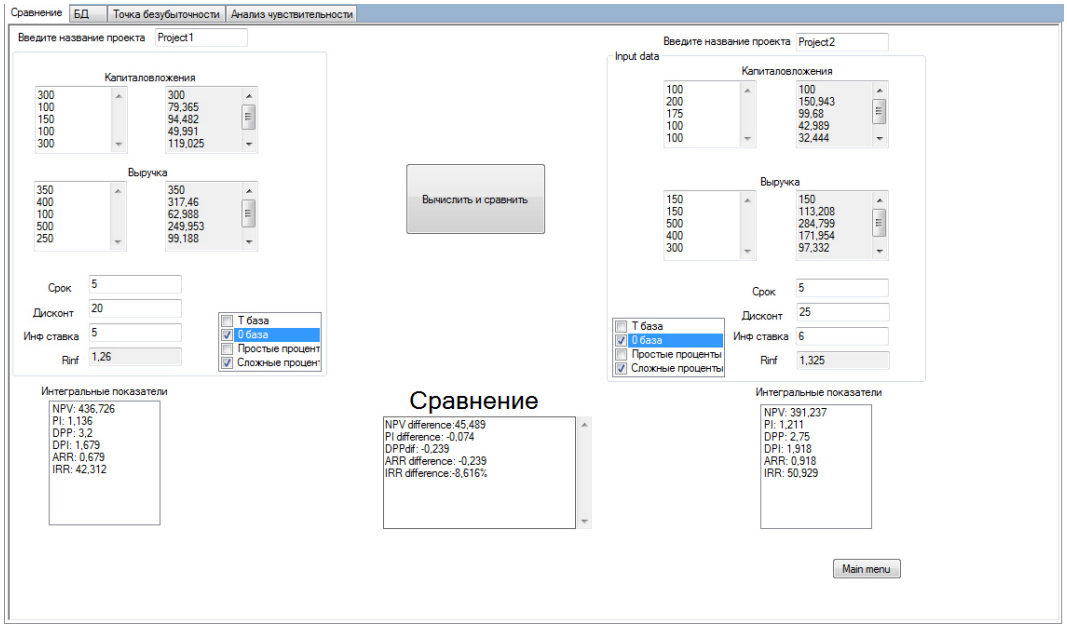


Рис. 3. Окно сравнения схем инвестирования

Анализ чувствительности проекта можно производить по разным показателям: объем прибыли, значение показателя инфляции, значение ставки дисконтирования и т.д (рис. 4). Данные для анализа берутся из расчитанных с окна для расчета интегральных показателей (рис. 1).

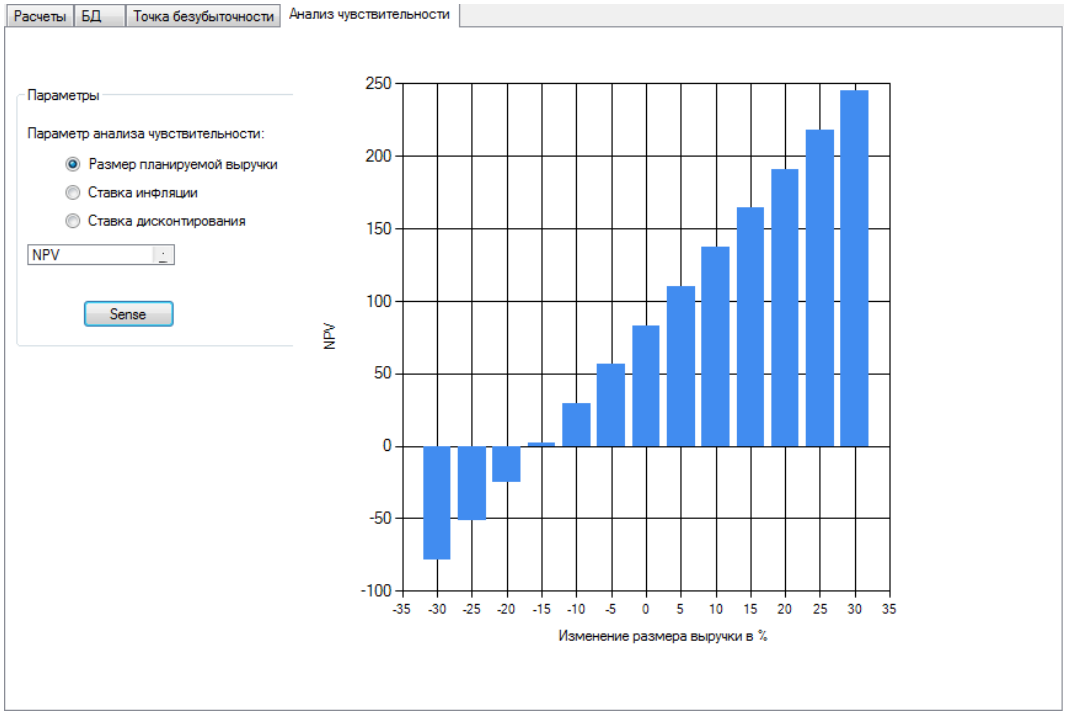


Рис. 4. Анализ чувствительности

Заключение

Разработано приложение, облегчающее расчетную работу при анализе инвестиционных проектов и являющееся удобным инструментом для расчета и последующего хранения расчетных данных. Приложение обладает следующим функционалом:

1. Позволяет пользователю синхронизировать денежные потоки, пересчитывая их по *T*- или 0-базе используя методы сложных или простых процентов.
2. Рассчитать интегральные показатели для инвестиционного проекта, численное значение которых позволит оценить его эффективность.
3. Рассчитать и сравнить показатели нескольких вариантов инвестиционного проекта для выбора наиболее выгодного с точки зрения инвестора.
4. Сохранить и использовать расчетные данные с помощью встроенной базы данных.
5. Имеется возможность формирования текстовых отчетов.
6. Имеется возможность графической интерпретация денежных потоков и нахождения точки безубыточности.
7. Проведения анализа чувствительности и его графической интерпретаций.

Как итог: приложение значительно облегчит расчетную работу инвестора, перед которым стоит задача выбор направления вложения средств с наибольшей для себя выгодой. Все неудобные расчеты запрограммированы и выполняются автоматизировано, от пользователя требуется лишь введение входных данных (рис. 1). База данных позволяет осуществлять важный для анализа экономики процесс накопления информации. Графические примитивы позволяют наглядно показать движение денежных потоков инвестиций и выручки, а также оценить устойчивость инвестиционного проекта на основе анализа его чувствительности по выбранным параметрам.

ЛИТЕРАТУРА

1. *Дасковский В., Киселев В.* Совершенствование оценки эффективности инвестиций // Экономист. – 2009. – № 1. – С. 42–57.

2. *Шапиро Л.Д.* Конспект лекций по курсу “Экономика производства”. – Томск, 2014. – 152 с.
3. *Баддин К.В., Передеряев И.И., Рукосуев А.В.* Антикризисное управление: макро- и микроуровень. – М.: Дашков и К, 2012. – 268 с.

ПРОЕКТИРОВАНИЕ И РАЗРАБОТКА ГИБРИДНОГО МОБИЛЬНОГО ПРИЛОЖЕНИЯ ДЛЯ НАУЧНОЙ ШКОЛЫ

Д.В. Копасов
Томский политехнический университет

Введение

На сегодняшний день информационные технологии используются практически во всех отраслях человеческой деятельности, а наблюдаемая в последние десятилетия тенденция к уменьшению размеров вычислительной техники и увеличению её доступности для широкого круга лиц привела к тому, что подавляющее большинство людей пользуется каким-нибудь мобильным гаджетом ежедневно (мобильные телефоны, смартфоны, планшеты, ноутбуки и прочее) в зависимости от личных нужд и предпочтений. Соответственно при необходимости иметь актуальную на текущий момент возможность взаимодействия с клиентами стоит обратиться к использованию мобильных приложений. С учётом того, что большое количество людей пользуются именно смартфонами [1], конечно стоит разрабатывать приложения именно для таких типов гаджетов.

За последние десять лет стремительно выросло число пользователей смартфонов, которые подразделяются на разные операционные системы: iOS, OS Android и другие [2]. На основании статистики распространённости основных мобильных операционных систем [3] более предпочтительным является разработка клиентского приложения под OS Android.

1. Описание предметной области

Данная международная научная школа будет проводиться в течение очень короткого срока (примерно в течение одной недели), а значит, участникам предстоит интенсивная работа. Всего в данном мероприятии выделяются три роли участников: студенты, докладчики и члены жюри.

Под студентами подразумеваются, как правило, учащиеся учебных заведений, которые привлекаются к получению и освоению новых знаний в определённой области образования.

В качестве докладчиков для международной школы приглашаются специалисты в сфере деятельности, связанной с основной идеей проводимых мероприятий, которые могут поделиться со студентами своим опытом, относящемся к их области образования и способствующем решению практических задач.

Поскольку задание будет сложным для выполнения в одиночку и в такой короткий срок (меньше одной недели), то студенты будут объединены в команды и выполнят задание совместно. После выполнения этих заданий их нужно будет оценить. Этим как раз и будут заниматься члены жюри, при этом каждый из них будет оценивать работу по девяти общим критериям и одному специальному, которое у каждого члена жюри будет своим, в зависимости от области специализации конкретного эксперта.

2. Описание возможностей пользователей мобильного приложения

В проводимой международной научной школе существует множество аспектов, касающихся различных ролей пользователей. Как уже упоминалось, выделяются три роли участников: студенты, докладчики и члены жюри. Однако, исходя из того, что согласно этим ролям участники выполняют некоторые уникальные действия, а также того, что существует ряд прочих действий, являющихся универсальными для всех пользователей приложения, в целом, для данного приложения можно выделить 6 категорий пользова-

телей. У каждой категории существуют определённые варианты использования приложения, из которых часть наиболее важных и отличительных указаны в соответствии с категориями пользователей приложения:

- пользователь приложения (просмотр расписания, информации о докладчиках и прочее);
- неавторизованный пользователь приложения (авторизация);
- авторизованный пользователь приложения (просмотр и скачивание учебного материала, просмотр и редактирование информации в личном кабинете);
- студент (выполнение задания и просмотр итоговой оценки и комментариев к работе);
- докладчик (добавление учебного материала);
- член жюри (проверка работ студентов).

Среди представленных категорий пользователей выделяется определённая иерархия наследования возможностей – «дочерний элемент» имеет доступ к функционалу «родителя». В данной же ситуации определяется следующая иерархия категорий пользователей приложения: пользователь является «родителем» для неавторизованного и авторизованного пользователя; у авторизованного пользователя есть 2 «дочерних категории» – студент и докладчик; у докладчика в свою очередь «дочерней» категорией является член жюри.

Для всех категорий пользователей приложения определён набор доступных действий, полный список которых можно более ёмко описать с помощью диаграммы вариантов использования в нотации UML, как представлено на рис. 1.

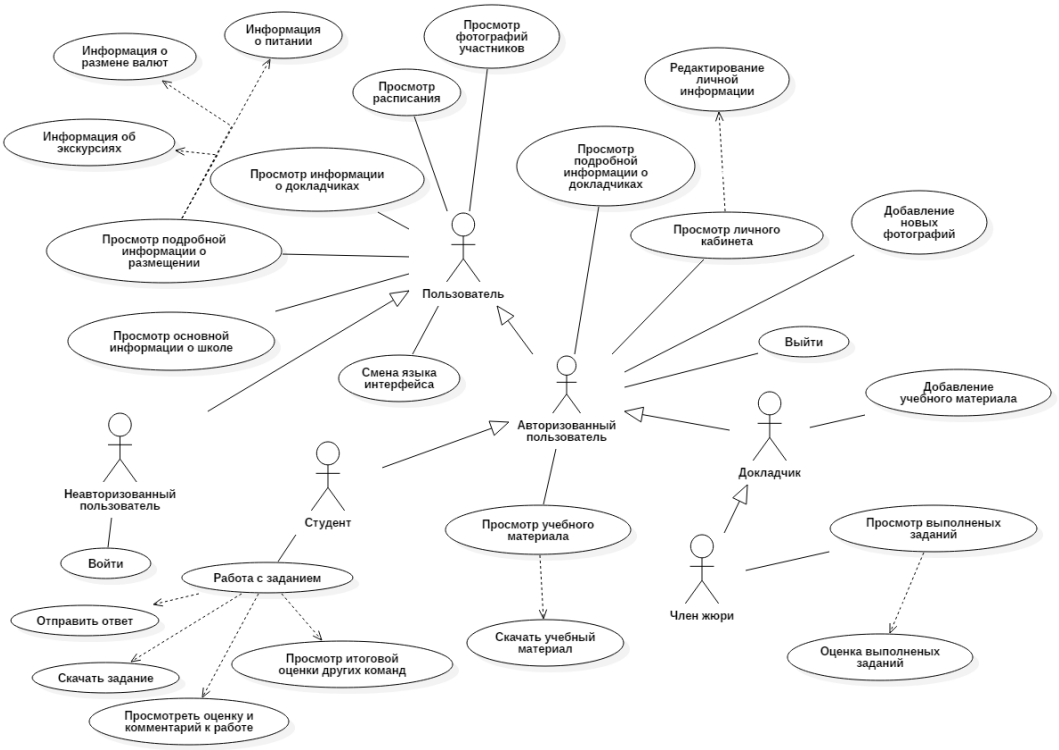


Рис. 1. Диаграмма вариантов использования мобильного приложения

3. Использование облачного хранилища Google Drive

Для хранения множества файлов, необходимых для проведения научной школы было принято решение использовать сервис облачного хранения файлов. Среди множества вариантов существующих решений (Dropbox, Google Drive, Яндекс Диск и прочие)

был выбран Google Drive из-за достаточного для проведения научной школы объёма предоставляемого пространства в 15 Гб, надёжности, узнаваемости, а также ввиду простой и доступной системы взаимодействия с помощью языка программирования Java, который используется для разработки всего приложения [4].

Для загрузки файлов в «облако» сначала необходимо произвести настройку машины, которая будет выступать в роли серверной части приложения. Всё последующее взаимодействие с сервисом облачного хранения файлов устроено следующим образом:

- по вышеупомянутой инструкции происходит загрузка любого файла в заданную директорию (загрузка файла проходит с мобильного приложения на серверное приложение, с которого файл загружается на Google Drive);
- по результатам загрузки файлу приписывается определённый индекс, который необходимо сохранить для дальнейшего использования (в данном случае используется база данных);
- имея идентификатор файла в хранилище (данные извлекаются из базы данным отдельным запросом), с помощью мобильного приложения можно напрямую открыть или скачать требуемый файл по ссылке вида «https://drive.google.com/open?id=12345» (12345 – предполагаемый идентификатор документа).

По представленному принципу проходит работа со всеми файлами в приложении, в частности: происходит загрузка и отображение фотографий профиля участников; загрузка и скачивание учебного материала; скачивание документа с подробным описанием задания; загрузка и проверка командных ответов на задание.

4. Структура взаимодействия мобильного приложения с веб-сервером

В разрабатываемом мобильном приложении осуществляется множество сетевых запросов к веб-серверу предназначенных для получения и отправки данных. Для обработки поступающих от клиента HTTP-запросов используются обработчики, соответствующие определённому модулю приложения. Таким образом, для функционирования приложения предусмотрены следующие обработчики:

- *UserHandler* – обработка запросов, связанных с пользовательскими данных;
- *PhotoHandler* – обработка запросов, связанных с фотографиями;
- *DocumentHandler* – обработка запросов, связанных с учебными материалами;
- *ScheduleHandler* – обработка запросов, связанных с расписанием;
- *SpeakerHandler* – обработка запросов, связанных с информацией о докладчиках;
- *TaskHandler* – обработка запросов, связанных с выполнением задания;
- *EstimationHandler* – обработка запросов, связанных с проверкой работ студентов.

Каждый из представленных обработчиков принимает на обработку свой запрос, например HTTP-запрос по URL http://server-name/task/{action} примет на обработку *TaskHandler*. При этом обработчики выполняют определённые действия, заданные в URL параметром action, с использованием передаваемых параметров в запросе. Структура обработки запросов сервером отображена на рис. 2.

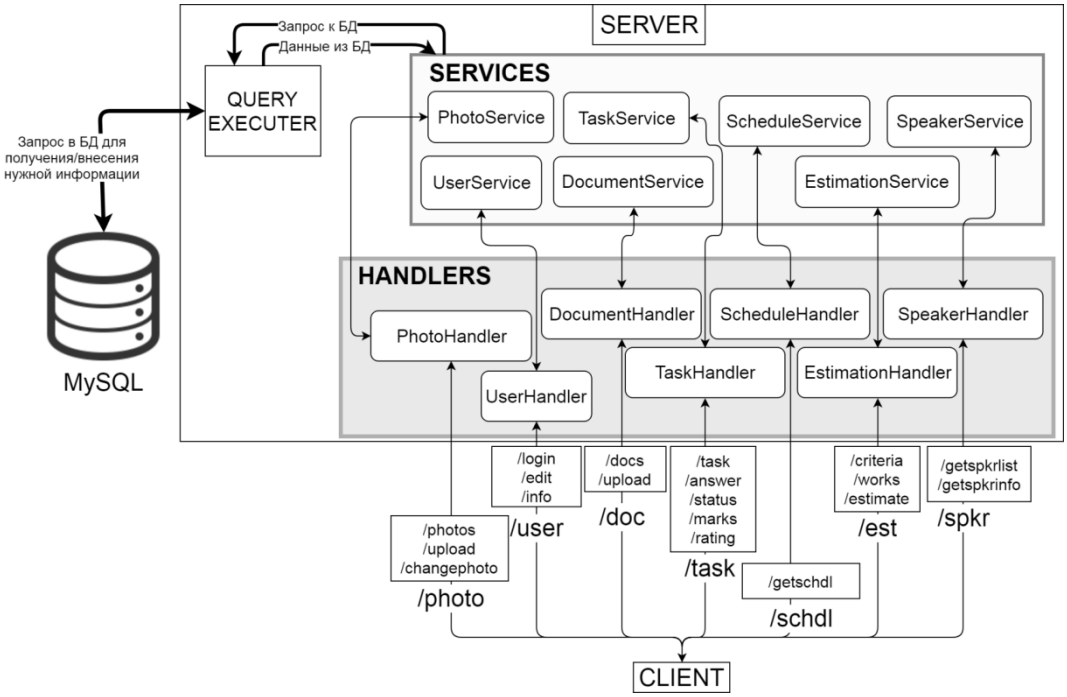


Рис. 2. Структура взаимодействия мобильного приложения с веб-сервером

Как видно, основной функционал приложения организован с помощью веб взаимодействия между клиентом и сервером, однако функционал мобильного приложения позволяет кешировать ранее полученные ответы от сервера, что обеспечивает дальнейший доступ к данным при отсутствии Интернет-соединения. Например, таким образом, работает расписание, что может быть весьма удобно тем, кто может испытывать затруднения с постоянным подключением к Интернету. Именно это позволяет назвать данное приложение гибридным.

5. Демонстрация основного функционала

В разработанном приложении разработан весь функционал, определённый в диаграмме вариантов использования (раздел 2), и ввиду своего значительного объёма лучше продемонстрировать некоторые основные возможности разработанного приложения, к которым относятся просмотр расписания, выполнение задания студентами и проверка выполненных работ членами жюри.

В проводимой международной школе проводятся различные мероприятия, направленные на освоение навыков и решения задач в области анализа больших данных, а также культурные и прочие мероприятия. Для всех это расписание является единым ввиду короткого срока проведения научной школы и относительно небольшого числа участников. С помощью мобильного приложения участники смогут ознакомиться с последовательностью проведения мероприятий, временем начала и продолжительностью, типом проводимых событий, списком докладчиков и местом проведения. В случае необходимости в расписании можно выбрать докладчика и узнать о нём более подробную информацию, а также, выбрав местоположение мероприятия, увидеть его на карте. Демонстрация данного функционала приведена на рис. 3.

