МИНИСТЕРСТВО ОБРАЗОВАНИЯ И НАУКИ РОССИЙСКОЙ ФЕДЕРАЦИИ

ВЕСТНИК ТОМСКОГО ГОСУДАРСТВЕННОГО УНИВЕРСИТЕТА

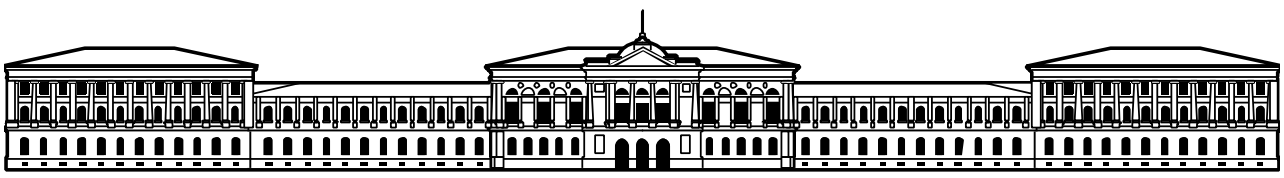
УПРАВЛЕНИЕ, ВЫЧИСЛИТЕЛЬНАЯ ТЕХНИКА И ИНФОРМАТИКА

TOMSK STATE UNIVERSITY JOURNAL OF CONTROL AND COMPUTER SCIENCE

Научный журнал

2015 **№ 4 (33)**

Свидетельство о регистрации: ПИ № ФС 77-29497 от 27 сентября 2007 г.



JORNAL INFO

Tomsk State University Journal of Control and Computer Science is an independent peer-reviewed research journal that welcomes submissions from across the world.
Tomsk State University Journal of Control and Computer Science is issued four times per year, and can be subscribed to in the Russian Press Joint Catalogue (Subscription Index 44031).
The publication in the journal is free of charge and may be in Russian or in English.

The topics of the journal are the following:
- *control of dynamical systems,*
- *mathematical modeling,*
- *data processing,*
- *informatics and programming,*
- *discrete function and automation,*
- *designing and diagnostics of computer systems.*

Rules of registration articles are given in a site:
http://journals.tsu.ru/informatics/

ISSN 2311-2085 (Online), ISSN 1998-8605 (Print).

О ЖУРНАЛЕ

Журнал «Вестник Томского государственного университета. Управление, вычислительная техника и информатика» выходит ежеквартально и распространяется по подписке, его подписной индекс 44031 в объединённом каталоге «Пресса России». Статьи публикуются на русском и английском языках.

Тематика публикаций журнала:
- *управление динамическими системами,*
- *математическое моделирование,*
- *обработка информации,*
- *информатика и программирование,*
- *дискретные функции и автоматы,*
- *проектирование и диагностика вычислительных систем.*

Журнал «Вестник Томского государственного университета. Управление, вычислительная техника и информатика» включен в Перечень изданий для публикации основных результатов кандидатских и докторских диссертаций, а также входит в систему Российского Индекса Научного Цитирования (РИНЦ). Правила оформления статей приведены на сайте:
http://journals.tsu.ru/informatics/

ВЕСТНИК ТОМСКОГО ГОСУДАРСТВЕННОГО УНИВЕРСИТЕТА
2015 Управление, вычислительная техника и информатика № 4 (33)

СОДЕРЖАНИЕ

ОБРАБОТКА ИНФОРМАЦИИ

ИНФОРМАТИКА И ПРОГРАММИРОВАНИЕ

ПРОЕКТИРОВАНИЕ И ДИАГНОСТИКА ВЫЧИСЛИТЕЛЬНЫХ СИСТЕМ

CONTENTS

DATA PROCESSING

INFORMATICS AND PROGRAMMING

DESIGNING AND DIAGNOSTICS OF COMPUTER SYSTEMS

ОБРАБОТКА ИНФОРМАЦИИ

УДК 519.237.8: 81’322.2
DOI: 10.17223/19988605/33/1

M.F. Ashurov, V.V. Poddubny

TEXT CLASSIFICATION STREAM-BASED *R*-MEASURE APPROACH USING FREQUENCY OF SUBSTRING REPETITION

The publication is prepared within the framework of the scientific project No. 14-14-70010 supported by the Russian Humanitarian Scientific Fund. This work is performed under the state order No. 1.511.2014/K of the Ministry of Education and Science of the Russian Federation.

Stream-based approach of *R*-measure using frequency of substring repetition in text classification is offered. Comparative quality analysis of classificators based on the truncated *R*-measure using frequencies of test text substring repetition and without one is performed on a text set of Russian fiction of the 19th century and the 90th of 20th century. An accuracy of text classification is estimated by Van Rijsbergen’s effectiveness measure known as *F*-measure. The fact that in case of genre mixing free into author’s text classes accounting frequency of test text substring repetition in supertexts increases the classification accuracy is confirmed.

Keywords: stream-based classification approach; *R*-measure; frequency of substring repetition; classification accuracy; *F*-measure.

Issues of automatic text classification arise in the creation and updating of search systems, scientific research linked with recognition style features of fiction or published texts, author text matching etc. The formal description of the text classification issue is described, as example, in [1]. As mentioned in [1] all text classification approaches can be divided on two groups: feature-based and stream-based. Feature-based approaches do not deal with text directly, but these use text-value representation as an array of feature values. It is necessary to determine a sufficient feature set by which the classificatory calibrated before using any feature-based approach of text classification. In case of feature set selection is poor; the classification result becomes extremely unsatisfactory. Stream-based approaches unlike feature-based do not require selection of any features for text classification and deal with text directly. Text X is considered as a sequence (stream) $x_1 x_2 \ldots x_n$ of n elements from an alphabet A.

Among the stream-based approaches are two main groups described in [1–4]. The first group is based on substring repetition counting (*R*-, *C*- and other measures). The second group uses data compression (for example, the off-the-shelf algorithm using PPM (Prediction by Partial Matching) compression method).

Khmelev [5] did a comparison of the methods based on substring repetition counting with the methods based on data compression. The comparison was held on the news articles (journalistic articles) in which an author style is involved much and a genre specifics influence is decreased. To estimate an accuracy of text classification better the additional comparison should be done on different types of texts, particularly, on fictions.

We offer a modification of *R*-measure which can use frequencies of substring repetition. This article describes the comparative quality analysis of the classification approaches based on the truncated *R*-measure and its modification in case of Russian fiction texts of different ages.

1. Stream-based approaches and truncated *R*-measure

Approaches proposed by Khmelev and Teahan in [6] are based on both of the above groups, but there is the approach which can be especially highlighted. This approach named *R*-measure (from repetition) counts all substrings of a test text that are present in a class supertext. The supertext is formed by concatenating all texts of the particular class. So *R*-measure has the following equation:

$$R(X \mid S) = \frac{\sum_{k=1}^{n} \sum_{i=1}^{n-k} c\left(x_i \ldots x_{i+k} \mid S\right)}{\frac{1}{2} n(n+1)}, \quad (1)$$

where X is a test text; S – a supertext of current class; n – length of the test text; k – length of pattern substring used in string search and $c\left(x_i \ldots x_{i+k} \mid S\right)$ – indicator function, indicating presence of a substring into the supertext:

$$c\left(x_i \ldots x_{i+k} \mid S\right) = \begin{cases} 1, x_i \ldots x_{i+k} \in S, \\ 0, x_i \ldots x_{i+k} \notin S. \end{cases} \quad (2)$$

As the volumes of fiction, texts are so huge R-measure calculation should be done quickly. In that case, the suffix arrays [1, 3, 4] are used in the R-measure calculation to search a substring quickly in the huge text.

The indicator function based on suffix array approach can be implemented in an algorithm using dichotomous search in the suffix array. So the indicator function should return 1 value if a pattern substring is matched with a suffix in the array at first time (other matching is not used here), or return 0 value, if no matching with the suffixes.

Also note that R-measure equation (1), in fact, has the number of used substrings in the denominator. This number is used to normalize the value of the proximity between the texts so that it should be in the interval [0,1], where 1 value means that a class supertext contains full test text (the maximum similarity of a test text with a class supertext), and 0 value means no matching between the test text and the supertext.

In the first place, the main disadvantage of R-measure is the fact that the substrings with length in 1 or 3 symbols and the substring with length over 200 symbols are absolutely unsuitable, because these substrings cannot be used to recognize specific characteristics for each class. The second disadvantage of R-measure is the issue that the length n of test fiction texts is around 150 thousand symbols on average so R-measure calculation under that condition of test text length requires a lot of computing resources and time on the typical personal computers in the classification using real data. Even if using the best on-time performance algorithms of substring search and data structures, R-measure calculation requires about 6 days for each class at the length of test text around 200 thousand symbols. A lot of time required in the R-measure classification is completely unsuitable in the real world. To solve these issues the modification of R-measure, named truncated R-measure, which uses only particular substring lengths from the full range of available substrings has been proposed.

Truncated R-measure similarity described in [1] counts all repetitions of substrings of test text length n in the supertext. These substrings have lengths from k_1 to k_2 unlike classical R-measure equation (1) where $k_1 = 1$, $k_2 = n$:

$$R(X \mid S) = r(X \mid S) / N, \quad (3)$$

where X is a test text; S – a supertext of particular class; N – a number of used substrings. The function to calculate a measure of similarity is described as follows:

$$r\left(X \mid S\right) = \sum_{k=k_1}^{k_2} \sum_{i=0}^{n-k} c\left(x_i \ldots x_{i+k} \mid S\right). \quad (4)$$

Normalizing of the measure is based on the number of used substrings N, which is calculated by following equation:

$$N = \left(2(n+1) - (k_1 + k_2)\right)\left(k_2 - k_1 + 1\right) / 2. \quad (5)$$

The issue of the k_1 and k_2 parameters selection used in truncated R-measure calculation has been solved by natural language features. The minimum length k_1 is equal to 10 symbols because shorter substrings can match with many words in Russian language, which are common for all authors so authors specifics cannot be pick out. The maximum length k_2 is equal to 45 symbols that allows to process several words from the test text. Using the greater length of substrings seems not useful as different authors cannot repeat such long phrases many times or it could be unlikely. So a time calculation of the truncated R-measure is decreased compared with the classical R-measure because of the reduction of substrings amount. For example, the time to calculate the truncated R-measure on a test text which length is equal to 150 thousand symbols for each class is reduced to the acceptable level about one or two minutes.