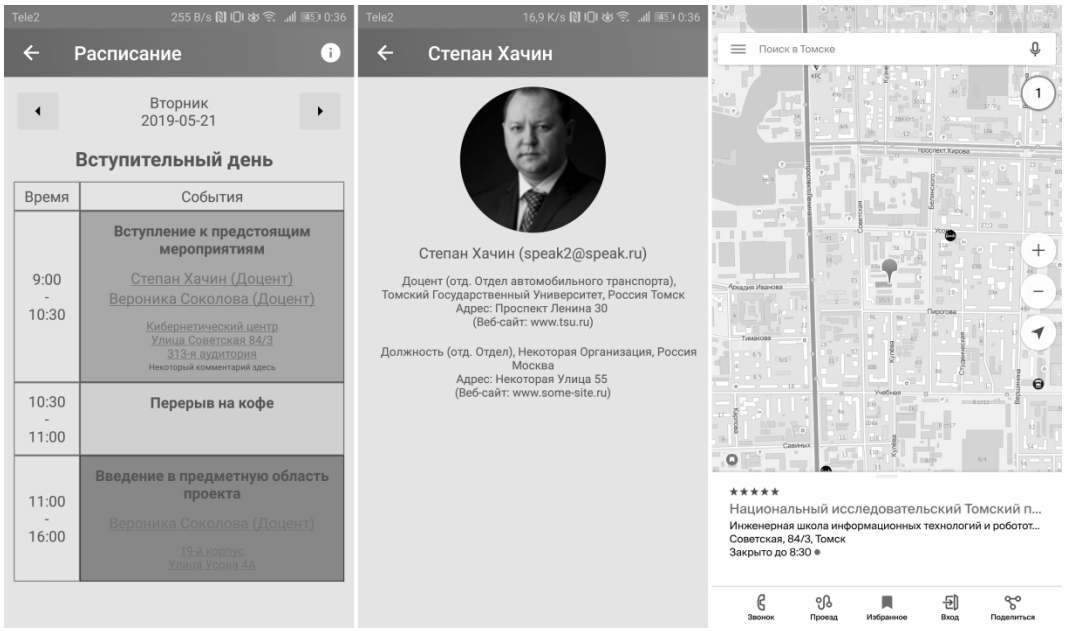


Рис. 3. Демонстрация возможностей просмотра расписания

В рамках проводимой научной школы студентам необходимо будет выполнить командное задание, которое направлено на решение заданной проблемы в области обработки больших данных. С помощью приложения студенты смогут ознакомиться с подробным описанием задания, а после выполнения один из членов команды отправит ответ на задание. После того, как члены жюри проверят работу и выставят оценку, участники команды смогут ознакомиться с выставленными оценками и общим рейтингом всех команд, как показано на рис. 4.

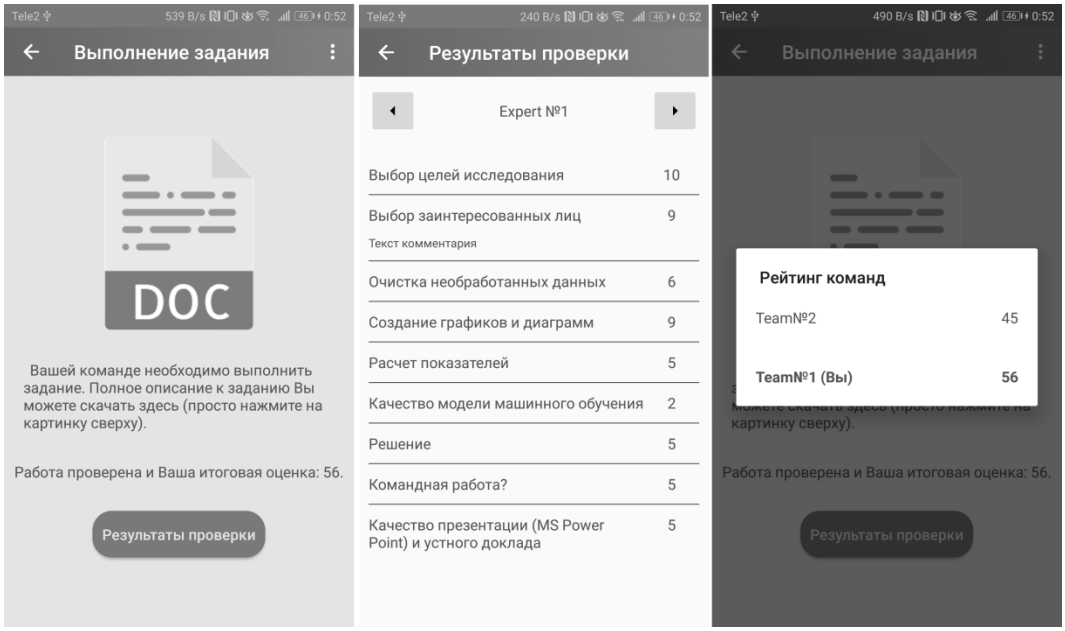


Рис. 4. Демонстрация возможностей по просмотру результатов выполнения задания

Для оценки результатов выполнения работы члены жюри, воспользовавшись мобильным приложением, могут скачать ответ на задание и ознакомиться с ним. После

окончательного изучения предоставленного ответа от студентов члены жюри, выбрав команду для проверки, выставляют оценки по требуемым критериям, а также могут добавить комментарии к оценке. По завершении процесса оценки всеми членами жюри, студенты смогут увидеть результаты оценки по своей работе. Демонстрация процесса проверки работы представлена на рис. 5.

Рис. 5. Демонстрация возможностей проверки работы студентов

Заключение

В результате проделанной работы спроектировано и реализовано гибридное мобильное приложение. Были рассмотрены основные возможности пользователей приложения и описаны с помощью диаграмма вариантов использования в нотации UML, а также поверхностно рассмотрены обрабатываемые веб-сервером HTTP-запрос и продемонстрирован основной функционал мобильного приложения.

ЛИТЕРАТУРА

1. Digital in 2019 [Электронный ресурс]. – URL: https://wearesocial.com/uk/blog/2019/01/digital-in-2019-global-internet-use-accelerates

2. Android на рынке мобильных ОС [Электронный ресурс]. – URL: https://habr.com/ru/post/101737/

3. Как iOS и Android разделили мобильный рынок [Электронный ресурс]. – URL: https://appcraft.pro/blog/kak-ios-i-android-razdelili-mobilnyi-rynok/

4. Манипулирование файлами и папками в Google Drive с использованием Java [Электронный ресурс]. – URL: https://o7planning.org/ru/11889/manipulating-files-and-folders-on-google-drive-using-java#a20602553

РАЗРАБОТКА ИНФОРМАЦИОННОЙ СИСТЕМЫ ДЛЯ РАСЧЕТА КАЛЬКУЛЯЦИИ СЕБЕСТОИМОСТИ ПРОДУКЦИИ

Д.Г. Максутова, А.С. Морозова
Томский государственный университет

Введение

Чем экономнее организация и рациональнее использует труд, материалы и финансы для производства продукции, оказания услуг, тем выше эффективность производственного процесса, и тем больше прибыль. Большой удельный вес в формировании себестоимости продукции составляет система оплаты труда, а вместе с ней система

стимулирования труда работников, так как основной целью системы стимулирования труда является реализация интересов предприятия, к которым относятся увеличение объемов выручки, интенсификация сбыта, повышение производительности труда и снижение себестоимости продукции. Чтобы определить, насколько цель достигнута предприятием, необходимо провести анализ эффективности стимулирования трудовых ресурсов, который показывает, в какой мере затраты на материальное и моральное стимулирование работников оправданы, произошел ли прирост продукции, снизилась ли себестоимость. Таким образом, одной из основных задач предприятия – выявление и реализация резервов улучшения использования трудовых ресурсов и повышения производительности труда работников организации [1]. В работе [2] рассмотрено исследование математических моделей стимулирования сбыта продукции, которые позволяют топ-менеджменту формировать более эффективную стратегию формирования цены продукции, а значит и ее себестоимости.

Расчет затрат, расходов на производство необходим компании для:

- оценки выполнение плана по производству и его динамика;
- определения рентабельности производства и отдельных видов продукции;
- введения внутреннего учета затрат;
- выявления резервов для снижения себестоимости продукции;
- определения цен на продукцию;
- расчета экономической целесообразности внедрения новых технологий и организационно-технических мероприятий.

Для формирования калькуляций себестоимости продукции создано достаточно большое количество прикладных решений, каждый из которых - обладает рядом достоинств и недостатков. Одним из главных недостатков является сложность адаптации системы к работе конкретного предприятия.

Разрабатываемая информационная система предназначена для учета затрат и формирования калькуляции с учетом всех особенностей хлебопекарного дела. Использование подходящей для конкретных условий системы позволяет облегчить процесс калькулирования, что дает возможность повысить эффективность работы, а также сократить время, затраченное сотрудником на решение определенных поставленных задач.

На рис. 1 приведена схема бизнес-процесса хлебопекарного производства.

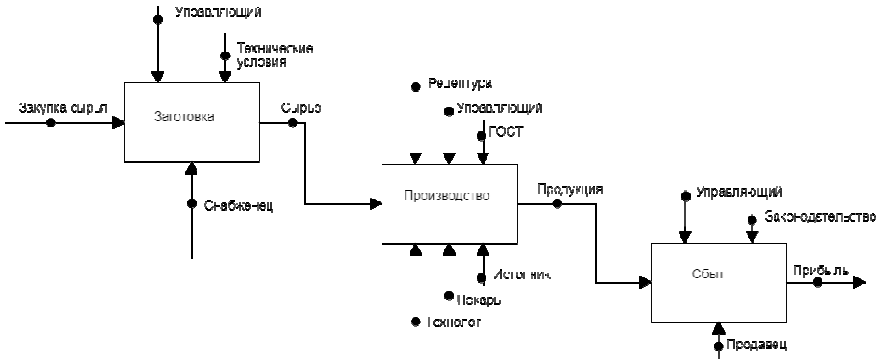


Рис. 1. SADT-модель деятельности хлебопекарного производства

Для расчета калькуляции себестоимости продукции на предприятии использовался метод прямого счета, без какой либо оптимизации. В связи с этим, был определен новый, наиболее правильный и соответствующий внутренним правила метод расчета калькуляции - расчетно-аналитический метод. Выбор данного метода связан со следующими факторами: расширение номенклатуры производства; увеличение объемов производства; в данном методе распределения затрат наиболее подходящие признаки расчета себестоимости; минимальные затраты на внедрение; на основании изучения струк-

туры и принятых нормативов предприятия расчетно-аналитический метод является самым актуальным.

1. Анализ IT-проектов рассматриваемой области. Обоснование выбора технологии проектирования хлебопекарного производства

ИТ-инфраструктура хлебопекарного предприятия на данный момент только начинает формироваться: используется ОС Windows, локальная сеть, веб-сайт – отсутствуют. Бухгалтерский, налоговый учет, учет кадров ведется с помощью «1С: Предприятия». Остальные элементы бухгалтерского учета выполняются вручную, поскольку нет разработанной конфигурации для их ведения. На предприятии возможно использование готовых информационных систем, представленных на рынке, но ни одна из них не соответствует полностью следующим особенностям пекарни: расчет калькуляции себестоимости расчетно-аналитическим методом, возможность редактировать метод расчета при последующем использовании, возможность заполнять и редактировать статьи калькуляции, возможность учесть при заполнении статей калькуляции все спецификации представляемой продукции.

Таблица 1

Анализ готовых решений

№	Наименование	Критерии оценки			
		Расчетно-аналитический метод	Редакция метода расчета в процессе работы	Заполнение и редакция статей калькуляции	Учет всех спецификаций
1	Базовая версия 1С: Предприятие.	+	+	+	–
2	Финансы в Excel. Калькуляция себестоимости.	–	+	+/–	–
3	OptiFood.	+	–	–	–

Каждое из вышеперечисленных решений было протестировано на предмет соответствия всем особенностям и требованиям. В результате, принято решение разработать конфигурацию «Калькуляция себестоимости» для программы «1С: Предприятие 8», удовлетворяющую всем требованиям.

2. Требования к разрабатываемой системе

Разрабатываемая информационная система должна удовлетворять следующим функциональным и нефункциональным требованиям, представленным в виде диаграмм вариантов использования [3].

Функциональные требования

Разрабатываемая система должна обеспечивать следующие функциональные возможности для Пользователя:

1. CRUD рецептуры продукции.
2. CRUD статей расходов.
3. Редактирование временного промежутка расчетов.
4. Формирование расходов по виду продукции.
5. Расчет суммы косвенных затрат на вид продукции.
6. Калькуляция себестоимости.

Функциональные возможности для Системного Администратора:

1. Доступ к системе в режиме «1С:Конфигуратор».
2. Возможность редактировать систему в режиме «1С: Конфигуратор».
3. Подготовка и сохранение резервных копий системы.

У системного Администратора должен быть доступ к возможностям Пользователя.

Нефункциональные требования

Использовать «1С: Предприятие 8».

На основе требований построены диаграммы использования.

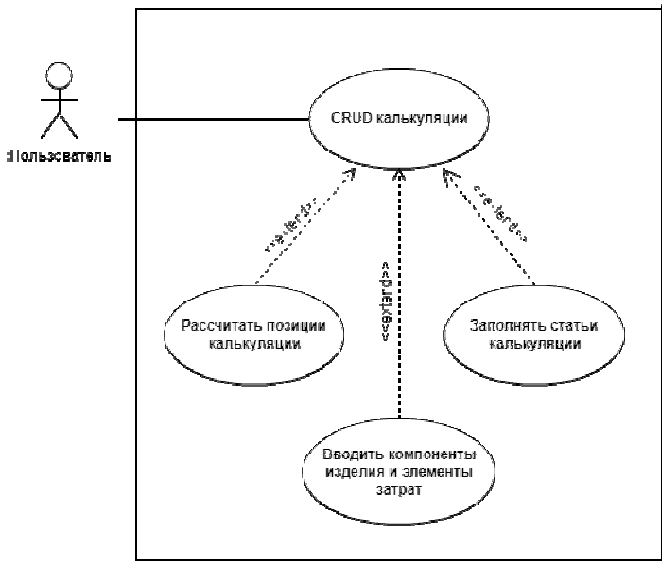


Рис 2. Диаграмма вариантов использования для роли «Пользователь»

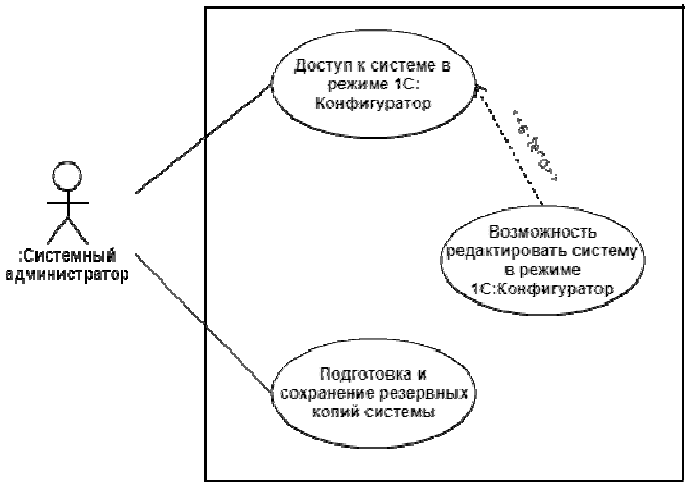


Рис 3. Диаграмма вариантов использования для роли «Системный администратор»

В ходе анализа требований разработана модель калькуляции себестоимости

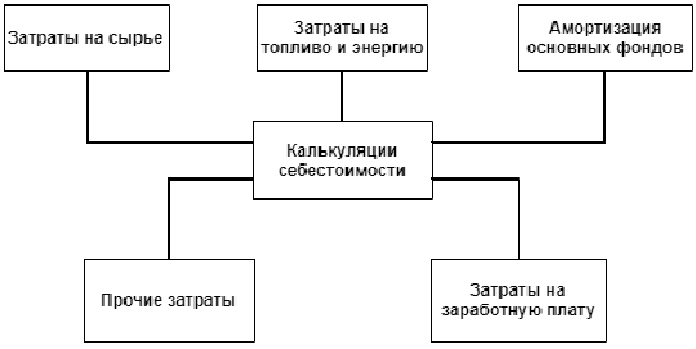


Рис 4. Модель калькуляции себестоимости

3. Реализация информационной системы

Для прикладного решения на базе «1С Предприятия 8» созданы подсистемы: документы, продукт и сотрудники. Подсистемы позволят разделить документы, справочники и отчеты по типу. Например, затраты будут располагаться в подсистеме продукт, данные по сотрудникам – в подсистеме сотрудники, созданы необходимые для решения задачи справочники и документы. В конфигураторе созданные подсистемы, справочники и документы выглядят следующим образом.

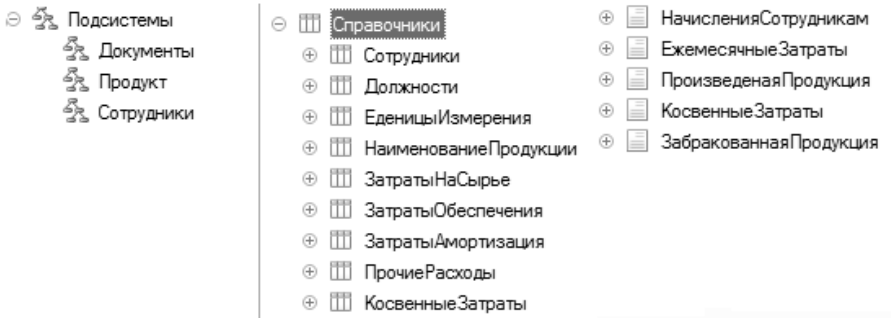


Рис 5. Созданные подсистемы, справочники, документы

Для каждого элемента созданы свои настройки.

На рис. 6 приведен вид элементов конфигурации.

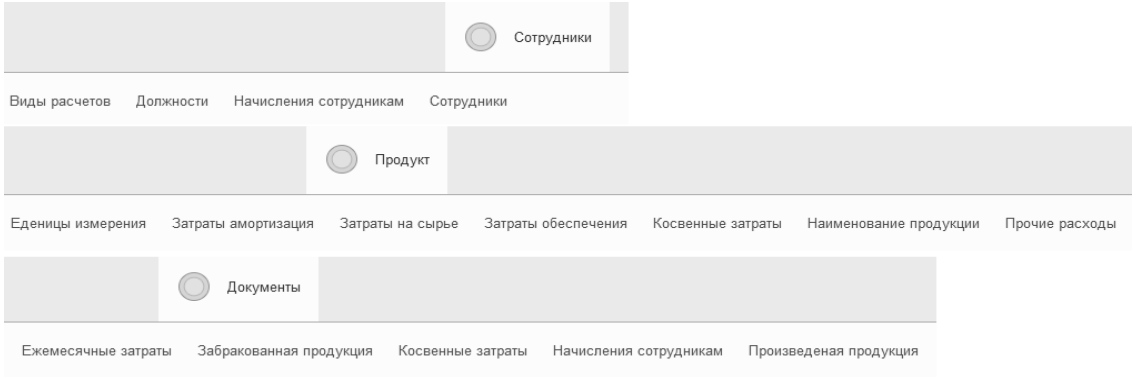


Рис. 6. Расположение элементов в конфигурации

Подсистема сотрудники

Виды расчетов. Справочник. Заполняется вид расчета заработной платы сотрудникам: оклад, премия штраф и т.д.

Должности. Справочник. Заполнены должности, согласно исходным данным по наличию должностей в пекарне «Хлеб на дровах».

Начисления сотрудникам. Начисления заработной платы сотрудникам отображаются в подсистеме сотрудники и в подсистеме документы, так как имеется связь с подсистемой сотрудники и в тоже время это документ, который создается бухгалтером.

Сотрудники. Справочник. Вносятся сотрудники, их ФИО.

Подсистема продукт

В данной подсистеме расположено все, что связано с продукцией и производством.

Единицы измерения. Единицы измерения продукции и затрат: кг, м. куб, шт. Такие единицы измерения как кг и шт. были созданы как предопределённые элементы справочника.

Следующие справочники и документы созданы с целью заполнения данных по статьям затрат: амортизация, сырье, обеспечение (обеспечение производства, т.е. затраты на топливо и электроэнергию), прочие расходы и косвенные затраты.

Как видно из рис. 6, косвенные затраты являются документом, так как отображаются и в подсистеме документов.

Подсистема документы

В данной подсистеме располагаются все документы, необходимые к заполнению и будет выводиться отчет о себестоимости производства.

Документ «Ежемесячные затраты». В данном документе заполняются в табличной части затраты на обеспечение (топливо и электроэнергия), амортизацию и прочие расходы.

Следующие документы по статьям затрат заполняются отдельно: «Забракованная продукция», «Произведенная продукция», «Косвенные затраты» и «Начисления сотрудникам». Для удобства заполнения начислений сотрудникам создана связь документов произведённой продукции и возможность создания на его основе документ начисления.

В документе «Произведённая продукция» заполняются следующие графы: сотрудник, продукция, количество. Заполняется не только продукция и количество, а еще и сотрудник, с целью дальнейшей доработки программы для отслеживания эффективности работы того или иного работника компании, так как планируется расширение производства, соответственно штата.

Документ «Забракованная продукция» позволяет проводить документы по браку на вид продукции (хлеба).

Отчет

Себестоимость производства был разработан с использованием системы Компоновки данных. В нем возможно:

1. Формировать отчет по виду продукции или полностью по всем видам продукции
2. Возможность выбора даты в отчете:
3. Себестоимость формируется с учетом распределения по статьям затрат.
4. Полностью расписана по рецептуре статья затрат на сырье.

Себестоимость производства

Продукция	Объем производства	На единицу продукции	Итого
Тип затрат			
Сырье/Статья			
Бездрожжевой ржаной подовый хлеб без добавления специй	**3 013**	**19,00**	**57 234,369**
Амортизация		0,01	25
Брак		0,03	100
Косвенные затраты		0,41	1 250
Прочие затраты		1,65	4 972,5
Статья ЗП		8,30	25 000
Сырье		8,41	25 348,369
Вода		0,10	295,274
Мука пшеничная 1 сорт		4,08	12 293,04
Мука ржаная обдирная		3,13	9 424,664
Соль		1,11	3 335,391
Топливо и электроэнергия		0,18	538,5

Рис. 7. Отчет себестоимости

Заключение

Таким образом, разработанная информационная система удовлетворяет всем требованиям для расчета калькуляции себестоимости хлебопекарной продукции.

ЛИТЕРАТУРА

1. *Морозова А.С., Шкуркин А.С.* Анализ системы мотивации и стимулирования труда в ООО «Анжерское молоко» // УЭкС. 2018. № 1 (107). С. 19. URL: http://uecs.ru/uecs-107-1072018/item/4751-2018-01-22-06-37-31 (дата обращения: 11.04.2018).

2. *Морозова А.С.* Исследование математических моделей стимулирования сбыта продукции: дис. канд. физ.-мат. наук: 05.13.18; Филиал Кемеровского гос. ун-т. в г. А.-Судженске – Анжеро-Судженск. 2007. 115 с.

3. *Максутова Д.Г.* Разработка и внедрение информационной системы по учету затрат и формированию калькуляции на предприятии // Информационные технологии. МНСК-2018: материалы 56-й Междунар. науч. студ. конф., 22-27 апр. 2018 г. – Новосибирск: ИПЦ НГУ, 2018. – С. 196.

VI. COMPUTER SCIENCE

THE USE OF STOETZEL METHOD IN PRICING OF NEW BIO-ENERGY DRINK

A.A. Andrievskaya, Z.N. Zenkova
Tomsk State University

Introduction

One of the most complicated decisions of an enterprise is to find how to price its products [1]. If the price is too low then you may not cover your costs or receive target profit, if it is too high then potential customers never turn into paying. Pricing strategies help a manager to answer such questions as [4, 5]:

– In what way the product should be priced?
– How much will sales fall if the price is increased?
– Are there price thresholds?
– Is it necessary for an enterprise to use price discrimination to increase profit?

The V. Wensterdorp approach [3] can be related to a similar method by Gabor and Granger and to a so-called "Psychological Price" developed in the early 1950s by Stoetzel. These models were probably inspired by the ideas of measuring the subjective probability of a "would-be bought" product, developed in games and utility theories. Price sensitivity techniques with some variations consist of data eliciting from the customers asking what is "Cheap" or "Expensive" price of a particular product.

1. Stoetzel Pricing Method

The method was proposed by Jean Stoetzel, French psychologist, and sociologist, in the middle of XX century [2]. The method is very simple, cheap and consists of the following stages. After the representatives of the target audience have used the product for some time period they were interviewed independently with two questions:

1) What price P_1 does this product start to seem too cheap that you start to suspect low quality and would not consider buying it?

2) What price P_2 does this product become too expensive that you would not consider buying it?

Therefore, we get an independent two-size sample $\left(P_{i1}, P_{i2}\right)$, $i = \overline{1,N}$, where N is a number of interviewed consumers. The sample is a result of observation on two prices P_j, $j = 1,2$ which are considered to be random variables with cumulative distribution function (cdf) $F^{(j)}(p) = P\left(P_j < p\right)$. Then the empirical cumulative distribution function (edf) is calculated for $j = 1,2$ by the formula $F_N^{(j)}(p) = \frac{1}{N}\sum_{i=1}^{N} \mathrm{I}_{(0,p]}\left(P_{ij}\right)$, where $\mathrm{I}_{(0,p]}(y) = \begin{cases} 0, & \text{if} \quad y \notin (0, p], \\ 1, & \text{if} \quad y \in (0, p] \end{cases}$ is indicator function. Edf is a step function and here it is smoothed. Required prices $P_{\min}$ and $P_{\max}$ are taken from the equation

$$d(p) = F_N^{(1)}(p) - F_N^{(2)}(p) = 0.5 .$$

An example of the implementation of the Stoetzel method is showed in fig. 1.

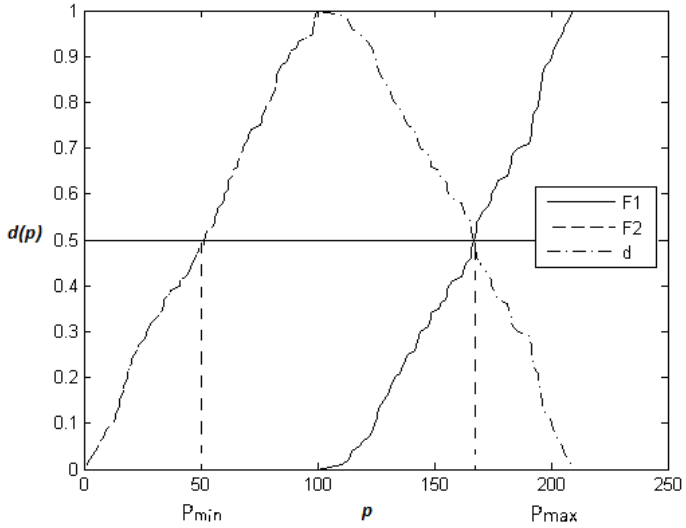


Fig. 1. An example of the Stoetzel method implementation

Note, that setting a price lower then $P_{\min}$ and higher then $P_{\max}$ leads to loss of the profit, because if the price is too low the majority of potential consumers refuses to buy it suspecting very bad quality and if the price is high the refusal accurse due to insolvency representatives of the target audience. If the enterprises targeted on wide market coverage and gaining more market share it is better to set a price about $P_{\min}$, but if the aim is gaining the status of goods the price should be closer to $P_{\max}$.

2. Quality Analysis of Price Estimations

The quality of the Stoetzel method was studied by means of simulations. The mean values and the standard deviation of prices, their distribution depending on the sample size N are analyzed. The number of simulation for every cases was $M = 100\,000$.

Firstly, we considered uniform distributed random variables $P_1 \in U_{[0;100]}(p)$, $P_2 \in U_{[110;210]}(p)$, and then the real prices can be found as

$$d(p) = F_1(p) - F_2(p) = R_{[0;100]}(p) - R_{[110;210]}(p) = \begin{cases} \dfrac{p}{100}, & p \in [0,100], \\ 1, & p \in [100,110], \\ \dfrac{p-110}{100}, & p \in [110,210]. \end{cases}$$

1) If $p < 100$ then $F_1(p) = \dfrac{p}{100}$, $F_2(p) = 0$. Consequently, $d(p) = \dfrac{p}{100}$.

2) For $p \in [100,110]$ $F_1(p) = 1$, $F_2(p) = 0$, therefore, $d(p) = 0$.

3) If $p \in [100,210]$ then $F_1(p) = 1$, $F_2(p) = \dfrac{p-110}{100}$ and $d(p) = \dfrac{100 - p + 110}{100} = \dfrac{210 - p}{100}$.

4) For $p > 210$ and $d(p) = 0,5$ we must consider two cases:

4.1) $\dfrac{p}{100} = \dfrac{1}{2}$, therefore, $P_1 = 50 < 210$,

4.2) $\frac{210-p}{100}=\frac{1}{2}$, then $P_1=160<210$.

Thus, the real values of the prices are $P_1=50$ and $P_1=160$.

Secondly, we simulate $N=\overline{20,300}$ uniform distributed prices and found the estimations of its mean values $\hat{E}P_i$ and mean square error $M\hat{S}E\left(P_i\right)$. The results are shown in fig. 2–4.

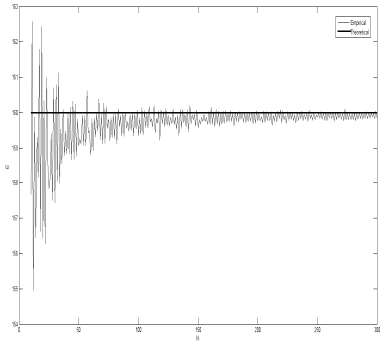


Fig. 2. The dependence of $\hat{E}P_1$ from N for $F_1(p)=U_{[0;100]}(p)$ and $F_2(p)=U_{[110;210]}(p)$

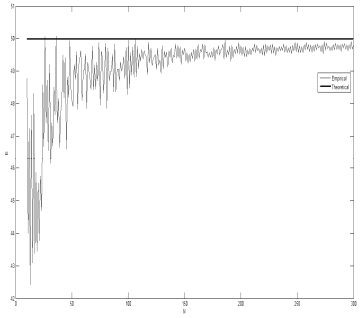


Fig. 3. The dependence of $\hat{E}P_2$ from N for $F_1(p)=U_{[0;100]}(p)$ and $F_2(p)=U_{[110;210]}(p)$

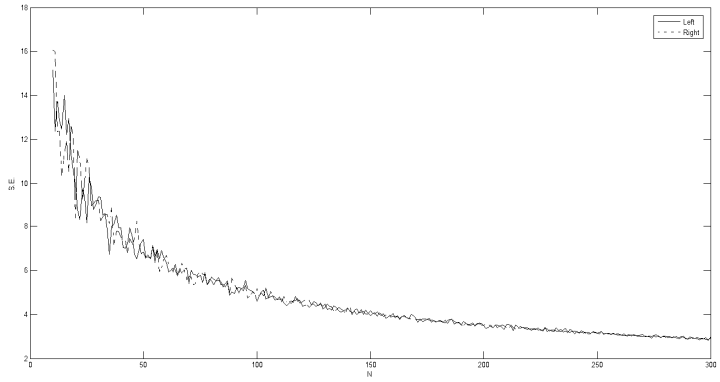


Fig. 4 The dependence of $M\hat{S}E\left(P_i\right)$ from N

As a result, we can conclude that the method is asymptotically unbias and its accuracy is increased with N.

3. Example

The Stoetzel method is applied to the real data of consumers' price preferences towards a non-alcoholic energy drink [6,7], which is launched by Tomsk production enterprise (see tab. 3).

Table 1

The results of the survey of energy drinkers using the Stoetzel method

№	P_1	P_2	№	P_1	P_2
1	18	52	14	48	60
2	25	45	15	58	90
3	48	69	16	48	75
4	28	41	17	43	63
5	23	39	18	28	90
6	29	42	19	40	100
7	28	89	20	58	85
8	18	66	21	48	98
9	28	59	22	38	74
10	32	67	23	48	88
11	16	40	24	48	65
12	28	48	25	38	55
13	28	48	26	46	112

As a result, we found that $P_1 = 45$ rub/item, $P_2 = 70$ rub/item. Enterprise managers are advised to set a price within the specified range. The results are shown in fig. 5.

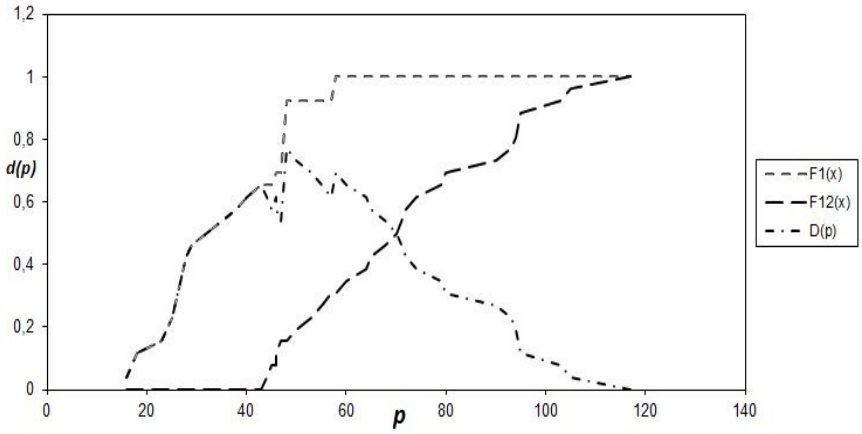


Fig. 5. The result of using method

Conclusion

In this paper, the quality of the Stoetzel method was studied by simulations. It was found that the method is asymptotically unbias and became more accurate with sample size increasing. The method was applied to a real data set about prices preferences of target consumers of a non-alcoholic energy drink launched by Tomsk production enterprise. The range of prices was calculated. The company managers are given practical recommendations.

REFERENCES

1. *Kotler P.* Marketing Management. – 11 Ed., Harlow: Prentice Hall International Editions, 2003.

2. *Stoetzel J.* Psychological/Sociological Aspects of Price // Bernard Taylor and Gordon Wills (Eds.).Pricing Research, Princeton, NJ: Brandon Systems Press. – 1970. – P. 70–74.

3. *Westendorp P.H.* NSS-Price Sensitivity Meter (PSM) – A New Approach to Consumer Perception of Prices // Venice Congress Main Sessions, European Marketing Research Society (ESOMAR), Amsterdam. – 1976. – P. 139–167.

4. *Nagle T.T., Holden R.K.* The Strategy and Tactics of Pricing // A Guide of Profitable Decision Making, 2nd edition, Prentice-Hall, Englewood Cliffs, NJ. – 1995.
5. *Comley P.* What do we know about Pricing research // WARC, Admap. – January 1997.
6. *Zenkova Zh., Krakovetckaia I.* Marketing of New Bio-Energy Drinks // Atiner conference paper series No: SME2013-0739. –2013.
7. *Zenkova Zh, Krakovetckaia I.* Marketing of New Bio-Energy Drinks // Theoretical and Empirical Reflections in Marketing. Athens Institute for Education & Research, ATINER. Greece. – 2014. – P. 179–192.

ONLINE PORTFOLIO SERVICE FOR MUSIC CONTENT CREATORS

V. Ivanov, A. Shkurkin
Tomsk State University

Introduction

With the advent of the Internet and the spread of online music platforms, free access to a huge amount of musical materials, instruments and opportunities has appeared. Music industry revenue grows nowadays [1]. It has led to the growth of people involved in music art.

There are still no acceptable ways for musicians to find each other. Such ways today are social networks and web applications primarily with the ability to download audio or video. But they are not designed specifically so that musicians can conveniently search for other users or share their skills and work.

So, the goal is to create a web-site that provides the functionality of creating an online portfolio for people involved in the creation of musical content.

This goal defines next objectives:

1. The definition of the main functionality.
2. Selection and study of development tools.
3. Creating a design layout.
4. Web application development.

1. Analysis of existing solutions

Consider some popular web applications.

«Soundcloud.com» [2]

It's an online platform for digital distribution of sound information with social networking features. Soundcloud provides users with uploading music or podcasts and some social network features like commenting and sharing. On Fig. 1 the main page of «Soundcloud.com» is represented.

«Bandcamp.com» [3]

This application also provides artists and labels with uploading music but it's more digital distribution platform than online-portfolio. Users can control how they sell their releases, setting their own prices, offering fans the option to pay more (which they do 40% of the time) and selling merchandise. The advantage of this application is advanced search and opportunity to sell music. On Fig. 2 the main page of «Bandcamp.com» is represented.

After analyzing the above and similar web-applications, it comes to conclusion:

– Applications don't provide instruments for mentioning users that participated in project.
– There is no detailed search with filters.