2. Modification of *R*-measure using frequencies of substring repetition

Previous comparison of approaches based on *R*-measure and PPMD (PPMd – Prediction by Partial Matching (dynamic) data compression algorithm without loss) compression described in [7, 8] shows that the accuracy of classification by *R*-measure is less than accuracy by the PPMD approach in some case. These cases are explained by the fact that *R*-measure approach does not use frequencies of substring repetition unlike the PPMD compression approach. To estimate feasible performance of using frequencies in classification the approach based on *R*-measure modification that can use frequencies of substring repetition is offered. This *R*-measure modification is named *RF*-measure (repetition frequency). The *RF*-measure detailed description and its comparison with *R*-measure are described in this article.

RF-measure is based on the same arrangement of truncation used in equations (3) and (4):

$$RF(X \mid S) = rf(X \mid S) / N, \tag{6}$$

$$rf(X \mid S) = \sum_{k=k_1}^{k_2} \sum_{i=0}^{n-k} rf(x_i \ldots x_{i+k} \mid S). \tag{7}$$

In that case, indicator function $c(x_i \ldots x_{i+k} \mid S)$ used in *R*-measure equation (3) is removed and the function, which determines the degree of similarity of substring frequencies between a test text and a supertext, is used in the similarity measure:

$$rf(x_i \ldots x_{i+k} \mid S) = \begin{cases} rfc(x_i \ldots x_{i+k}), x_i \ldots x_{i+k} \in S, \\ 0, x_i \ldots x_{i+k} \notin S. \end{cases} \tag{8}$$

During the selection of similarity measure, it was decided to use the relative frequencies of substring repetition of the test text and the supertext in the measure calculation. The main condition of the measure selection is that calculated value should satisfy the normalization principles. It means that the similarity measure value should be in the interval [0,1], where 1 value means that compared frequencies are the same. This condition can be satisfied easily if similarity measure is based on the ratio of the minimal frequency to maximum one. Accordingly, the similarity measure using the ratio of frequencies of substrings repetition of test texts and supertexts is following:

$$rfc(x_i \ldots x_{i+k}) = \frac{\min(f_X(x_i \ldots x_{i+k}), f_S(x_i \ldots x_{i+k}))}{\max(f_X(x_i \ldots x_{i+k}), f_S(x_i \ldots x_{i+k}))}.$$

Thus, any difference in the relative frequencies results in a reduction of the similarity measure, whereby a test text belonging to the class is reduced too.

The relative frequencies of substrings repetition for test text *X* and supertext *S* is as follows:

$$f_X(x_i \ldots x_{i+k}) = \frac{cc_X(x_i \ldots x_{i+k})}{N_X} \text{ and } f_S(x_i \ldots x_{i+k}) = \frac{cc_S(x_i \ldots x_{i+k})}{N_S},$$

where a function $cc(x_i \ldots x_{i+k})$ unlike the indicator function (2) returns the amount of pattern substring matching in the test text and supertext (relevant marks *X* and *S* is in the subscripts), and *N* is the lengths of texts *X* and *S* respectively.

The function $cc(x_i \ldots x_{i+k})$ can be implemented in an algorithm using the dichotomous substring search in a suffix array. This algorithm after the completion of substring search process and the first success of full pattern matching selects next suffixes in series that satisfy the pattern matches and counts the total number of that matches. To decrease time of counting matches the special array containing the lengths of a common prefix of the next suffix for each particular suffix can be used. In the algorithm the procedure of full matching substring pattern and suffix can be skipped in case a length of the next suffix prefix less than a length of the substring pattern. Using the common prefix array calculation time of counting all substring matches for substring lengths more than 10 symbols could be decreased about 5–10%.

3. Accuracy analysis of classification

In this article the analysis of classification accuracy of *R*-measure and *RF*-measure have been done for two text samples. The first text sample consists of 126 prose fiction texts of nine authors of the 19th century rep-

resenting the classes. This text sample is described in [9, 10]. More than 90 texts have been used as a training sample to build corresponding supertexts. The second text sample used in [8] work is the prose fiction texts of the 90th of 20th century. This text sample contains 138 texts by 21 authors. The volume of the supertexts is the same for each class necessarily and text lengths of test samples are equal to 100 thousand symbols.

To test the classification accuracy software module that allows classify a text by *R*-measure and *RF*-measure has been designed and implemented. The issue of parameters k_1 and k_2 selection has been solved by natural language features as described above.

Classification accuracy for each class is estimated by Van Rijsbergen's measure [1, 11] known as *F*-measure. *F*-measure is harmonic mean of precision *p* (the number of correct positive results divided by the number of all positive results) and recall *r* (the number of correct positive results divided by the number of positive results that should have been returned):

$$F = 2\frac{r \times p}{r + p}.$$

In general, the rating of the classification accuracy is calculated as an arithmetical mean of *F*-measures of all classes, named *F*-macro, or as *F*-measure is calculated on average of precisions and recalls of all classes, named *F*-micro.

4. Results

As for the first text sample represented by 9 classes the result characteristics have been calculated for truncated *R*-measure and *RF*-measure, shown in table 1.