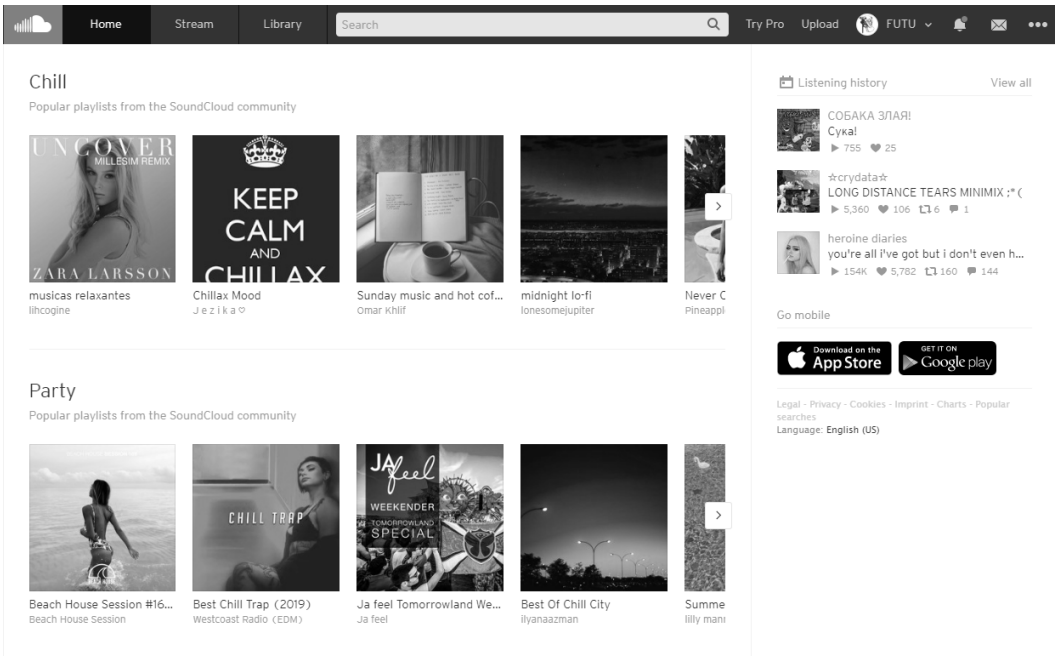


Fig. 1. The main page of «Soundcloud.com»

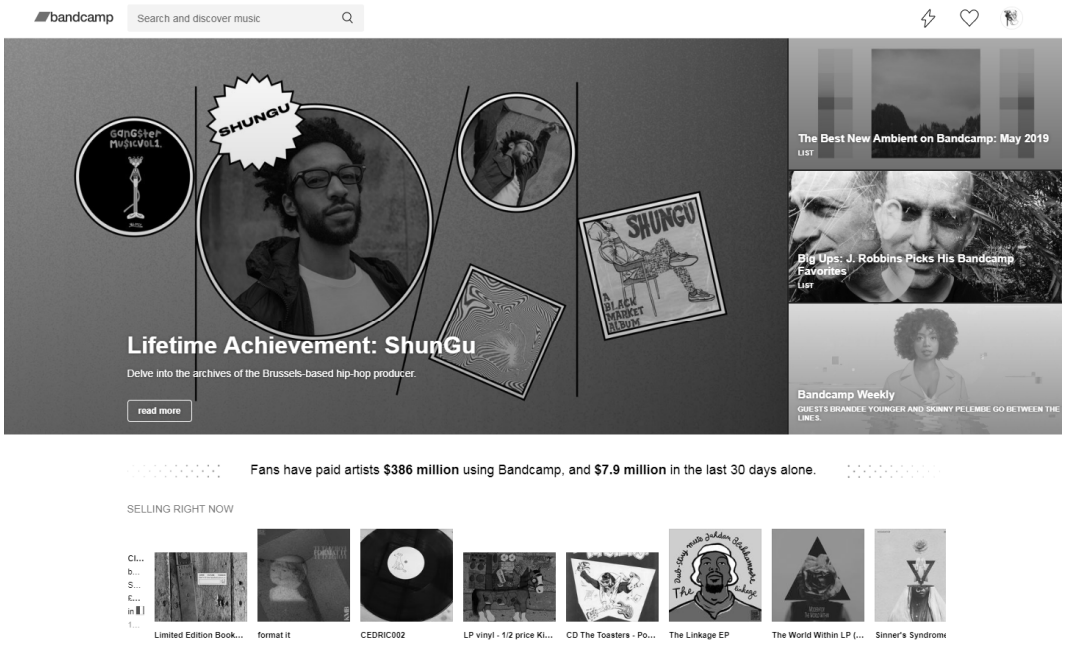


Fig. 2. The main page of «Bandcamp.com»

2. Domain modelling

The first step of web application modelling is domain modelling which shows the main details of system. Based on the fact that Javascript is not an Object Oriented Programming language It's hard to show all the relations between entities and business logic so on the Fig.3 you can see how Mongoose classes look like.

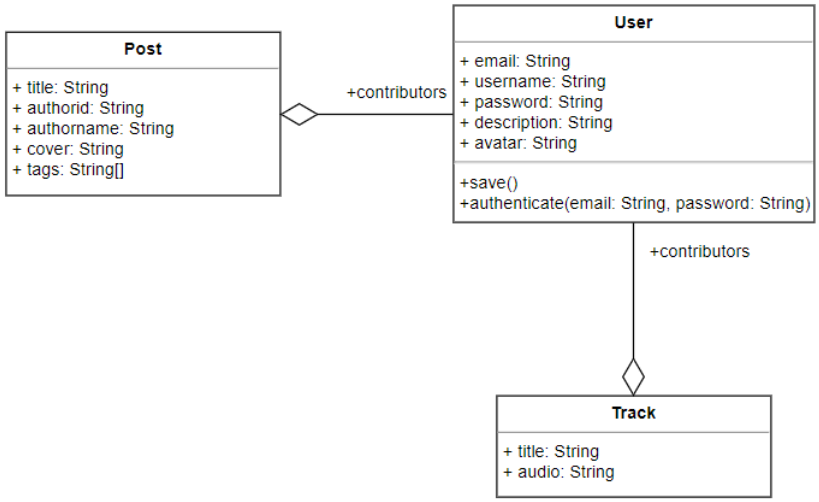


Fig. 3. Domain model of web application

Description of classes is next:

1. «Post» is a specific project and it consists of 6 attributes:

– Title of project, string;

– Authorid: id of author who created projects, string;

– Authorname: name of performer, string;

– Cover: name of project cover image file, string;

– Tags: array of string;

– Contributors: contains of id of participated users, array of string.

2. «User» consists of 5 attributes:

– Email: email of user that used to login, string;

– Username: name of user, string;

– Password: encrypted password of user, string;

– Description: description of user profile, string;

– Avatar: name of avatar image file, string.

3. «Track» represents certain song and it contains 3 attributes:

– Title: name of song, string;

– Audio: name of song audio file, string;

– Contributors: contains of id of participated users, array of string.

3. Defining of requirements

For the implementation of web application, the following requirements were formed:

1) Ability to create and manage personal portfolio page. It means that user can customize his own profile with avatar and information;

2) The function of adding project to the portfolio. When adding a project user has to upload cover of project, audio files and add information;

3) The function of mentioning persons involved in the creation of musical project;

4) Ability to search projects and content creators;

5) Home page with the latest projects of users.

4. Use case diagram

The origin and conceptual representation of the system in designing and development process is use case diagram. It helps to determine all main functional moments of interaction between system and user.

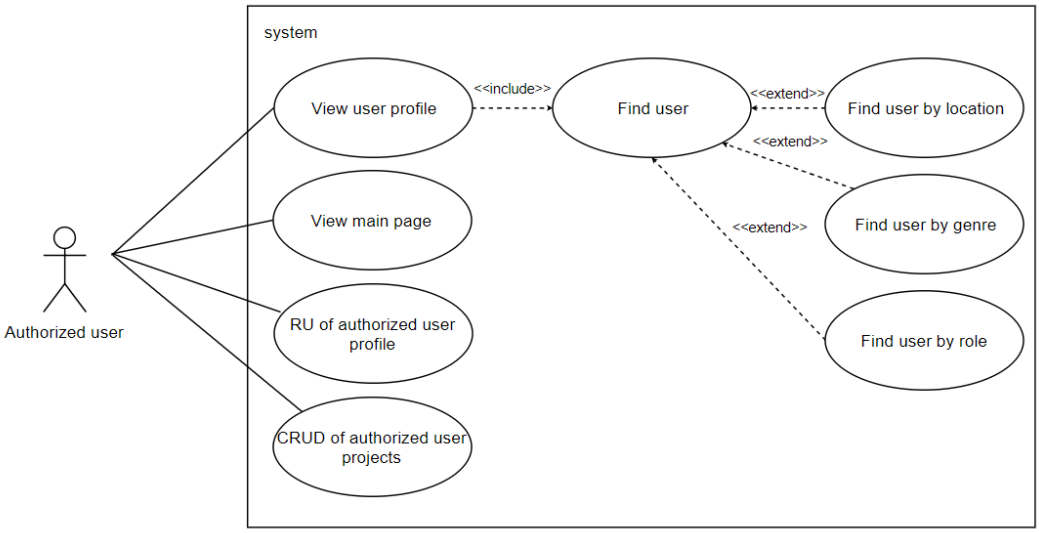


Fig. 4. Authorized user use case diagram

There are 3 roles in web-application – Authorized User (Fig. 5), Unauthorized User and Administrator (Fig. 6). The main actor of web application is authorized user so he has access to all the features of site while unauthorized user has not. The only difference between them it's that authorized user can manage his profile and create his own projects. Administrator has full rights and can manage all projects.

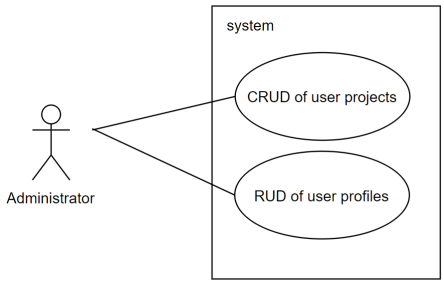


Fig. 5. Administrator use case diagram

5. Implementation tools

For server part of web application Node.js was chosen. Node.js is an open-source, cross-platform JavaScript run-time environment that executes JavaScript code outside of a browser. It gives programmer an ability to create server-side applications using Javascript. It has a huge number of additional libraries which expand opportunities for building applications and make development faster.

Express is a web application framework for Node.js designed for building web applications and APIs, which helps to organize web-application into MVC-architecture on server side. For creating a front-end template engine called Jade with UI framework Bootstrap 4 were used.

For development of database MongoDB was chosen. MongoDB – is a cross-platform document-oriented database program that classified as a NoSQL database program and it uses JSON-like documents. To manage relationships between data and to translate between objects in code and the representation of those object in MongoDB used Object Data Modeling (ODM) library Mongoose.

6. User's guide

The main page contains of user projects or releases with covers, names of projects and usernames. User can find other users, log in, sign up and click on projects for more information (Fig. 6).

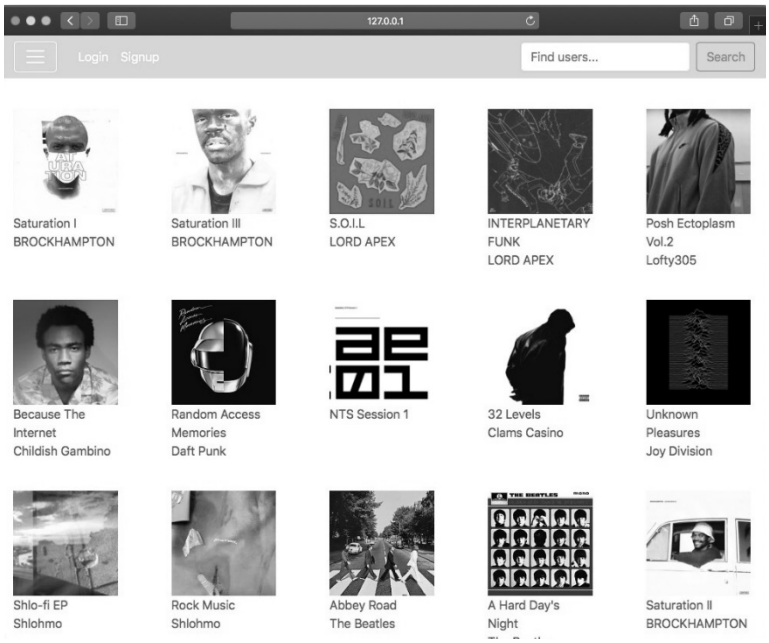


Fig. 6. The main page

The album page with elements such as name of release, name of user, all the participated users and some tags. Next to album information, we see songs with attached and playable audio files, their names and people involved in creating of song (Fig. 7).

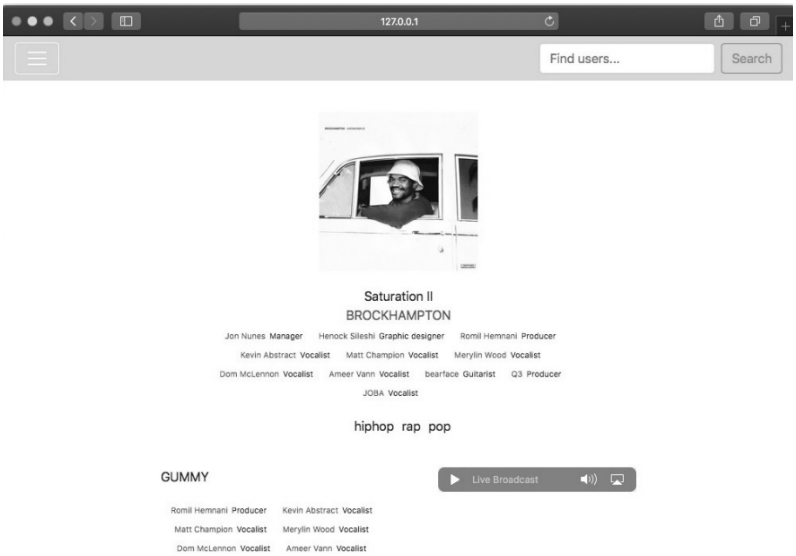


Fig. 7. Album page

Adding a project page provides with text fields such as title, tags, participated user names. Also users have to upload cover of project and choose roles of mentioned users. In

user profile users have ability to upload their avatars, to add information about them and contacts. Projects of user are presented under description block. (Fig. 8 and Fig. 9).

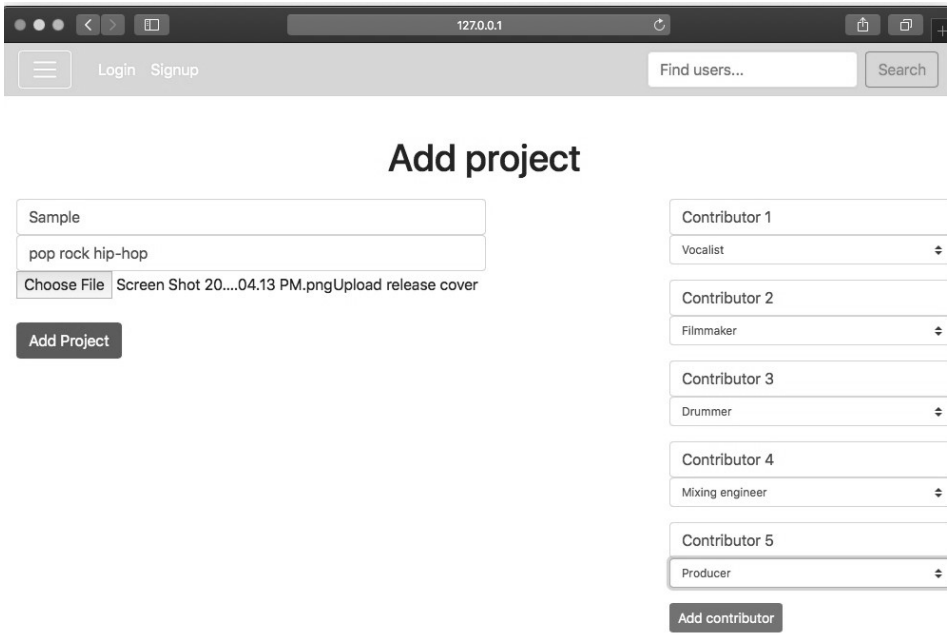


Fig. 8. Adding a project

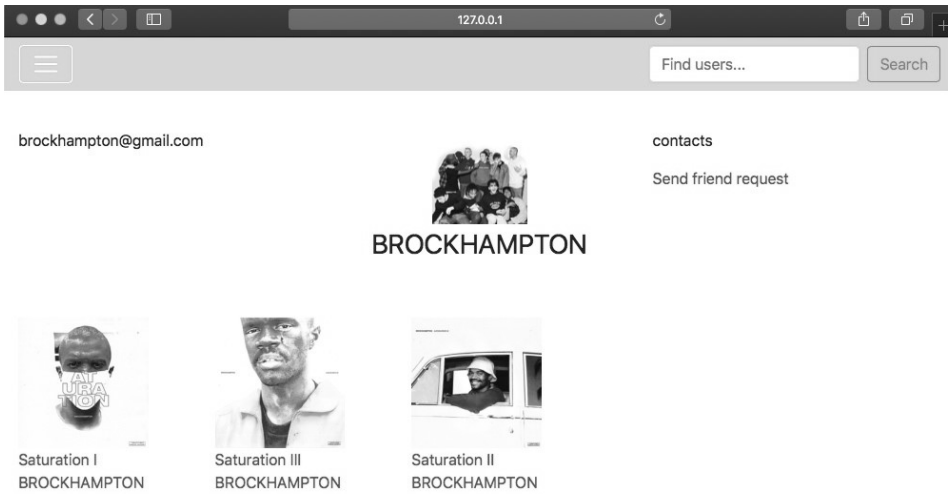


Fig. 9. User's profile page

Conclusion

The goal of this project was development of a web application that allows people involved in music to create online-portfolio to which they can attach their projects or be marked as participating in the creation of project. Also, this application has advantages compared to already existing solutions.

All objectives, which were defined in the beginning, are done. Application satisfies all requirements which were formed before implementation. So In that way, the goal and all objectives were accomplished.

REFERENCES

1. Global Music Report 2018 – IFPI, representing the recording industry worldwide. [online] Available at: https://ifpi.org/downloads/GMR2018.pdf [Accessed 31 May 2019]

2. Soundcloud.com – Music and podcast streaming platform. [online] Available at: URL: https://soundcloud.com/ [Accessed 31 May 2019]

3. Bandcamp.com – Music distribution service. [online] Available at: URL: https://bandcamp.com/ [Accessed 31 May 2019]

DETECTION OF UNDERLYING SURFACE, HORIZONTAL AND VERTICAL EDGES OF THREE-DIMENSIONAL OBJECTS BASED ON LASER SCANNING DATA

Y.L. Kostyuk, M.I. Litovchenko, A.A. Pichugina
Tomsk State University

Introduction

Laser scanning technology is widely used in various industries and solves a lot of production problems. The result of applying this technology is a three-dimensional point cloud. There is a lot of different software for handling obtained laser scanning data. An overview of a number of such software products is given in [1]. Each of these software products has various implemented algorithms for detection objects in point cloud, which are distinguished by detection accuracy, computational complexity, input data, user parameters, etc. So each algorithm has its own approach for handling point cloud and could be applied to detection certain objects.

A lot of algorithms are based on the idea of Hough Transform [2,3]. The main problem of using Hough Transform for detection objects is high computational complexity, which depends on how compound the object is. Therefore, actual task is to develop more efficient detection algorithms for some type-object.

1. Point cloud structuring

The first step includes preparation of point cloud for further research. Proposed approach of point cloud structuring in [4] allows to reduce the computation complexity of most detection algorithms to a linear order $O(n)$.

The space of point cloud is divided on voxels which mean cells in three-dimension spaces with regular size *d*. Each point of cellular partition uniquely belongs to a certain voxel: the point coordinates (x_i, y_i, z_i) are divided on cell size *d* and obtained indexes define the voxel. The list of entered points with their quantity are formed for each voxel.

2. Detection of the underlying surface

Underlying surface includes all components referring to the earth`s surface. In a simple case the underlying surface means the terrain surface. So the points reflected from the terrain surface have the lowest *Z* coordinates. Using obtained data structure it is easy to detect dead-level underlying surface by scanning the low level voxels.

If the surface presents ambiguous terrain, then it is necessary to use algorithm from [5]. Succinctly describing of algorithm is demonstrated as follows:

- consistently searching for cells in point cloud structuring that have voxel on zero level and collect them into list;
- checking eight neighboring cells for each item from certain list. If a neighboring cell has a voxel whose lower level is not more than 1 from current cell, then it is added to the end of the list.

The model of detected surface is represented by flat triangulation on selected points (Fig. 1). Handled points are deleted or ignored from point cloud for further research.

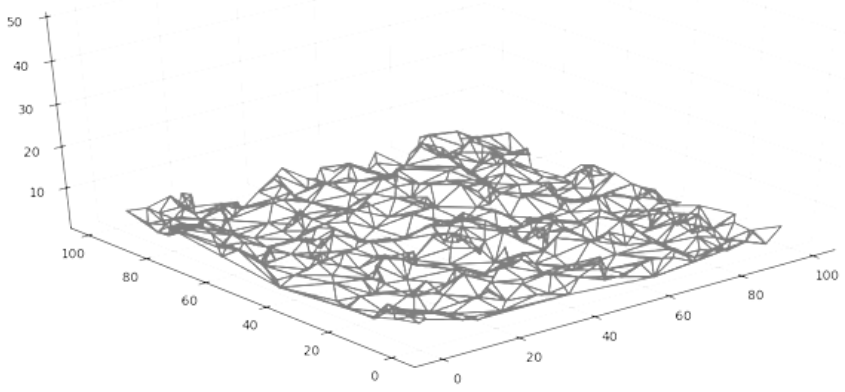


Fig. 1. The triangulated model of the underlying surface

3. Detection of horizontal edges of the objects

The algorithm for detection of horizontal edges is based on the idea of Hough Transform. This approach is effective since the horizontal edge is given by an equation with only one parameter: $z = D$, where D is a distance from origin to plane. So the accumulator array would be one-dimensional with regular size m and would be set to zero.