Table 1

Precision, recall and *F*-measure for authors of the 19th century

Author	Precision		Recall		*F*-measure	
	***R*-measure**	***RF*-measure**	***R*-measure**	***RF*-measure**	***R*-measure**	***RF*-measure**
Chehov	0,60	0,56	0,75	0,75	0,67	0,64
Dostoevskii	1,00	0,56	0,42	0,42	0,59	0,48
Gogol	0,92	1,00	1,00	1,00	0,96	1,00
Goncharov	0,67	0,63	1,00	1,00	0,80	0,77
Kuprin	1,00	0,50	0,50	0,42	0,67	0,45
Leskov	0,71	1,00	0,83	0,75	0,77	0,86
Saltikov-Shedrin	0,40	0,40	0,17	0,17	0,24	0,24
Tolstoi	1,00	1,00	0,92	0,92	0,96	0,96
Turgeniev	0,57	0,71	1,00	1,00	0,73	0,83

For clarity, the total ratings of classification accuracy on the first text sample are shown in table 2.

Table 2

Values of *F*-macro and *F*-micro for each approach

Approach	*F*-measure (macro)	*F*-measure (micro)
R-measure	0,71	0,75
RF-measure	0,69	0,71

For the first text sample the *F*-measure of the classificator based on *R*-measure is 2-4% better than the *F*-measure of the classificator based on *RF*-measure. Further research has illustrated the fact that stylistic characteristics of the test texts and supertexts of some classes are different visibly. In case using *RF*-measure that difference between texts can significantly affects the numerical value of similarity measure and results of text classification.

To minimize the effect of the above issue, depending on the quality of training and test samples, it has been decided to make these samples randomly from all texts of the class. The process of sample generation must be satisfy two conditions – test texts should not be included in the training sample, which is used to build a supertext, and vice versa, and the volumes of all supertext should be the same for each class. After repeated classifying all values of *F*-measure calculated for each class on the randomly generated samples are average for

each class. Average values of the class *F*-measure for the truncated *R*-measure и *RF*-measure, and total *F*-macro of the classificators are presented in table 3.

Table 3

Average *F*-measure for authors of the 19th century on randomly generated samples

Author	*F*-measure	
	***R*-measure**	***RF*-measure**
Chehov	0,95	0,90
Dostoevskii	0,41	0,57
Gogol	0,88	0,92
Goncharov	0,91	0,89
Kuprin	0,89	0,69
Leskov	0,89	0,90
Saltikov-Shedrin	0,61	0,73
Tolstoi	0,94	0,94
Turgeniev	0,73	0,85
Total F-mesuare (macro)	0,80	0,82

As illustrated by results in table 3 total (average for the classes) the *F*-measure value of *RF*-measure is higher than the *F*-measure value of *R*-measure. Note that only two classes have *F*-measure results using *RF*-measure worse than results using *R*-measure on 2–5%. It means using *RF*-measure improves the classification accuracy in general. This fact also can be confirmed by the classification results on the second text sample. The text sample is specific because it contains modern text of little-known authors, texts are in most cases represented by one genre and the used set of words is not wide as for prose authors of the 19th century.

As for the second text sample based on 21 classes the results for truncated *R*-measure and *RF*-measure are shown in table 4.

Table 4

Precision, recall and *F*-measure for authors of the 20th century

Author	Precision		Recall		*F*-measure	
	***R*-measure**	***RF*-measure**	***R*-measure**	***RF*-measure**	***R*-measure**	***RF*-measure**
Agafonov	0,88	0,88	1,00	1,00	0,93	0,93
Aristov	1,00	1,00	1,00	1,00	1,00	1,00
Azarov	0,96	1,00	1,00	1,00	0,98	1,00
Baganov	0,79	0,84	1,00	1,00	0,88	0,91
Belkin	1,00	1,00	1,00	1,00	1,00	1,00
BelobrovPopov	1,00	1,00	1,00	1,00	1,00	1,00
Belomlinskaya	1,00	1,00	1,00	0,80	1,00	0,89
Belov	1,00	1,00	0,78	0,78	0,88	0,88
Bonch	1,00	1,00	0,27	0,27	0,43	0,43
Bronin	1,00	1,00	1,00	1,00	1,00	1,00
Burmistrov	0,50	0,50	1,00	1,00	0,67	0,67
Galkin	1,00	1,00	0,29	0,29	0,45	0,45
Gergenreder	1,00	1,00	1,00	1,00	1,00	1,00
Glushkin	1,00	1,00	1,00	1,00	1,00	1,00
Svetlana	1,00	1,00	1,00	1,00	1,00	1,00
Velboi	0,10	0,26	0,20	0,90	0,13	0,40
Vershovskii	1,00	1,00	0,19	0,19	0,32	0,32
Veter	0,57	0,80	1,00	1,00	0,73	0,89
Vitkovskii	1,00	1,00	0,33	0,33	0,50	0,50
Voronov	1,00	1,00	1,00	1,00	1,00	1,00
Vulf	1,00	1,00	1,00	1,00	1,00	1,00

For clarity, the total ratings of classification accuracy on the second text sample are shown in table 5.

Table 5

Values of *F*-macro and *F*-micro for each approach

Approach	*F*-measure (macro)	*F*-measure (micro)
R-measure	0,80	0,85
RF-measure	0,82	0,88

The classificator based on *RF*-measure is a little better than one based on *R*-measure on the text sample. And only one class has the *F*-measure value of *RF*-measure less than one of *R*-measure.

Moreover, note the fact that there are classes, classification accuracy which is extremely low. Additional researches show that this aspect occurs because the similarity measure between the supertext of that classes and the supetext of another classes is more than 20%. The aspect is illustrated in figure 1.