The voting process runs for each voxel where voxels` points vote for certain cell in accumulator A: $A[z_i/m]$, where z_i is coordinate of current point. Instead of using all the points it is possible to use only the average points in voxel to speed up the voting process.

The global maximum in the accumulator indicates the layer with one or several horizontal edges. Single edge could be separated from the layer by algorithm for finding connected components (Fig. 2). The model of detected edge is represented by triangulation in the edge of the plane and convex hull that illustrates boundaries of the edge.

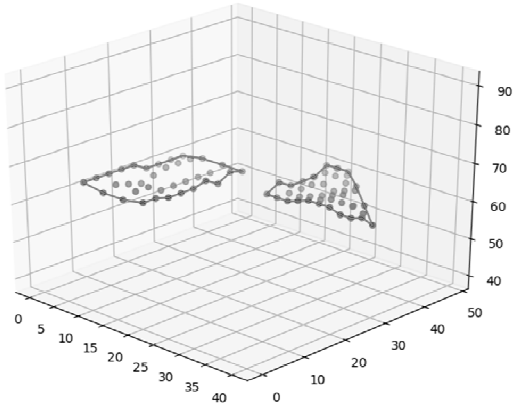


Fig. 2. Two horizontal edges of the same layer

After researching the global maximum the next one is found and the same operations are done with it.

4. Detection of vertical edges of the objects

The detection of vertical edges works more effective after detecting and deleting points belonging to the underlying surface, horizontal edges and single high objects. The vertical edges usually refer to objects such as buildings.

The vertical edge is represented as a projection on the xOy plane as a straight line segment. The algorithm for detection one is also based on the idea of Hough Transform. It is nec-

essary to search for a part of the projection of the vertical edge that includes the maximum number of entered points.

The laser scanning data always has unpredictable nature, so it is rational to use a variable parameter such as length of the edge which consists of number of cells. Let`s pretend that minimum length of the searching part of the projection is probably passing through four cells. Then for each cell (i_1, j_1) of the cellular partition we have twelve unique positions for the other end of the segment (i_2, j_2).

The end of the segment (i_2, j_2) rotates clockwise from the initial position (Fig. 3) to the final position (Fig. 4).

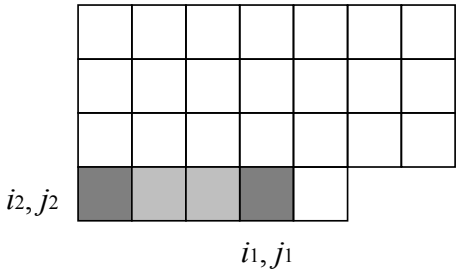


Fig. 3. Initial position

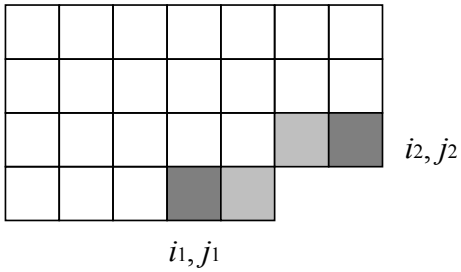


Fig. 4. Final position

The global maximum among fours of cells contains sought-for part of the projection. The detected segment is approximated by method of least squares. If the both ends of segment have neighbor cells, which continuous current segment and have enough points, then the segment would be extended and approximated again.

The equation of the approximated edge is $Ax + By + 1 = 0$, where coefficients A, B are calculated on the formulas [6]:

$$A = \frac{ce - bd}{c^2 - ab}, B = \frac{cd - ae}{c^2 - ab},$$

$$a = \sum_{i=1}^{n} x_i^2,\ b = \sum_{i=1}^{n} y_i^2,\ c = \sum_{i=1}^{n} x_i y_i,\ d = -\sum_{i=1}^{n} x_i,\ e = \sum_{i=1}^{n} y_i.$$

In order to build the model of the detected edge it is appropriate to do some transformations. Points with coordinates (x_i, y_i) that belongs to the edge are transformed to coordinate system xOy of the edge. Coordinates z_i stay without changes and become coordinates on Y axis on the edge plane.

There are three cases:

1. $A = 0$, coordinates x_i stay without changes, among them finding maximum and minimum.
2. $B = 0$, coordinates y_i become coordinates x_i, among them finding maximum and minimum.
3. $A \neq 0$ and $B \neq 0$, coordinates of the intersection point of the perpendicular from the current point (x_i, y_i) to the projection of the edge (x_{0i}, y_{0i}) are calculated in accordance with the formulas given in [6]:

$$x_{0i} = -\left(A + B\left(Ay_i - Bx_i\right)\right) \Big/ \left(A^2 + B^2\right),\ y_{0i} = \left(A\left(Ay_i - Bx_i\right) - B\right) \Big/ \left(A^2 + B^2\right).$$

On acquired coordinates (x_{0i}, y_{0i}) minimums $(x_{\min}, y_{\min})$ and maximums $(x_{\max}, y_{\max})$ are found. Point $(x_{\min}, y_{\min})$ become origin of X axis in the plane of the edge. Original coordi-

nates (x_i, y_i) of point recalculated and become new coordinate $x_{\text{new } i}$ on X axis in the plane of the edge with formulas [6]: $x_{\text{new } i} = \sqrt{(x_{0i} - x_{\min})^2 + (y_{0i} - y_{\min})^2}$.

The model of detected edge is represented by triangulation in the edge of the plane and convex hull that illustrates boundaries of the edge (Fig. 5).

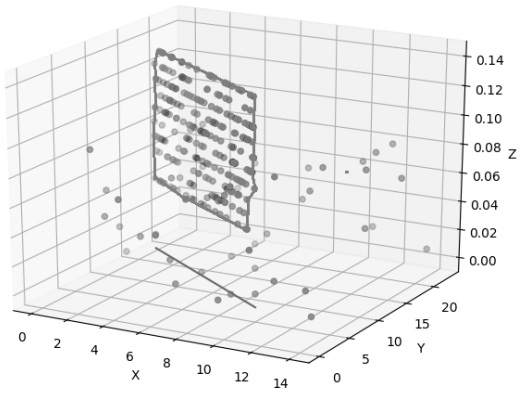


Fig. 5. The model of vertical edge with convex hull

Conclusion

The proposed point cloud structuring allows us to develop effective algorithms for detection different type-objects. The required memory capacity for such structure reaches $O(n)$. When applying an efficient sorting algorithms the computation complexity of creation this structure reaches $O(n \cdot \log n)$, where n is the number of points in the cloud. In this paper we propose several algorithms for detection the underlying surface, horizontal and vertical edges of three-dimensional objects. The previous two usually compose complex objects such as buildings or facilities.

REFERENCES

1. *Medvedev V.I., Sarychev D.S., Skvortsov A.V.* Predvaritel'naya obrabotka dannykh mobil'nogo lazernogo skanirovaniya v sisteme IndorCloud [Preliminary processing of mobile laser scanning data in the IndorCloud system] // SAPR i GIS avtomobil'nykh dorog. – 2014. – No. 2(3). – P. 67–74.

2. *Khoshelham K.* Extending Generalized Hough Transform to Detect 3D Objects in Laser Range Data // ISPRS Workshop on Laser Scanning 2007 and SilviLaser 2007. Espoo, Finland, 12-14 September 2007. – Finland, 2007. – Vol. 36. – Part 3. – P. 206–210.

3. *Dalitz C., Schramke T., Jeltsch M.* Iterative Hough transform for line detection in 3D point clouds // Image Processing On Line. – 2017. – No. 7. – P. 184–196.

4. Rodríguez-Cuenca B. et al. Automatic detection and classification of pole-like objects in urban point cloud data using an anomaly detection algorithm // Remote Sensing. – 2015. – Vol. 7. – No. 10. – P. 12680–12703.

5. *Kostyuk Y.L., Litovchenko M.I., Smirnova A.A.* Raspoznavaniye trekhmernykh ob"yektov na mestnosti po dannym lazernogo skanirovaniya [Detection of three-dimensional objects on the terrain according to laser scanning data] // Information technologies and mathematical modeling (ITMM-2017) Materials of the XVI International Conference named after A.F. Terpugov, 10-15 Sept. 2018. – Tomsk: NTL. – 2018. – P. 193–198 (In Russian).

6. *Kostyuk Y.L., Litovchenko M.I.* Raspoznavaniye graney trekhmernykh ob"yektov po dannym lazernogo skanirovaniya [Detection of edges of three-dimensional objects according to laser scanning data] // Information technologies and mathematical modeling (ITMM-2017) Materials of the XVI International Conference named after A.F. Terpugov. – 2017. – P. 55-61 (In Russian).

MULTIPLAYER BLACKJACK

A. Rodionov, A. Shkurkin
Tomsk State University

Introduction

Over the years the origin of Blackjack, like many other games, has eluded researchers for a long time and which continues to be hotly debated to this day. Till date there has been no clear consensus, but most agree it probably originated in French casinos around 1700, where it was called Vingt-et-Un which, translated, means twenty-one [1]. Most believe it was probably derived from the French card games, Chemin de Fer, and French Ferme, which were in vogue at that time.

Another of its version was prevalent in Spain, which was called 'One and Thirty'. The basic rule of the Spanish version was to reach 31 with a minimum of three cards.

A theory that's also doing the rounds is that Blackjack was an invention by the Romans. It is believed that Romans used to play this game with wooden blocks of different numerical values. This theory holds some weight as Romans loved gambling, but that by no means confirms the theory.

Among the various versions of the game, 'Vingt-et-Un' or '21' continued to grow in popularity and reach. It gradually spread to North America thanks to French colonists and soon after it was played throughout the continent. The rules of Blackjack were then different from those of modern or contemporary Blackjack. Just to elaborate, in this form of Blackjack, only the dealer was permitted to double. Also, a betting round was there between each of the playing cards dealt.

The game was still termed '21' when it gained popularity in Nevada in 1931 as the State first chose to make gambling legal. To draw more people to the game, some casinos then offered a special bet: A hand featuring either of the black jacks (the Jack of Spades or that of Clubs) plus the Ace of Spades would pay 10-to-1 odds on the lucky player's bet. Although casinos later discontinued this peculiar payout, the name 'blackjack' or 'black jack' remained, as the game is known today [2].

Blackjack was not conceived by any one person at one definite point of time. On the contrary, blackjack has evolved over centuries; it continues to evolve and grow even today, thanks to the Internet. Fuelling its popularity the online casino industry has developed various versions of the game which has truly revolutionized it.

The goal is to develop the web-application that allows you to play blackjack in multiplayer mode.

Objectives:

- Formulate requirements for the developing application.
- To study the necessary technical tools for the development.
- To design the web-application architecture.
- To develop the web-application.

Disadvantages of online Blackjack equivalents

Despite the abundant variety of online blackjack platforms, there are have some drawbacks like:

- **Advertisement**, that's nobody like;
- **Singleplayer mode**, that's opposed to multiplayer;
- **Non-adaptive layout**, that's bad for UX.

Blackjack lobby

Before playing blackjack the player will face the game lobby where 2 things can be done:

- **Create room**.

• **Join room**.

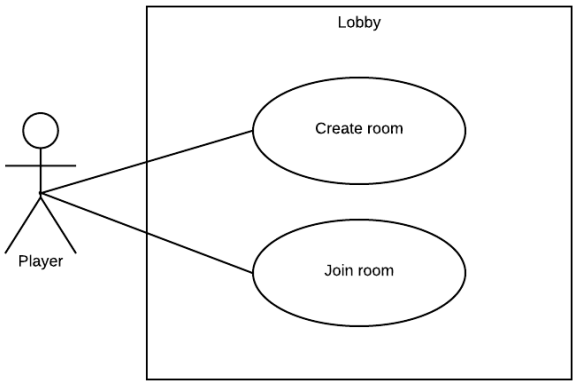


Fig. 1. Use case lobby diagram

Each room can accommodate up to three players, inclusive, but if a player lacked space in the rooms or for some other reason does not want to be in the available rooms, he can create a new room and start playing in it.

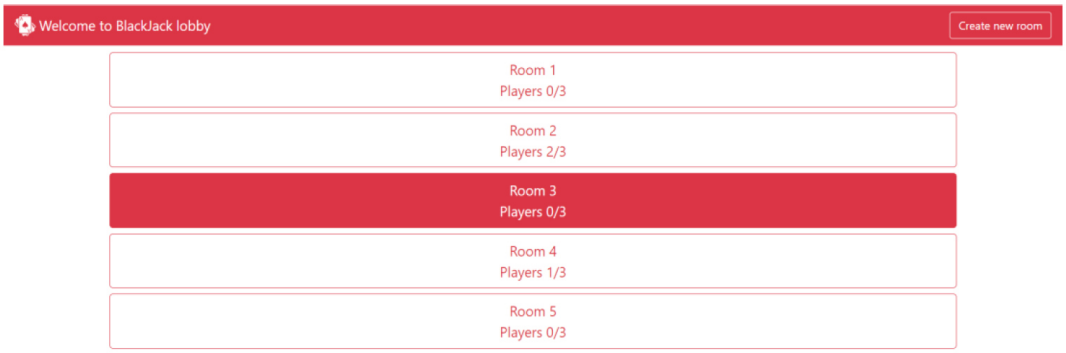


Fig. 2. Lobby screen

Blackjack gameplay

A blackjack game consists of a recurring round. In turn, each round is divided into four stages:

• **Stage №1** – players are makes a bet onetime until time runs out;

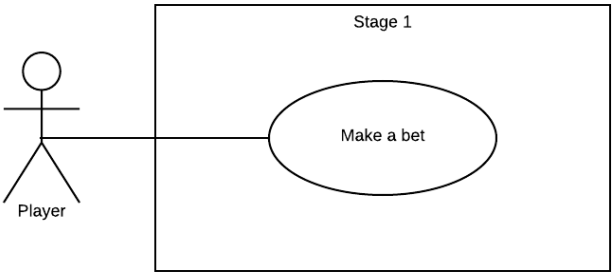


Fig. 3. Use case Stage №1 diagram

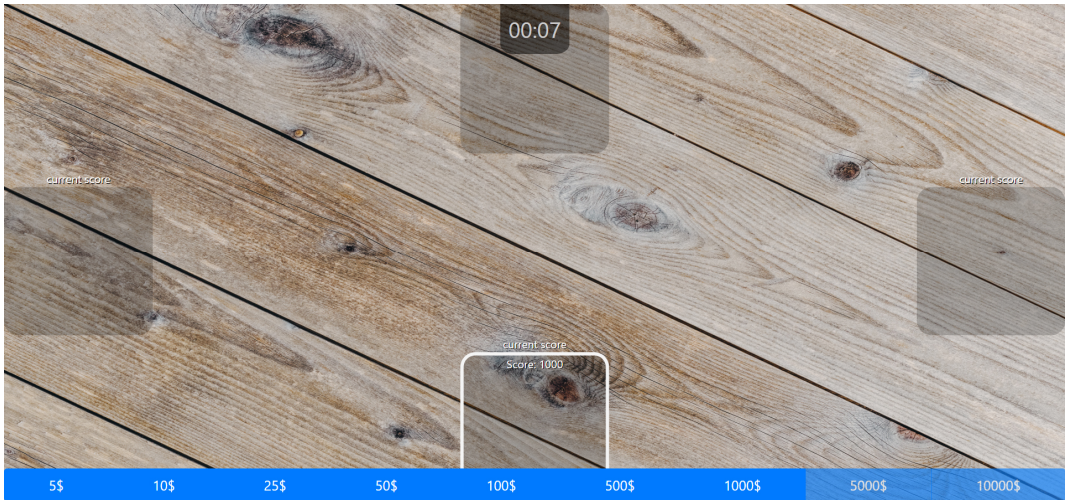


Fig. 3.Example of Stage №1

- **Stage №2** – players are takes cards from the deck (hit) until everyone says stand or time runs out; Player takes the cards until he decides that he has enough or the time runs out or he does not have more than 20 points;
- **Stage №3** – the dealer takes cards from the deck until he decides that he has enough cards or he does not have more than 20 points;
- **Stage №4** – determining the winners of the round and the distribution of the bets between players from room.

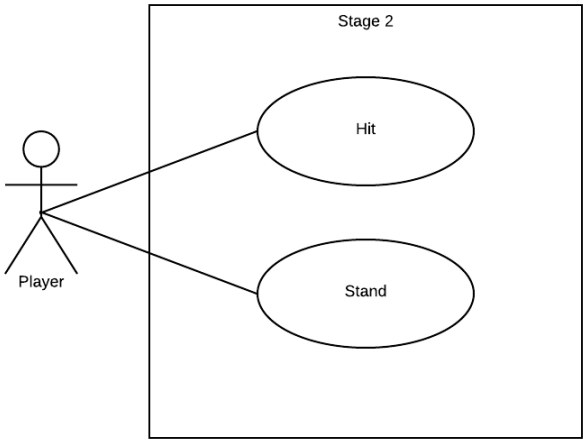


Fig. 4. Use case Stage №3 diagram

Development tools

Frontend:
- **HTML5**;
- **CSS3**;
- **JavaScript**;
- **Greensock** – javascript library for frontend animation.

Backend:
- **NodeJS** – open-source, cross-platform JavaScript run-time environment that executes JavaScript code outside of a browser [3];

- **Express** – server framework for Node.js;
- **Socket.io** – javascript library that enables realtime, bi-directional communication between web clients and servers.
- **Webpack** – open-source JavaScript module bundler.
- **Nodemon** – is a utility that will monitor for any changes in your source and automatically restart server.

Conclusion

Blackjack has an amazing story, a lot of realized variations and now presented took the basic advantages and discarded, a lot of minuses and turned into a new variation of the well-known gambling game. Nowadays lobby and 2 stages of the blackjack are done. In the future, the remaining stages of the game will be implemented.

REFERENCES

1. History of Blackjack – The World of Playing Cards. [online] Available at: https://www.wopc.co.uk/history/blackjack/blackjack [Accessed 29 May 2019].

2. Blackjack. En.wikipedia.org. [online] Available at: https://en.wikipedia.org/wiki/Blackjack [Accessed 29 May 2019].

3. What exactly is Node.js? – Medium, freeCodeCamp. [online] Available at: https://medium.freecodecamp.org/what-exactly-is-node-js-ae36e97449f5 [Accessed 29 May 2019].

DEVELOPMENT OF A WEB APPLICATION TO DETERMINE THE HAZARD CATEGORY OF FOOD ADDITIVES

E. Skripachenko, A. Shkurkin
Tomsk State University

Introduction

In nowadays world almost all food products contain food additives. These food additives are used in technological purposes in production process, packing, transportation and also for giving wishful features, such as taste, color, consistency and etc.

Today there is a problem that not all additives are absolutely safe for human organism. Often food additives are involved into frequent causes of allergy, asthma and others problems with health, which takes place in XXI century. In pursuit to follow healthy lifestyle, some people try to refuse from them and other go to concessions and discover contents written on the package. With a high probability customer will understand hazard of this additive not right away and there is a need to study an additional source.

So, the goal is to develop web application to determine the hazard category of food additives.

This goal defines next objectives:

1. Analysis of existing solutions.
2. Choosing tools for implementation.
3. Web application designing.
4. Web application implementation.