Identify author's stylistic features of such text requires using additional instruments. It can be an analyzer, which takes class specific substrings from the text, or a filter, which excludes substrings that are absolutely common for the classes (for example, it can be a common phrase or idioms of Russian language or a series of function words without semantic). These instruments can become a dynamic dictionary, which can be extended by adding new texts in the class supertext. Moreover, the proposal instruments are heuristic tools based on the language features and its performance is questionable in general if a stream contains other entities instead of text characters (stream codes of the human genome for example).

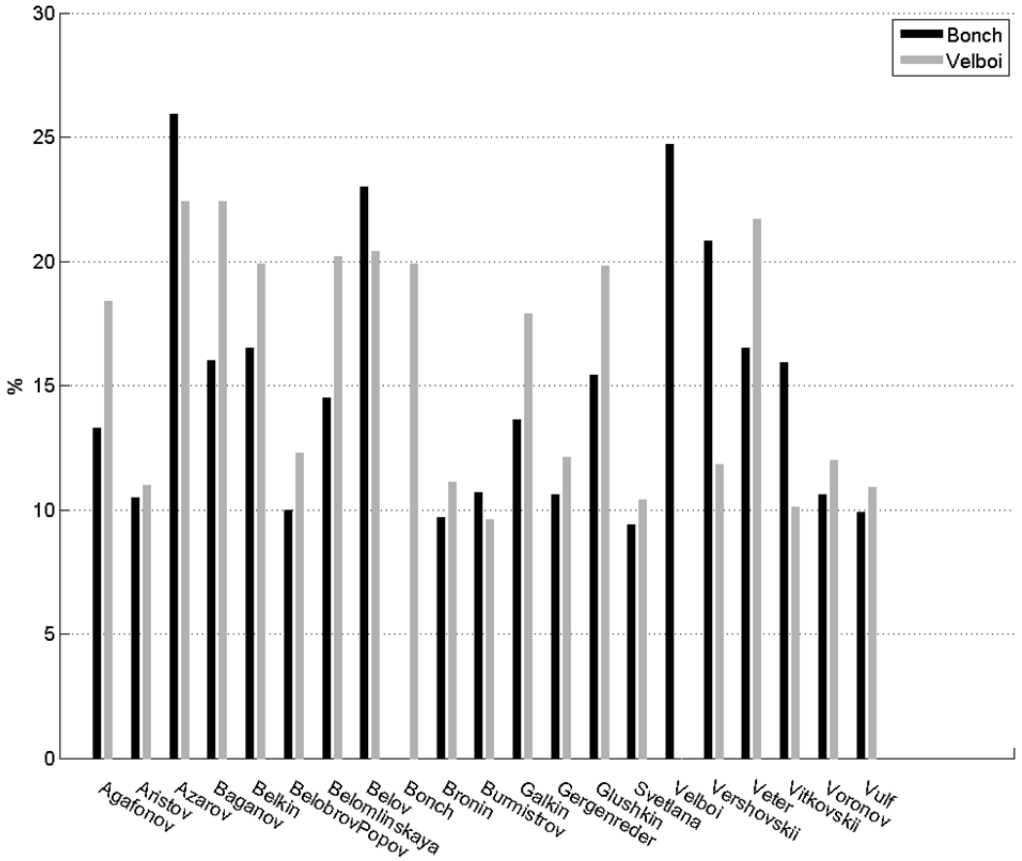


Fig. 1. Similarity measure of Bonch and Velboi supetexts by other classes

As for the second text sample, note that *RF*-measure has a better classification accuracy for that class types allowing highlight the stylistic features of author style well. So the accuracy of the classification based on *RF*-measure is better on text samples, class supetext of which can be very similar. In that case, the precisions of two classes have increased more than 50%.

Conclusions

Classification accuracy of both classificators on considered text data is more than 70%. Furthermore, the approach proposed in this article has a prospect of further improving. Individual features of the constructed training and test samples can explain the issue that *RF*-measure has the low values or it is less than *R*-measure for some classes. The impact of this issue can be decreased by frequentative classifications on the training and test text samples created randomly from all texts of the class. In the future, we see an opportunity to improve a classification accuracy of *RF*-measure by:

1) using a threshold filtering;

2) modification of counting relative frequencies of a supertext;

3) modification of the function using ratio of the minimum of relative frequencies to the maximum to the weighted average.

The first aspect is connected with the issue that natural language texts have some substrings that do not contains the semantic content (author's thoughts or natural logic), but the number of this substring repetition can reach 200 or 300 repetitions in the supertext. It affects relative frequencies vastly so these substrings, in the simplest case, can be skipped.

The second aspect occurs when a class supertext contains many different texts, so a length of the supertext is many times greater than a length of the test text. The current approach uses the relative frequency calculation as a ratio of the number of pattern substring repetitions to length of the supertext. So relative frequency of the supertext is less than the relative frequency of the test text if the numbers of repetitions are the same. Moreover, as for natural languages, if the supertext consists of different texts then the relative frequency of the supertext cannot recognize clearly author style used in particular text included in the supertext. Therefore, it requires finding a function which calculates a relative frequency of the supertext removing the effects of absolute value of the supertext length.

The third aspect is based on the above issue that the difference between relative frequencies of the supertext and test text results in little value of the frequencies ratio. This means that, in practice, calculated value of the ratio is less than 1 if there is a little absolute difference between the relative frequencies. The little difference of frequencies is a normal in fiction texts. Therefore, a function, which returns values around 1 in little difference of relative frequencies, should be found.

The most interesting to improve the *RF*-measure is using an additional markup of the text in classification. It allows creating a better similarity measure which can use more text features for each substring. This aspect is useful not only for the fiction text classification. In general, many streams (streams of characters, codes, markup entities) can be used in the similarity measure calculation. Also note that the calculation can be done in case of availability of additional information or if it is not.

REFERENCES

1. *Шевелёв О.Г.* Методы автоматической классификации текстов на естественном языке : учеб. пособие. Томск : ТМЛ-Пресс, 2007. 144 с.
2. *Humnisett D., Teahan W.J.* Context-based methods for text categorization // Proceedings of the 27th Annual International ACM SIGIR Conference (SIGIR). The University of Sheffield. UK. 2004.
3. *Ukkonen E.* Constructing Suffix-trees On-Line in Linear Time // Algorithms, Software, Architecture: Information Processing. 1992. № 1(92). 484 p.
4. *Kärkkäinen J., Sanders P.* Simple linear work suffix array construction // ICALP 2003, LNCS 2719 / eds. by J.C.M. Baeten et al. 2003. P. 943–955.
5. *Хмелёв Д.В.* Классификация и разметка текстов с использованием методов сжатия данных. Краткое введение. 2003. URL: http://compression.graphicon.ru/download/articles/classif/intro.html
6. *Khmelev D.V., Teahan W.J.* Verification of text collections for text categorization and natural language processing // Technical Report AIIA 03.1. School of Informatics. University of Wales. Bangor. 2003.
7. *Ашуров М.Ф., Поддубный В.В.* Метод классификации текстов художественной литературы на основе *R*-меры // Новые информационные технологии в исследовании сложных структур : материалы Десятой рос. конф. с междунар. участием. Томск : Издательский Дом Томского государственного университета, 2014. С. 63–64.
8. *Ашуров М.Ф.* Сравнение потоковых методов классификации текстов художественной литературы на основе сжатия информации и подсчета подстрок // Вестник Томского государственного университета. Управление, вычислительная техника и информатика. 2014. № 4(29). С. 16–22.

9. *Ашуров М.Ф., Поддубный В.В.* Потоковый метод классификации текстов художественной литературы на основе *C*-меры // Информационные технологии и математическое моделирование (ИТММ-2013) : материалы XII Всерос. науч.-практ. конф. с междунар. участием им. А.Ф. Терпугова (29–30 ноября 2013 г.). Томск : Изд-во Том. ун-та, 2013. Ч. 2. С. 85–89.
10. *Shevelyov O.G., Poddubnyj V.V.* Complex investigation of texts with the system «StyleAnalyzer» // Text and Lanquage / ed. by P. Grzyber, E. Kelih, J. Macutek. Wien : Praesens Verlag, 2010. P. 207–212.
11. *Van Rijsbergen C.J.* Information Retrieval. London : Butterworths, 1979.

Ashurov Mikhail Faritovich. E-mail: ashurov.mf@gmail.com
Poddubny Vasiliy Vasilievich, doctor of science, professor. E-mail: vvpoddubny@gmail.com
Tomsk State University, Russian Federation. Поступила в редакцию 24 сентября 2015 г.

Ашуров М.Ф., Поддубный В.В. (Томский государственный университет, Российская Федерация).
Основанный на *R*-мере подход к потоковой классификации текстов, использующий частоты повторения подстрок.
Ключевые слова: потоковые методы классификации; *R*-мера; частота повторения подстрок; качество классификации; *F*-мера.

DOI: 10.17223/19988605/33/1

Предлагается потоковый метод классификации текстов на основе *R*-меры с использованием частот повторения подстрок. На материале текстов русской художественной прозы XIX в. и 90-х гг. XX в. проводится сравнительный анализ качества работы алгоритмов распознавания авторских классов классификаторами, построенными на основе усечённой *R*-меры как с использованием, так и без использования частот повторения подстрок испытуемого текста. Качество классификации оценивается известной *F*-мерой Ван Ризбергена. Показано, что в отсутствие жанрового смешения текстов внутри авторских классов учёт частот повторения подстрок исследуемого текста в текстовых суперклассах приводит к повышению качества классификации.

Усечённая *R*-мера близости текстов учитывает все возможные повторения всех подстрок длин от k_1 до k_2 (для неусечённой *R*-меры $k_1 = 1$, $k_2 = n$) испытуемого текста длины n в супертексте:

$$R(X \mid S) = r(X \mid S) / N, \quad r\left(X \mid S\right) = \sum_{k=k_1}^{k_2} c_k\left(X \mid S\right), \quad N = \left(2(n+1) - (k_1 + k_2)\right)\left(k_2 - k_1 + 1\right) / 2,$$

$$c_k\left(X \mid S\right) = \sum_{i=k}^{n} c\left(x_{i-k+1} \ldots x_n \mid S\right), \quad c\left(x_{i-k+1} \ldots x_n \mid S\right) = \begin{cases} 1, x_{i-k+1} \ldots x_n \subset S, \\ 0, x_{i-k+1} \ldots x_n \not\subset S, \end{cases}$$

где X – испытуемый текст, S – супертекст (объединённый текст) исследуемого класса, N – число подстрок.