1. Subject area research

The main source of knowledge about food additives is Codex Alimentarius [1]. In this Codex you can find any information about products whether it was processed or not (even semi-products). Besides standards on every single product, also there are rules of marking, food hygiene, pesticide content and regulates questions about food additives, which is the most interesting direction in domain scope.

All additives are divided in three big groups: Natural, Artificial and Synthetic. Natural additives are already used for over 1000 years and divided on animalistic (carmine dye from cochineal aphid), herbal (from algae and plants) and mineral (chalk, salt and etc.).

Artificial additives are characterized by the fact that they occur naturally, but due to technological process development they are synthesized in production in order to reduce the end price of product, because mining of them is usually very expensive process. Now there is no clear evidence of benefits and harms of the artificial additives, but often customers do not trust "Identical to natural" elements written on a package in products' content, and that fact influences on their final decision.

Synthetic additives are obtained only by synthesizing and they cannot be met in nature. In recent years, many additives have appeared that make life easier for people (for example, sugar substitutes for diabetics) and protect against the viral bacteria found in some products, such as meat. But these additives also aren't fully safe for peoples' health, and consuming in big quantities can cause bad consequences.

Among food additives exists next dividing into categories:

– *colorants* (is used to reach necessary color);
– *preservatives* (is used for part or full stop microbial growth in foods and as sequence prevents changing of food's color, taste and etc.);
– *antioxidants* (prevents excessive oxidation of cells of an organism);
– *stabilizers* (is used to giving to products wishful form, texture and saving consistency);
– *emulsifiers* (to make an emulsion from unstirring liquids, e.g. oil and water);
– *anti-caking agents* (to prevent lumps in powder products);
– *flavor and aroma enhancers*;
– *antifoams, coatings, flour improvers*.

All sources usually divide food additives into five hazard categories: very low, low, medium, high, very high. So it was decided to use same division on categories of danger.

2. Analysis of existing solution

Consider some popular web applications about food additives.

«Добавкам.нет» [2]

This application can provides food additive origin, synonyms and dangerous level of these additives. Also shows composition of products with additives, but it hasn't fast search. On fig. 1 the main page of «Добавкам.нет» is represented.

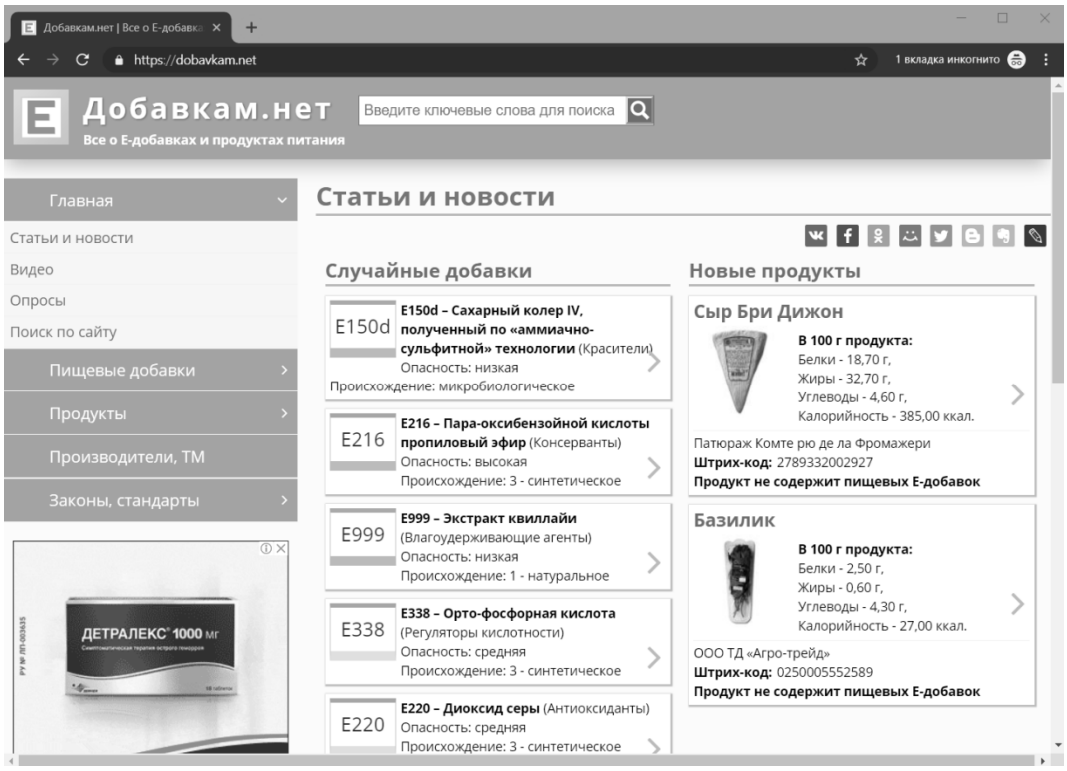


Fig. 1. The main page of «Добавкам.нет»

«БазаДобавок.ру» [3]

This application also can provides food additive origin, dangerous level and synonyms of additives, but has not ability to show composition of products. The advantage of this application is fast search, which gives you results of query at once. On fig. 2 the main page of «БазаДобавок.ру» is represented.

«Prodobavki.com» [4]

Third application's name is «Prodobavki.com». It doesn't give you full information about products because there is no ability to view hazard category or origin of food additives. It is only good for discovering synonyms for them. On fig. 3 the main page of «Prodobavki.com» is represented.

Fig. 2. The main page of «Prodobavki.com»

Fig 3. The main page of «Prodobavki.com»

After analyzing the above and similar web-applications, it comes to conclusion:

– Not all applications provide full information about products and theirs food additives;
– Fast search is necessary for user because it helps to find information much more faster than discovering them through sections of application;
– Good and friendly design is also required and influences on users' experience.

3. Domain modelling

Domain modelling is the first stage of web application modelling, which represents the main aspects of developing system. Also it allows to reveal all the relations between entities, theirs business logic and to formalize all information about them.

On fig. 4 domain model of web application is represented.

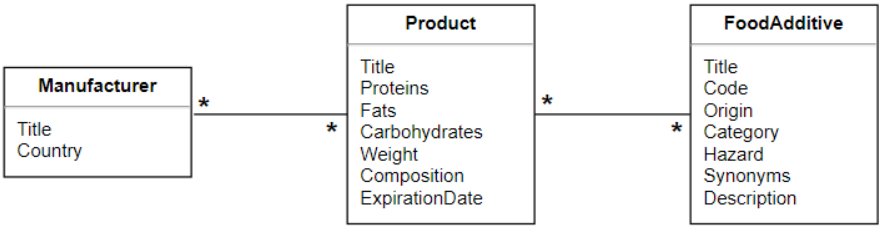


Fig. 4. Domain model of web application

Description of classes is next:

1. «Manufacturer» consists of 2 attributes:
– Title of manufacturer, string;
– Country, where manufacturer is registered, string.
2. «Product» consists of 7 attributes:
– Title: title of product, string;
– Proteins: how many proteins is in product, integer;
– Fats: how many fats in product, integer;
– Carbohydrates: how many carbohydrates in product, integer;
– Weight: weight of product, integer;
– Composition: composition of product, string;
– ExpirationDate: conditions and expiration date of product, string.
3. «FoodAdditive» represents food additive and consists of 7 attributes:
– Title: title of an additive, string;
– Code: code of an additive, string;
– Origin: one of three origins of an additive, string;
– Category: one of eight categories of an additive, string;
– Hazard: hazard of an additive, from 0 – very low to 4 – very high, integer;
– Synonyms: synonyms, which is used in different compositions, string;
– Description: description and useful information about additive, string.

There is «Many-to-Many» connection between class «Manufacturer» and class «Product» and between and class «FoodAdditive».

4. Defining of requirements

For the implementation of web application, the following requirements were formed:

1. Providing full information (such as: title, origin, code, hazard, category, synonyms and description) about food additive. This information should be understandable for user.

2. Showing show all food additives in concrete products and their manufacturers, and user should easily go through links between them.

3. Providing fast search, when user just type their request in search field and application shows results without excesses.

4. Admin page functionality, where administrator can CRUD additives, products and manufacturers and set dependencies between them.

5. Application should have easy-to-use, not overloaded UI which will not discourage users with very bright colors.

5. Use case diagram

The origin and conceptual representation of the system in designing and development process is use case diagram. It helps to determine all main functional moments of interaction between system and user. On fig. 5 use case diagram of web application is represented.

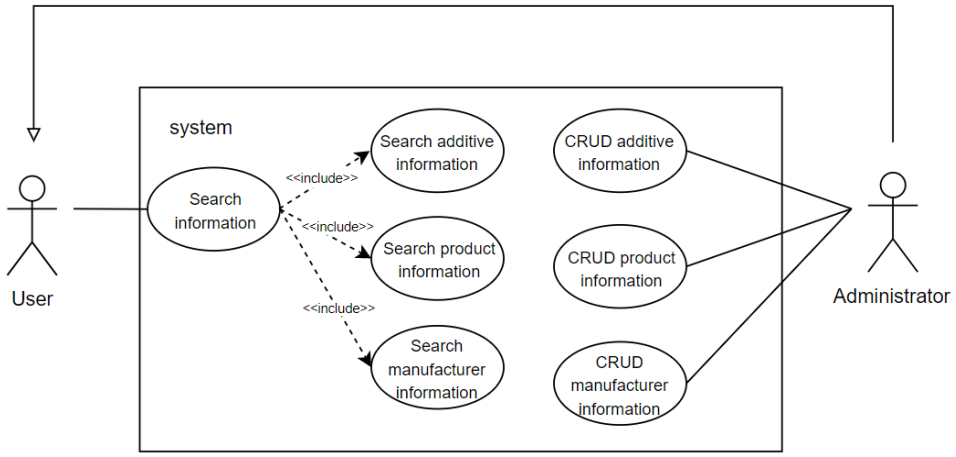


Fig. 5. Web application use case diagram

There are 2 roles in web-application – User and Administrator. User has restricted rights – he can only search and read information about additives, products and manufacturers. Administrator has full rights and can CRUD all information about additives, products and manufacturers and set dependencies between them. Also, Administrator can be User in testing purposes.

6. Implementation tools

For server part of web application Node.js was chosen [5]. Node.js is an execution environment and server platform for working with JavaScript. This is a good tool for creating simple (and not only simple) web applications due to powerful development instruments. Also Node.js is known for huge library of additions which may be required in development and implementation process.

Express.js is a web application platform, which is used with Node.js to organize web-application into MVC-architecture on server side. In conjunction with this EJS was chosen as view engine due to syntax, which was borrowed from the engine views of the Web Forms platform on ASP.NET. In addition to EJS Bootstrap 4 was selected for UI implementation.

MySQL was defined as the DBMS.

7. User's guide

For User the main page allows to type request and make a search by clicking Search button. Also User can choose between 3 categories – 'Additives', 'Products' and 'Manufacturers'. Clicking of them leads to page where all chosen entities are represented, sorted by alphabet order (fig. 6).

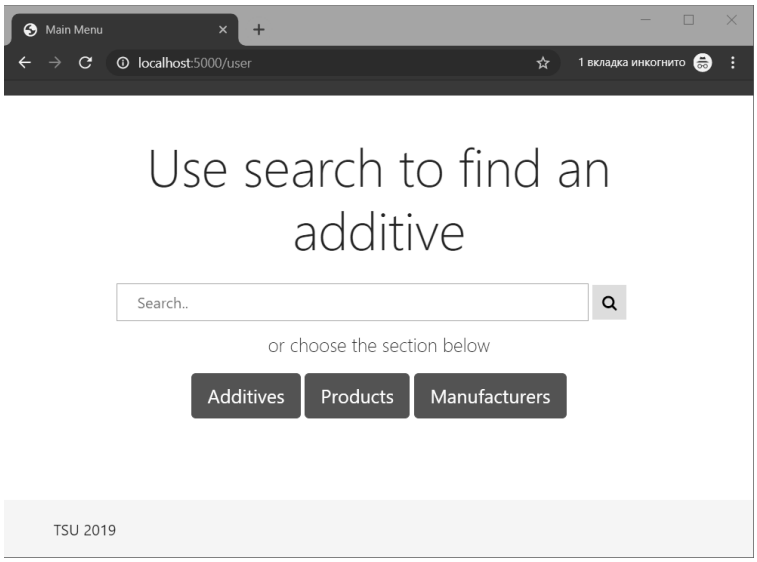


Fig. 6. User's interface

After section's and record's choosing, user sees all information about additive, product or manufacturer (fig. 7).

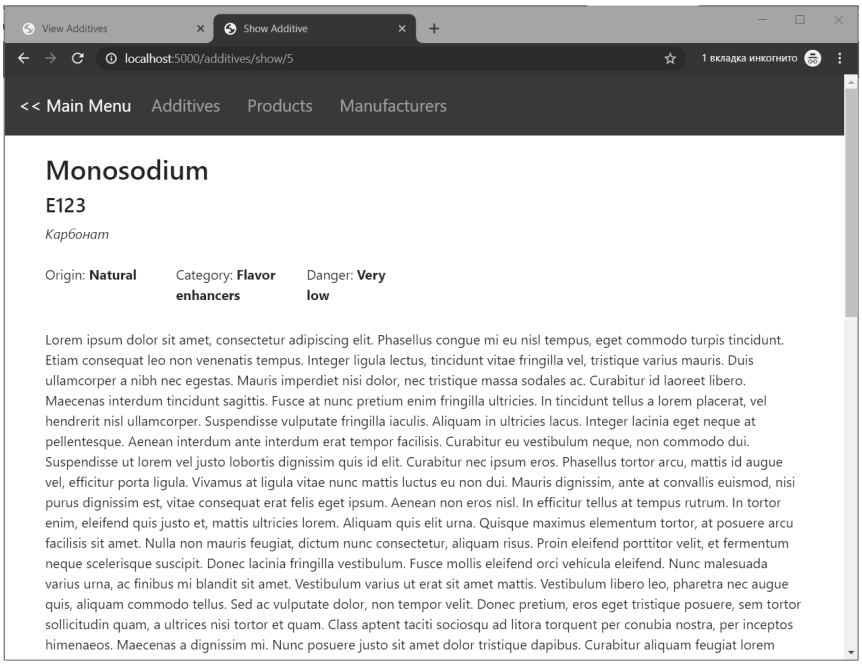


Fig 7. Information about additive

Admin page almost looks like as User's, but has not search field. By clicking section's button, application leads to all entities in this section (fig. 8, 9).

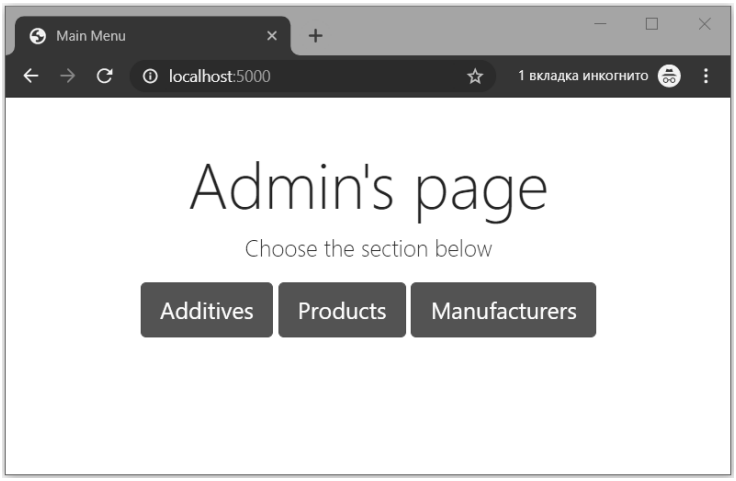


Fig. 8. Admin page

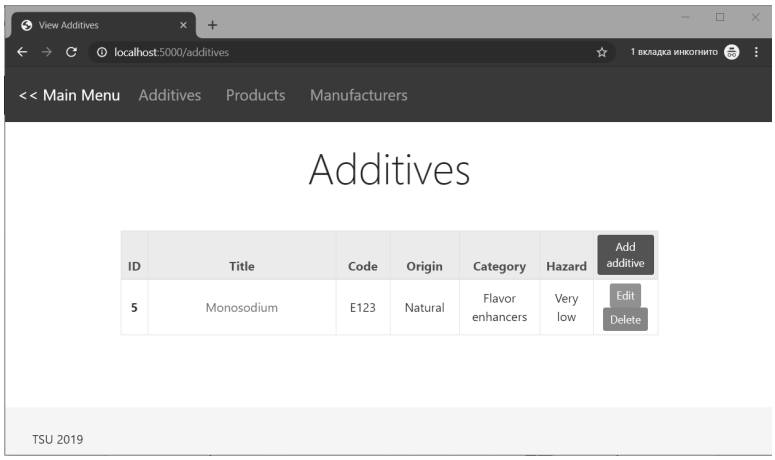


Fig. 9. Admin's page of Additives section

Admin can CRUD all entities, and after clicking "Edit" button administrator edit an entity. "Delete" button deletes entity, "Add additive" adds new entity. Read operation is done by clicking on entity's title. This functionality is the same for the rest of sections: read operation is the same as for User, and entity's adding page looks like editing page, but without prefilled fields according to attributes of entity (fig. 10).

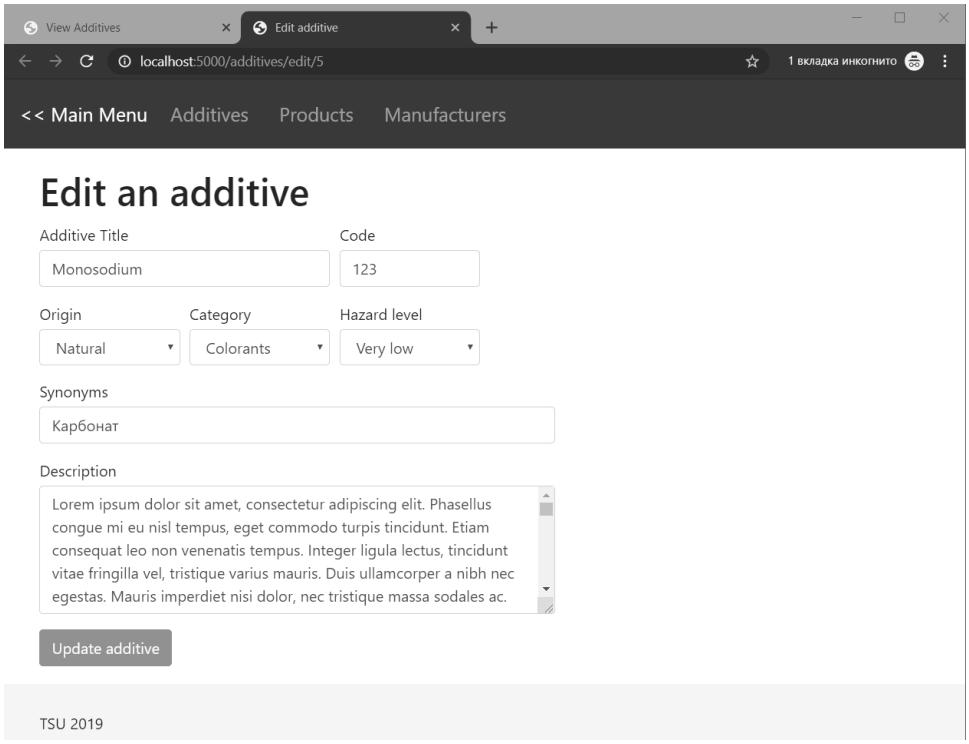


Fig. 10. Entity's edit page

Conclusion

The goal of this project was development of a web application to determine the hazard level of food additive. This web application will be useful for those who is indifferent to content of product they eat. And, also, this application is good alternative to already existing solutions.

All objectives, which were defined in the beginning, done, namely: existing solutions were analyzed, tools for implementation were chosen, web application was designed and implemented. Application satisfies all requirements which were formed before implementation.

In that way, the aim and all objectives were accomplished. Improvements like mobile device support and product's barcode camera scanner were already planned in the next versions of this project.

REFERENCES

1. Codex Alimentarius. International Food Standards. [online] Available at: http://www.fao.org/fao-who-codexalimentarius/home/en [Accessed 29 May 2019].

2. Добавкам.нет – Все о Е-добавках и продуктах питания. [online] Available at: https://dobavkam.net/ [Accessed 29 May 2019].

3. БазаДобавок.ру – справочник пищевых добавок и БАДов. [online] Available at: http://bazadobavok.ru/ [Accessed 29 May 2019].

4. Prodobavki.com – Вредные пищевые добавки. [online] Available at: https://prodobavki.com/. [Accessed 29 May 2019].

5. What exactly is Node.js? – Medium, freeCodeCamp. [online] Available at: https://medium.freecodecamp.org/what-exactly-is-node-js-ae36e97449f5 [Accessed 29 May 2019].

THE RECOGNITION OF SINGLE HIGH OBJECTS ON TERRAIN ACCORDING TO THE LASER SCANNING DATA

Y.L. Kostyuk, A.A. Smirnova, A.A. Pichugina
Tomsk State University

Introduction

There are many software products for points cloud [1]. The algorithms that are used in such programs are diverse, but not universal and work ambiguously for objects of different types.

Many researches are devoted to development and modification of algorithms aimed to classification of points and recognition of objects in the cloud. Many works consider algorithms based on neural networks, on triangulation. A lot of algorithms are based on Hough Transform [2, 3].

1. The proposed approach

The Hough Transform structure is partially used in the implementation of the proposed algorithms.

The phases of proposed approach:

1) initial processing of points;

2) cellular dividing of the initial points cloud [4];

3) recognition of single high objects.

Let's consider in detail the last stage. Before applying this procedure, it is recommended to recognize and remove points belonging to the «underlying» surface.

In the resulting structure, we look for the global maximum — the cell in which the point with the maximum height is located, and check for the absence of unfilled voxels.

For a voxel containing a maximum, its neighbors are traversed at the same level and form a single layer. Similarly, the group of voxels of the layer below is highlighted, etc. For each layer in the (horizontal plane) it is possible to build a convex hull to determine the boundaries of the object in the layer.

2. Recognition of cylindrical objects by method of convex hull

For the identified single high object, assuming that the object is a vertical cylinder, one can determine the center and diameter of each part of the object located in a certain layer. For an ideal cylindrical object, after constructing a convex hull for the coordinates in X and Y the differences between the maximum and minimum values are calculated:

$$\Delta x = X_{\max} - X_{\min},\ \Delta y = Y_{\max} - Y_{\min},$$

after which the average diameter D is calculated:

$$D = \left(\Delta x + \Delta y\right)/2 .$$

Most often, in practice, none of the objects is an ideal cylinder entirely, but appears as a complex object, in which only a certain part forms a cylinder. Therefore, this approach will not always give accurate results.

3. Hough Transform for section cylinder in a plane *xOy*

An alternative option that gives the best estimates of the parameters. The section of the cylinder in the *xOy* plane is a circle that can be specified in the form:

$$\left(x - x_c\right)^2 + \left(y - y_c\right)^2 = R^2 ,$$

$\left(x_c, y_c\right)$ – coordinates of the center of the circle, *R*– its radius.

We define the radius *R* as the maximum possible radius for objects of a given type.

Of all the cells in the selected object, it is proposed to find the cell with the largest number of points in it, that is, for all voxels. We outline this cell with a 5x5 area, and the cell itself is located in the center of the region. Let the radius of a cylindrical object be limited to the size of a single cell, so the circle can be located in a 3x3 area, as shown in fig. 1.

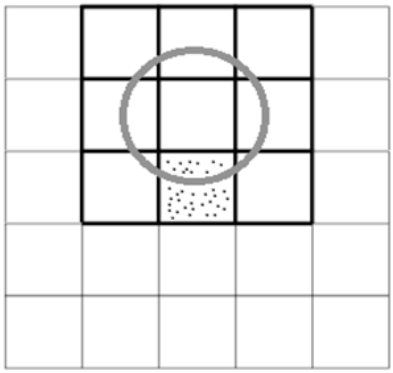


Fig. 1. Viewing of the area 3x3

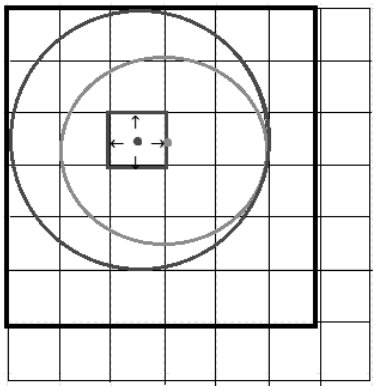


Fig. 2. Displacement of the centre

Among all 9 possible 3x3 positions within the 5x5 area, a global maximum is searched, which contains the greatest number of points. Further in this area you need to find the center of the circle, each cell is divided into smaller cells, so that the results are more accurate. Each point of the 3x3 area votes for one or another center of a circle with a maximum radius. Further, the radius decreases with a given step and the vote is again held, while the center of the circle may remain the same or shift, as shown in fig. 2.

4. Construction of the circle on three points

Another alternative option search for the center and diameter of the desired cylindrical object is based on the principle of constructing a circle by three points that define the circle passing through them. The center of the circle is the point of intersection of the mid-perpendiculars to two segments, constructed from the points P_1, P_2, P_3.

As in the previous version, use the area of 5x5 and 3x3 with the highest number of points. The 3x3 area is divided by a fine grid and we hold a vote, going through all the three points.

The center of the required circle (x_0, y_0) is calculated by the formulas [5]:

$$\begin{cases} (x_2 - x_1)x_0 + (y_2 - y_1)y_0 = A, \\ (x_3 - x_2)x_0 + (y_3 - y_2)y_0 = B, \end{cases}$$

$$A = (x_2 - x_1)(x_1 + x_2)/2 + (y_2 - y_1)(y_1 + y_2)/2,$$

$$B = (y_3 - y_2)(y_2 + y_3)/2 + (x_3 - x_2)(x_2 + x_3)/2.$$

Next is the global maximum - the largest number of points that voted for a certain center of the circle (x_0, y_0). Mean square of radius for the found circle:

$$R_m^2 = \frac{1}{k}\sum_i \left[(x_i - x_0)^2 + (y_i - y_0)^2 \right],$$

k – the amount of points which voted for given center (x_0, y_0), (x_i, y_i) – coordinates of current looked through point.

In fig. 3, the marker '◦' denotes points of the surface of a cylindrical object. The marker 'x' indicates noises. Fig. 4 shows an example of layering a single high cylindrical object. The centers of the layers are found using the convex hull algorithm.

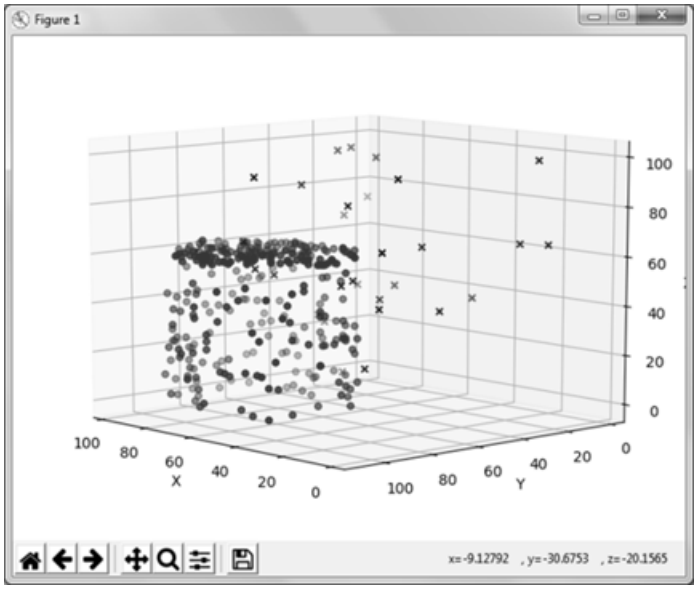


Fig. 3. Surface points cylindrical object and noise

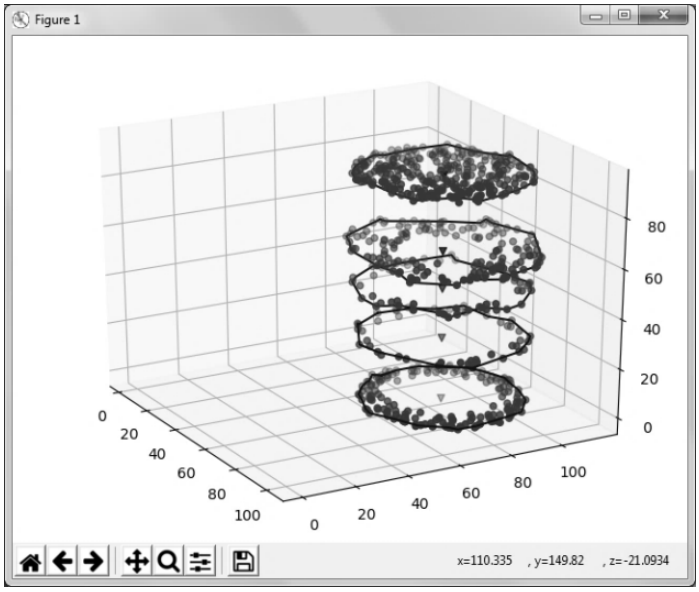


Fig. 4. Recognized vertical object of cylindrical form

Conclusion

The proposed approach is aimed at minimizing the time spent by typical objects, from which the vertical cylinder can be fully or partially distinguished.

In this paper, three methods were considered for detecting the centers and diameters of high single objects, after the convex hull was built for each layer.

The first method is simple, but it is suitable for ideal cylindrical objects, and for the rest it will produce results with a large error that do not correspond to the actual parameters of the object.

The second and third methods are fairly accurate, but lose on runtime to the first method.

REFERENCES

1. *Medvedev V.I., Raikova L.S.* Programs for processing of laser scanning data of a location // CAD and GIS roads. 2017. – N 2 (9). – P. 10–31.

2. *Borrmann D. et al.* The 3d Hough transform for plane detection in point clouds: A review and a new accumulator design // 3D Research. – 2011. – V. 2. – N 2. P. 1–13.

3. *Dalitz C., Schramke T., Jeltsch M.* Iterative Hough Transform for Line Detection in 3D Point Clouds // Image Processing On Line. – 2017. – V. 7. – P. 184–196.

4. *Kostyuk Y.L., Litovchenko M.I., Smirnova A.A.* Recognition of three-dimensional objects on the ground according to laser scanning data // The 17th International Conference named after A. F. Terpugov «Information Technologies and Mathematical Modeling» (ITMM-2018). – 2018. – P. 193–198.

5. *Merkin D.R.* Algebra of free and sliding vectors. M: Publishing house Fizma-tgiz, 1962. – 165 p.

CURRENT PROBLEMS IN UX-DESIGN SPHERE

Y.A. Khovansky, A.S. Yurovskaya, A.A Pichugina
TomskStateUniversity
yakovkhovansky@mail.ru

Introduction

The field of UX-design is an actively developing and changing field in science, it is also noted in scientific research [1]. Analysis of the scientific literature allows us to study the problems of UX design and their solutions, which are presented in the latest modern studies of 2016-2019. According to research, you can trace new directions in the study of UX-design.

1. Theoretical issues of UX research

Starting to talk about UX design, it is impossible not to consider human-computer interaction as well HCI initially focused on the work of a person with a computer interface, but now this concept has become much broader and includes all types of devices with displays (fig. 1).

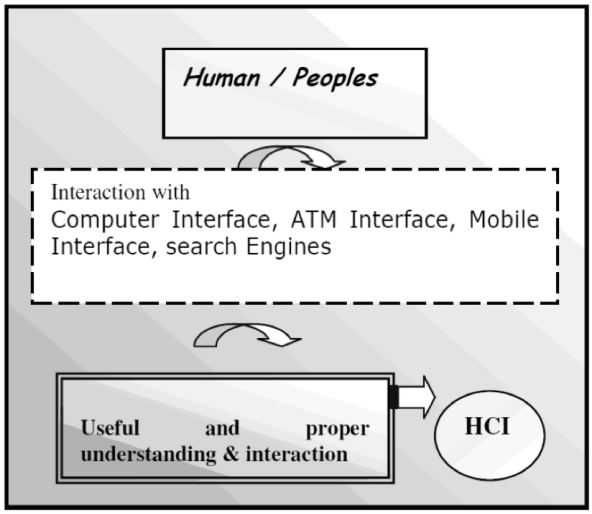


Fig. 1. Human interaction with computer and other devices

The main tasks of HCI are the creation of: more convenient interfaces, methodologies and interface design processes, an attractive interface, methods of using interfaces, a convenient information environment that allows you to process information faster [2]

Thus, UX-design can be considered a subsection of HCI, which focuses on creating efficient and user-friendly interfaces.Literature review by Tomasz Nedzielka [3] identified about 100 UX-methods of research. But despite this, the most common are the following: observa-

tion, interviews, prototyping, task analysis, cognitive passage, questionnaires, brief and longitudinal field studies as well as mixed methods.In addition, all of them can be divided into two global groups: individual and group laboratory studies.

Joy Robinson, Candice Lanius, and Ryan Weber [1] overlap in conclusions regarding the popularity of methods with previous work. However, they identified surveys and interviews as particularly common.It is worth noting that most of the UX research is based on methods imported from other disciplines, but there are a number of proprietary methods in the UX, such as contextual queries and eye tracking.

Despite the lack of similarity in the definition of the concept UX, due toits interdisciplinarity, Stefan Hellweger with co-authors [4], based on the literature, built a concept in which the elements of UX intersect with fields of application:

Element\ Field	Design	Information Technology	Psychology
Product	Novelty	Mobility, adaptability	Ethics
User	Ergonomic Issues	Productivity/reliability	Requirements (Needs), self-expression
Interaction	Brand	Engagement rate	Emotional fulfillment

On the basis of existing research methods UX, Hussein J., Khan V.A., Khur T. and other authors in their study [5] created a special research platform combining the collection of information using different methods and sensors (audio, video, sensors of fixing biometric indicators), as well as through direct user interaction (surveys, self-analysis, logging), fig. 2.

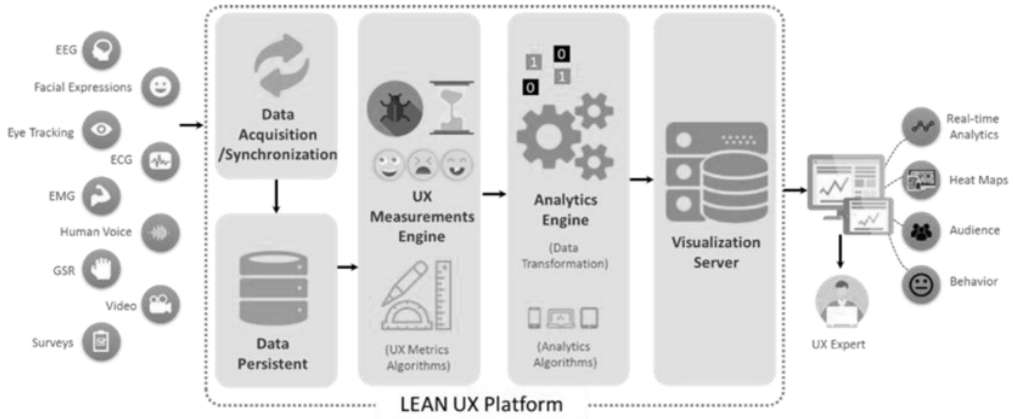


Fig. 2. Diagram of a research platform based on Lean UX

The considered research platform is based on Lean UX, which includes product development through continuous measurement of the so-called "learning loop» (construction – measurement – training) (fig. 3), and is able to receive and store data obtained from all sources, which are further analyzed. Data analysis allows us to conclude the user's perception of a product in order toits further improvement.

Fig. 3. The “tlearning loop” scheme

Marieke Van Camp and his co-authors also wrote about the effectiveness of methods of monitoring physical condition [6]. They analyze in detail the method of electroencephalography (EEG) to collect data on the bioelectric activity of individual zones, areas, brain, which occurs when the user interacts with the product, but the study also summarizes that for an adequate assessment of the user experience it is necessary not to limit the EEG results and combine both objective and subjective measurements.

Also, speaking about the UX study, one cannot fail to mention the cases when users showed a lot of negative emotions during interaction with the product, but afterwards they were convinced that their experience was positive. One of the reasons for this phenomenon can be called the gap in memory, defined as "the discrepancy between averaged experiences and the overall assessment of the experience, which is usually more intense than the averaged emotions." Hence the need to use a set of empirical research methods with methods that use machinery [7].

Lin Feng [8] in her work suggested using an approach that helps determine the quality of user experience, both in primary and in continuous use. Its essence is in the combination of the Attrakdiff questionnaire with the Hassenzal model [9] and the UX Curve method.

2. Practical applications of UX research

2.1. The use of UX for commercial purposes

In addition to the analysis of the UX sphere as a whole, as well as the development of new methods, researchers pay considerable attention to the application of methods in practice and evaluation of their effectiveness.

So, Rendani M. Kruger and colleagues conducted work in this direction [10]. They investigated the potential value of comparative UX for e-commerce. By comparative UX between companies, they understand a detailed study of the similarities or differences between user behavior when using interfaces of various organizations, with comparable goals. The main tools in the study were the case method and the eye tracker.

Florian Lachner et al. [11] in their study, tested the applicability of UX-based web analytics metrics as part of a cross-cultural experience of interacting with sites. The result of the work was the identification of significant differences in the perception of the same element of users from France, Germany and Italy, and hence the need to take this fact into account when developing products.

2.2. Application of UX techniques to help users with disabilities

UX techniques can be considered not only for commerce, but also on the social agenda. Scientists are paying a lot of attention to developing interfaces for people with various disabilities.

Papri Ghosh et al. [12] describe collective approaches to designing human-computer interaction for people with disabilities. At the moment, there are a lot of interfaces that embody these principles, such as an audio navigation system for people with visual impairments, a video navigation system for people with hearing problems and a voice command wheelchair for people with limited movement abilities. However, the complexity of designing such systems lies in maintaining a balance between the embodied functionality and the cost of implementation, since it is necessary to make devices accessible to audiences with different financial capabilities.

A particular difficulty lies in the fact that most user experience assessment (UX) tools require users to reflect and express their thoughts on their own (for example, thoughts out aloud, retrospective interviews, questionnaires). However, in the context of development for people with dementia, conditions such as aphasia and overall cognitive decline limit the applicability of these methods. To do this, Stefan Huber and co-authors [13] developed a separate application, the essence of which is to record human emotions and their subsequent video comparison.

Continue to work with people with dementia and Parkinson's disease and Naomi Kokubu with co-authors,[14] consider a cognitive function testing device sensitive to the determination of neuropsychological parameters in patients with these diseases. EXperience test-Trail Making User (UX-TBL) is a modified test ways of solving problems, which considers the achievement of the objectives on the screen of the tablet. UX-TY consists of 3 main components: 1) neurocognitive assessment, 2) cognitive training and 3) logging. In addition, UX-TMT has moderate Cronbach alpha values, indicating good internal consistency. UX-TBL is currently only available in Japanese language, but can be used for mass screening and monitoring the progression of dementia disorders.