В работе предлагается метод потоковой классификации, основанный на модификации *R*-меры с учётом частот повторений подстрок. Эта мера (обозначим её как *RF*-мера) использует тот же механизм усечённости:

$$RF(X \mid S) = rf(X \mid S) / N, \quad rf\left(X \mid S\right) = \sum_{k=k_1}^{k_2} \sum_{i=0}^{n-k} rf\left(x_i \ldots x_{i+k} \mid S\right),$$

$$rf\left(x_i \ldots x_{i+k} \mid S\right) = \begin{cases} rfc(x_i \ldots x_{i+k}), x_i \ldots x_{i+k} \in S, \\ 0, x_i \ldots x_{i+k} \notin S, \end{cases} \quad rfc(x_i \ldots x_{i+k}) = \frac{\min(f_X(x_i \ldots x_{i+k}), f_S(x_i \ldots x_{i+k}))}{\max(f_X(x_i \ldots x_{i+k}), f_S(x_i \ldots x_{i+k}))}.$$

Относительные частоты повторения подстрок для исследуемого текста X и супертекста S вычисляются следующим образом:

$$f_X(x_i \ldots x_{i+k}) = \frac{cc_X(x_i \ldots x_{i+k})}{N_X} \text{ и } f_S(x_i \ldots x_{i+k}) = \frac{cc_S(x_i \ldots x_{i+k})}{N_S},$$

где функция $cc\left(x_i \ldots x_{i+k}\right)$, в отличие от функции-индикатора $c\left(x_i \ldots x_{i+k} \mid S\right)$, возвращает количество совпадений подстроки поиска в исследуемом тексте или супертексте (обозначенные соответствующими индексами X и S), а N_X и N_S – количество символов в текстах X и S соответственно.

Качество классификации по каждому классу оценивалось по текстам контрольной выборки *F*-мерой Ван Ризбергена – средним гармоническим между полнотой r (долей текстов, приписываемых классу из всех текстов этого класса) и точностью p (долей текстов, правильно приписываемых этому классу из всех текстов, приписываемых этому классу).

Оба классификатора на рассматриваемых данных (текстах русской художественной литературы XIX и XX вв.) дают качество классификации более 70%, причём в условиях отсутствия жанрового смешения текстов внутри авторских классов учёт частот повторения подстрок исследуемого текста в супертекстах приводит к повышению качества классификации.

REFERENCES

1. Shevelev, O.G. (2007) *Metody avtomaticheskoy klassifikatsii tekstov na estestvennom yazyke* [Approaches of automatic natural language text classification]. Tomsk : TML-Press.
2. Humnisett, D. & Teahan, W.J. (2004) Context-based methods for text categorization. *Proceedings of the 27th Annual International ACM SIGIR Conference (SIGIR)*. The University of Sheffield. UK. 2004. DOI: 10.1145/1008992.1009129
3. Ukkonen, E. (1992) Constructing Suffix-trees On-Line in Linear Time. *Algorithms, Software, Architecture: Information Processing*. 1(92).

4. Kärkkäinen, J. & Sanders, P. (2003) Simple linear work suffix array construction. *ICALP 2003, LNCS 2719*. pp. 943-955. DOI: 10.1007/3-540-45061-0_73
5. Khmelev, D.V. (2003) *Klassifikatsiya i razmetka tekstov s ispol'zovaniem metodov szhatiya dannykh* [Text classification and markup using data compression approaches]. [Online] Available from: http://compression.graphicon.ru/download/articles/classif/intro.html
6. Khmelev, D.V. & Teahan, W.J. (2003) Verification of text collections for text categorization and natural language processing. *Technical Report AIIA 03.1. School of Informatics*. University of Wales. Bangor.
7. Ashurov, M.F. & Poddubnyy, V.V. (2014) [Approach based on R-measure of fiction text classification]. *Novye informatsionnye tekhnologii v issledovanii slozhnykh struktur* [New information technologies in complex structure research]. Proc. Of the 10th Conference with International Participation. Tomsk: Tomsk State University. pp. 63-64.
8. Ashurov, M.F. (2014) Comparison of stream-based fiction text classification methods based on data compression and counting substrings. *Vestnik Tomskogo gosudarstvennogo universiteta. Upravlenie, vychislitel'naya tekhnika i informatika – Tomsk State University Journal of Control and Computer Science*. 4(29). pp. 16-22. (In Russian).
9. Ashurov, M.F. & Poddubnyy, V.V. (2013) [Approach based on C-measure of fiction text classification]. *Informatsionnye tekhnologii i matematicheskoe modelirovanie (ITMM-2013)* [Information technology and mathematical modeling (ITMM-2013)]. Proceedings of the 12th Russian Scientific-Practical Conference with International Participation named after A.F. Terpugov. Tomsk. 29th – 30th November 2013. Tomsk: tomsk State University. pp. 85-89. (In Russian).
10. Shevelyov, O.G. & Poddubnyy, V.V. (2010) Complex investigation of texts with the system “StyleAnalyzer”. In: Grzyber, P., Kelih, E. & Macutek, J. (eds) *Text and Lanquage*. Vienna: Praesens Verlag. pp. 207-212.
11. Van Rijsbergen, C.J. (1979) *Information Retrieval*. London: Butterworths.