Tiago Nogueira and Deller Ferreira [15] work with people with visual impairments. They study in their work the perception of web trends by visually impaired and blind users. They highlight the fact that websites are fundamental tools for communication, information and services, especially for people with visual impairments. For people with disabilities, the Internet is necessary for social and professional integration. However, millions of people around the world have difficulty interacting with websites and web applications. The reason is the low level of usability of these products.

3. The latest ways to improve human-computer interaction

In addition to researching existing UX design methods, scientists are also working to create new methods and devices. So, in April 2019, a patent was registered for a UX enhancement device based on the voice query history [16].

Another group of inventors moved away from the UX enhancement model for a large group of users and focused on creating a device to personalize user preferences[17]. B. Alenlyung, J. Lindblom, R. Anderason and T. Zimke [18], however, proceeded to consider the prospects for further development of user experience. In their opinion, socially interactive robots will be of increasing importance in everyday life for a large number of people. Recently, the number of socially interacting robots in the human environment has increased, and their level of participation in everyday activities has become increasingly diverse. For robots – as with all other types of interactive systems, products, and devices – a positive user experience is needed to achieve the intended goals.If the use of a robot entails a negative user experience, this can have negative consequences, such as unwillingness to use a particular robot or robots in general. Therefore, it is important for developers of robots to make serious efforts to

create robots that users consider positive. The study emphasizes the importance of researching not only a negative user experience, but also a positive one.

Conclusion

In the course of this work, the most recent studies in the field of UX-design were studied. Note that the literature highlights the problems of collecting and analyzing user experience, both theoretical and practical. Summarizing the above, we can say that today UX-design is an actual research direction not only in practical activities, but also in science. The user experience thus makes it possible to improve existing products and create new ones so that they meet the needs of users for efficiency, reliability, usability, and also evoke positive emotions. UX-research is a complex and multifaceted process both from a theoretical and practical point of view, therefore this topic is relevant for research in science.

REFERENCES

1. *Robinson J., Lanius C., Weber R.* The Past, Present, and Future of UX Empirical Research // Communication Design Quarterly Review. 2017. – V. 5. – Is. 3. – P. 10–23.

2. *Paul P.K., Kumar A., Ghosh M.* Human Computer Interaction and its Types: A Types // International Conference on Advancements in Computer Applications and Software Engineering (CASE 2012), At Chittorgarh, India. – 2012.

3. *Niedziółka T.* User experience research and development of knowledge in the field of human-computer interaction // Studia i Prace WNEIZ US. – 2018. – № 54/2. – P. 71–80.

4. *Hellweger S., Wang X., Abrahamsson P.* The Contemporary Understanding of User Experience in Practice. – URL: https://www.researchgate.net/publication/272828260_The_Contemporary_Understanding_of_User_Experience_in_Practice

5. *Hussain J., Khan.W.A., Hur T., Bilal H.S.M., Bang J., Hassan A.U., Afzal M., Lee S.A.* Multimodal Deep Log-Based User Experience (UX) Platform for UX Evaluation // Sensors (Basel). – 2018. – 18(5).

6. *Van Camp M., De Boeck M., Verwulgen S., De Bruyne G.* EEG Technology for UX Evaluation: A Multisensory Perspective // Advances in Neuroergonomics and Cognitive Engineering: proceedings of the AHFE 2018 International Conference on Neuroergonomics and Cognitive Engineering, July 2125, 2018, Loews Sapphire Falls Resort at Universal Studios, Orlando, Florida USA. – Springer, 2018. – P. 337–343.

7. *Nakamura W.T., De Oliveira E.H.T., Conte T.* Negative Emotions, Positive Experience: What Are We Doing Wrong When Evaluating the UX // Proceeding CHI EA '19 Extended Abstracts of the 2019 CHI Conference on Human Factors in Computing Systems.

8. *Feng L., Wei W.* An Empirical Study on User Experience Evaluation and Identification of Critical UX Issues // Sustainability/ – 2019. – 11.

9. AttrakDiff [Electronic resource]. – URL:http://attrakdiff.de/index-en.html

10. *Kruger R.M., Gelderblom H., Wynand B.* The value of comparative usability and UX evaluation for e-commerce organisations // CONF-IRM 2016 Proceedings. – 2016. – 9.

11. *Lachner F., Fincke F., Butz A.* UX Metrics: Deriving Country-Specific Usage Patterns of aWebsite Plug-in FromWeb Analytics // 2017.

12. *Ghosh P., Chingtham T.S., Ghose M.K., Dutta R.* A Review of Modern HCI Challenges for Betterment of Differently Abled People // January 2013.

13. *Huber S., Bejan A., Radzey B., Hurtienne J.* Proxemo or How to Evaluate User Experience for People with Dementia // May 2019.

14. *Kokubo N., Yokoi Y., Saitoh Y., Murata M., Maruo K., Takebayashi Y., Shinmei I., Yoshimoto S., Horikoshi M.* A new device-aided cognitive function test, User eXperience-Trail Making Test (UX-TMT), sensitively detects neuropsychological performance in patients with dementia and Parkinson's disease // BMC Psychiatry. – 2018, Jul 5. – 8 (1).

14. *Nogueira T., Ferreira D.* Systematic Review of Visually-Impaired and Blind User Experience of Web Trends // December 2018.

16. United States Patent Application Publication. Human machine interface system and method for improving user experience based on history of voice activity. Apr .4 , 2019

17. United States Patent Application Publication. Method and system for user experience event processing and analysis. Dec .13 , 2018.

18. *Alenljung B., Lindblom J., Anderasson R., Ziemke T.* User experience in social human-robot interaction // International Journal of Ambient ComputingandIntelligence (IJACI). – 2017. – 8(2). – P. 12–31.

АВТОРСКИЙ УКАЗАТЕЛЬ

Аду Кпангни Ивь Бэрэнжэ 225
Андреева Валентина Валерьевна 190
Андриевская Анна Андреевна 125,328
Балашова Ольга Михайловна 107
Беккерман Екатерина Николаевна 197
Бояркина Анна Александровна 229
Буторина Наталья Борисовна 202
Васинев Александр Владиславович 197
Веткина Анна Васильевна 40
Власкина Анастасия Сергеевна 232
Володенок Анастасия Викторовна 85
Гайдамака Юлия Васильевна 250
Галилейская Анастасия Александровна 295
Гендрина Ирина Юрьевна 118
Голышев Валерий Константинович 269
Горцев Александр Михайлович 40,58
Грекова Татьяна Ивановна 101
Гудкова Ирина Андреевна 225,300
Давыдова Анастасия Игоревна 118
Дмитренко Анатолий Григорьевич 107
Дымова Полина Игоревна 232
Емельянова Татьяна Вениаминовна 168
Завгородняя Мария Евгеньевна 67
Зенкова Жанна Николаевна 125,149,328
Зимин Михаил Павлович 236
Зимина Мария Михайловна 300
Ибрагимова Эллада Ибрагимовна 276
Иванов Владислав Дмитриевич 332
Идрисова Анастасия Ринатовна 281
Карим Пешанг Хасан Карим 284
Карлина Дарья Павловна 76
Кеба Анастасия Владимировна 23
Килин Никита Михайлович 310
Ким Константин Станиславович 79
Кожемякина Алиса Александровна 202
Копасов Дмитрий Владимирович 316
Костюк Юрий Леонидович 338,354
Кущазли Анна Ивановна 300
Лившиц Климентий Исаакович 90,94

Лизюра Ольга Дмитриевна 241
Лисовская Екатерина Юрьевна 246,257
Литовченко Марина Игоревна 338
Ложкин Алексей Сергеевич 183
Лопсанчап Буян Орланович 32
Максутова Дарья Гумаровна 321
Маркова Екатерина Викторовна 225
Матросова Анжела Юрьевна 183
Моисеев Александр Николаевич 236
Моисеева Светлана Петровна 229
Морозова Анна Сергеевна 321
Назаров Анатолий Андреевич 241
Нежельская Людмила Алексеевна 3,14,23,32,49
Никифоров Артём Дмитриевич 220
Охрименко Александра Сергеевна 131
Пальянова Юлия Евгеньевна 142
Пауль Светлана Владимировна 241
Перелевский Святослав Сергеевич 159
Першина Анна Александровна 49
Пичугина Анастасия Александровна 338,354,357
Платонова Анна Алексеевна 250
Поволоцкая Татьяна Александровна 90
Полуэктова Ксения Олеговна 101
Поляков Никита Алексеевич 232
Провкин Виктор Алексеевич 172,177
Просвиров Владислав Анатольевич 250
Пупков Андрей Викторович 154,165
Пчелинцев Евгений Анатольевич 159
Решетникова Галина Николаевна 85
Родионов Алексей Владимирович 342
Савич Валерия Николаевна 232
Самохина Светлана Ивановна 210,220
Серебренникова Кристина Андреевна 149
Сидорова Екатерина Филипповна 3
Скрипаченко Евгений Витальевич 345
Смагин Валерий Иванович 79
Смирнова Анастасия Александровна 354
Сущенко Сергей Петрович 284
Табольжина Юлия Евгеньевна 67
Тумашкина Диана Александровна 14

Тычинский Вячеслав Зиновьевич ... 190
Ульянова Екатерина Сергеевна ... 94
Фатеева Эвелина Сергеевна ... 183
Фомченко Татьяна Дмитриевна ... 250
Хованский Яков Александрович ... 357
Холкина Арина Владиславовна ... 73
Цыбульникова Нина Романовна ... 136
Чаллуб Акил Люайби ... 284
Чалых Егор Петрович ... 210
Чернышова Елизавета Николаевна ... 246
Чупрасова Людмила Витальевна ... 94
Шапиро Людмила Дмитриевна ... 310
Шерстобитова Анна Олеговна ... 168
Шкабара Анастасия Игоревна ... 293
Шкуркин Алексей Сергеевич ... 332,342,345
Шманкеева Анастасия Вячеславовна ... 58
Шуленин Виктор Петрович ... 154
Юровская Алёна Сергеевна ... 357
Якимук Алексей Юрьевич ... 215
Ярцева Ирина Сергеевна ... 250

СОДЕРЖАНИЕ

I. Управление и обработка информации 3

Нежельская Л.А., Сидорова Е.Ф. Аппроксимация реального трафика одним из дважды стохастических потоков событий 3

Нежельская Л.А., Тумашкина Д.А. Оценивание методом моментов параметров плотности вероятности длительности интервала между событиями рекуррентного полусинхронного потока второго порядка в особом случае 14

Кеба А.В., Нежельская Л.А. Апостериорные вероятности состояний обобщённого MAP-потока событий с произвольным числом состояний 23

Лопсанчап Б.О., Нежельская Л.А. Уравнение моментов для оценки длительности мёртвого времени в MAP-потоке событий 32

Горцев А.М., Веткина А.В. Имитационное моделирование полусинхронного дважды стохастического потока событий с непродлевающимся случайным мёртвым временем 40

Нежельская Л.А., Першина А.А. Статистические эксперименты на имитационной модели обобщённого асинхронного потока событий при непродлевающемся случайном мёртвом времени 49

Горцев А.М., Шманкеева А.В. Статистические эксперементы на имитационной модели асинхронного потока событий при непродлевающемся случайном мертвом времени 58

Завгородняя М.Е., Табольжина Ю.Е. Оценивание параметров простейшего потока событий и мертвого времени, распределенного по закону Рэлея, методом моментов 67

Холкина А.В. Моделирование случайных величин методом обратных функций 73

Карлина Д.П. Имитационное моделирование случайных величин методом суперпозиции 76

Ким К.С., Смагин В.И. Фильтрация в дискретных системах со случайными скачкообразными параметрами и неизвестным входом 79

Решетникова Г.Н., Володенок А.В. Синтез адаптивных систем управления при неполной информации 85

Лившиц К.И., Поволоцкая Т.А. Модель управления запасами однородной продукции с релейным управлением темпом производства и MMP-потоком моментов продаж 90

Лившиц К.И., Чупрасова Л.В., Ульянова Е.С. Розничная продажа многономенклатурной партии продукции с ограниченным сроком годности при MMP-потоке моментов продаж 94

Полуэктова К.О., Грекова Т.И. Оптимальное распределение ресурсов в двухсекторной модели экономики с государственным регулированием 101

Дмитренко А.Г., Балашова О.М. Программное обеспечение для решения задачи электромагнитного рассеяния на тонких параллельных идеально проводящем и диэлектрическом цилиндрах 107

Давыдова А.И., Гендрина И.Ю. Применение гармонического анализа для изучения характеристик систем видения через атмосферу 118

II. Системный анализ и математическое моделирование 125

Андриевская А.А., Зенкова Ж.Н. Подход Липовецкого к модификации метода ценообразования Штоцеля 125

Охрименко А.С. Разработка программной надстройки для ABC-XYZ-анализа по случайно цензурированным данным в среде MS Excel ... 131

Цыбульникова Н.Р. Расчет цены спроса на программный продукт методом PSM для зависимых наблюдений ... 136

Пальянова Ю.Е. Прогнозирование на основе моделей растущих систем ... 142

Зенкова Ж.Н., Серебренникова К.А. Расчет чистой нетто-премии по полным и цензурированным данным о страховых выплатах по комплексному автомобильному страхованию, кроме ответственности (КАСКО) ... 149

Шуленин В.П., Пупков А.В. Непараметрический БКР-критерий независимости ... 154

Перелевский С.С., Пчелинцев Е.А. Асимптотически эффективное оценивание в непараметрической гетероскедастичной регрессии ... 159

Пупков А.В. Обнаружение разладки в гауссовском процессе AR(1) с неизвестным параметром ... 165

Шерстобитова А.О., Емельянова Т.В. Построение одноэтапной последовательной процедуры оценивания параметров авторегрессионной модели с непрерывным временем ... 168

III. МАТЕМАТИЧЕСКОЕ И ПРОГРАММНОЕ ОБЕСПЕЧЕНИЕ ВЫЧИСЛИТЕЛЬНЫХ МАШИН И КОМПЬЮТЕРНЫХ СЕТЕЙ ... 172

Провкин В.А. Маскирование константных неисправностей логических схем внешними вентильными подсхемами ... 172

Провкин В.А. Синтез частично программируемых схем, ориентированный на маскирование трудно обнаружимых неисправностей ... 177

Ложкин А.С., Матросова А.Ю., Фатеева Э.С. Построение частичных функций для подсхем ... 183

Тычинский В.З., Андреева В.В. О свойствах ROBDD-графов, представляющих тестовые пары для робастно тестируемых неисправностей задержек путей ... 190

Васинев А.В., Беккерман Е.Н. Применение расширенного автомата для обучения алгоритмам с вариативной траекторией решения ... 197

Кожемякина А.А., Буторина Н.Б. Система сжатия информации с выбором оптимального алгоритма ... 202

Чалых Е.П., Самохина С.И. Виртуальная машина EmeraVM ... 210

Якимук А.Ю. Влияние вибрато на качество распознавания нот в вокальном исполнении ... 215

Никифоров А.Д., Самохина С.И. Создание API-клиентов с помощью код-генераторов ... 220

IV. ТЕОРИЯ МАССОВОГО ОБСЛУЖИВАНИЯ И ТЕЛЕТРАФИКА ... 225

Аду К.И., Маркова Е.В., Гудкова И.А. Методы расчета показателей эффективности СМО, функционирующей в случайной среде ... 225

Бояркина А.А., Моисеева С.П. Исследование бесконечнолинейных СМО с групповым обслуживанием ... 229

Дымова П.И., Савич В.Н., Поляков Н.А., Власкина А.С. Вероятностная модель нарезки ресурсов беспроводной сети между двумя виртуальными операторами ... 232

Зимин М.П., Моисеев А.Н. Применение отрицательного биномиального распределения для аппроксимации стационарного распределения числа заявок в системе HIGI|GI|∞ ... 236

Лизюра О.Д., Назаров А.А., Пауль С.В. Асимптотический анализ RQ-системы MMPP|M|1 с n типами вызываемых заявок в условии предельно редких изменений состояний входящего потока 241

Чернышова Е.Н., Лисовская Е.Ю. Численный анализ суммарного объема занятого ресурса на блоках системы с параллельным обслуживанием и входящим MMPP-потоком 246

Просвиров В.А., Ярцева И.С., Платонова А.А., Фомченко Т.Д., Гайдамака Ю.В. Модель обслуживания трафика мультивещания точкой доступа миллиметрового диапазона с учетом фонового трафика 250

Галилейская А.А., Лисовская Е.Ю. Стационарное асимптотическое распределение вероятностей суммарного объёма занятых ресурсов в бесконечнолинейной ресурсной СМО с входящим MMPP-потоком 257

V. КОМПЬЮТЕРНЫЕ НАУКИ, ИНТЕЛЛЕКТУАЛЬНЫЙ АНАЛИЗ И ВИЗУАЛИЗАЦИЯ ДАННЫХ .. 269

Голышев В.К. Методы нечёткой логики в анализе формальных понятий 269

Ибрагимова Э.И. Исследование 2-сбалансированных знаковых графов 276

Идрисова А.Р. Структура независимости случайных событий и матроиды 281

Карим П.Х., Сущенко С.П., Чаллуб А.Л. Численный анализ быстродействия транспортного протокола в режиме группового отказа с прямой коррекцией ошибок . 284

Шкабара А.И. Исследование методов кластеризации для анализа задач в больших вычислительных сетях 293

Зимина М.М., Кущазли А.И., Гудкова И.А. Разработка и оптимизация модели бизнес-процесса поддержки системы электронного документооборота 300

Килин Н.М., Шапиро Л.Д. Разработка программного обеспечения для оценки эффективности инвестирования 310

Копасов Д.В. Проектирование и разработка гибридного мобильного приложения для научной школы 316

Максутова Д.Г., Морозова А.С. Разработка информационной системы для расчета калькуляции себестоимости продукции 321

VI. COMPUTER SCIENCE 328

Andrievskaya A.A., Zenkova Z.N. The Use of Stoetzel Method in Pricing of New Bio-Energy Drink 328

Ivanov V., Shkurkin A. Online Portfolio Service for Music Content Creators 332

Kostyuk Y.L., Litovchenko M.I., Pichugina A.A. Detection of Underlying Surface, Horizontal and Vertical Edges of Three-Dimensional Objects Based on Laser Scanning Data 338

Rodionov A., Shkurkin A. Multiplayer Blackjack 342

Skripachenko E., Shkurkin A. Development of a Web Application to Determine the Hazard Category of Food Additives 345

Kostyuk Y.L., Smirnova A.A., Pichugina A.A. The Recognition of Single High Objects on Terrain According to the Laser Scanning Data 354

Khovansky Y.A., Yurovskaya A.S., Pichugina A.A. Current Problems in Ux-Design Sphere 357

АВТОРСКИЙ УКАЗАТЕЛЬ 362

Научное издание

МАТЕРИАЛЫ
VII Международной молодежной научной конференции «МАТЕМАТИЧЕСКОЕ И ПРОГРАММНОЕ ОБЕСПЕЧЕНИЕ ИНФОРМАЦИОННЫХ, ТЕХНИЧЕСКИХ И ЭКОНОМИЧЕСКИХ СИСТЕМ»

Томск, 23–25 мая 2019 г.

Под общей редакцией
кандидата технических наук И.С. Шмырина

Издание подготовлено в авторской редакции

Подписано в печать 15.10.2019 г. Формат 70×108 1/16
Печ. л. 23; усл. печ. л. 32,2.
Тираж 500 экз. Заказ № .

Отпечатано на оборудовании
Издательского Дома
Томского государственного университета
634050, г. Томск, пр. Ленина, 36
Тел. 8+(382-2) 52-98-49
Сайт: http://publish.tsu.ru; e-mail: rio.tsu@mail.ru