ВЕСТНИК ТОМСКОГО ГОСУДАРСТВЕННОГО УНИВЕРСИТЕТА

2015 Управление, вычислительная техника и информатика № 4 (33)

УДК 519.21
DOI: 10.17223/19988605/33/2

А.М. Горцев, А.А. Соловьев

СРАВНЕНИЕ МП- И ММ-ОЦЕНОК ДЛИТЕЛЬНОСТИ НЕПРОДЛЕВАЮЩЕГОСЯ МЕРТВОГО ВРЕМЕНИ В МАР-ПОТОКЕ СОБЫТИЙ

Изучается МАР-поток событий, функционирующий в условиях непродлевающегося мертвого времени, когда длительность мертвого времени – неизвестная фиксированная величина. Проводится сравнение качества оценок непродлевающегося мертвого времени, получаемых (по наблюдениям за моментами наступления событий МАР-потока) методом максимального правдоподобия (МП-оценка) и методом моментов (ММ-оценка).

Ключевые слова: МАР-поток; непродлевающееся мертвое время; метод максимального правдоподобия; оценка максимального правдоподобия; метод моментов; оценка метода моментов.

Широко применяемой математической моделью реальных физических процессов являются случайные потоки событий. Изучаемый МАР-поток событий относится к классу дважды стохастических потоков и является одной из адекватных математических моделей потоков элементарных частиц (фотонов, электронов и т.д.), информационных потоков событий, функционирующих в цифровых сетях интегрального обслуживания, информационно-телекоммуникационных системах, спутниковых сетях связи и т.д. Задачи по оценке состояний и параметров случайных потоков событий возникают в оптических и лазерных системах, функционирующих в режиме счета фотонов, например, при лазерном зондировании высотных слоев атмосферы, в оптических системах обнаружения, распознавания и сопровождения, работающих через атмосферу на предельно большие расстояния, а также в оптических системах загоризонтной связи.

В настоящей работе проводится дальнейшее исследование потоков физических событий, функционирующих в условиях непродлевающегося мертвого времени, начатое в работах [1–9].

Параметры потоков событий, функционирующих в реальном времени, неизвестны частично, либо полностью, либо представляют собой функцию времени. В подобных ситуациях наиболее рациональным является применение адаптивных систем, которые в процессе функционирования оценивают неизвестные параметры либо состояния потоков событий и изменяют дисциплины обслуживания в соответствии с полученными оценками [10]. Вследствие этого возникают задачи: 1) оценки состояний потока (задача фильтрации интенсивности потока) по наблюдениям за моментами наступления событий [11–14]; 2) оценки параметров потока по наблюдениям за моментами наступления событий [15–24].

Одним из искажающих факторов при оценке состояний и параметров потока событий (фотонов, электронов и т.д.) выступает мертвое время регистрирующих приборов [25], которое порождается зарегистрированным событием (фотоном, электроном и т.д.). Другие же события, наступившие в течение периода мертвого времени, недоступны наблюдению (теряются). Можно считать, что этот период продолжается некоторое фиксированное время (непродлевающееся мертвое время). Таким образом, эффект мертвого времени влечет за собой потери событий потока (частично теряется информация о потоке событий), что в итоге отрицательно сказывается на оценивании как состояний, так и параметров потока [2–9]. Для того чтобы оценить потери событий потока, возникающие из-за эффекта мертвого времени, необходимо оценить значение его длительности.

При оценке параметров потока событий обычно используется метод моментов [3–8, 16–18, 21–23], как более простой в аналитическом исполнении реже используется метод максимального правдоподобия [9, 24, 26–28]. Последнее связано с возникающими при использовании этого метода аналитическими трудностями.

В настоящей статье производится численное сравнение оценок значений длительности мертвого времени в МАР-потоке событий, полученных методом максимального правдоподобия (МП-оценка) и методом моментов (ММ-оценка).

1. Математическая модель MAP-потока событий

Рассматривается MAP-поток с интенсивностью, представляющей собой кусочно-постоянный случайный процесс $\lambda(t)$ с двумя состояниями: $\lambda(t) = \lambda_1$ либо $\lambda(t) = \lambda_2$ ($\lambda_1 > \lambda_2$). Длительность пребывания процесса $\lambda(t)$ в i-м состоянии есть случайная величина с экспоненциальной функцией распределения $F_i = 1 - e^{\lambda_i t}$, $i = 1, 2$. В момент окончания i-го состояния процесса $\lambda(t)$ возможны следующие ситуации, каждая из которых протекает мгновенно: 1) процесс $\lambda(t)$ переходит из i-го состояния в i-е и наступает событие потока в i-ом состоянии; совместная вероятность этой ситуации $P(\lambda_i \to \lambda_i, 1) = P_1(\lambda_i \mid \lambda_i)$, $i = 1, 2$; 2) процесс $\lambda(t)$ переходит из i-го состояния в j-е и наступает событие потока; совместная вероятность этой ситуации есть $P(\lambda_i \to \lambda_j, 1) = P_1(\lambda_j \mid \lambda_i), i, j = 1, 2; i \neq j$; 3) процесс $\lambda(t)$ переходит из i-го состояния в j-е, и событие потока не наступает; совместная вероятность этой ситуации есть $P(\lambda_i \to \lambda_j, 0) = P_0(\lambda_j \mid \lambda_i)$, $i, j = 1,2; i \neq j$. При этом $P_1(\lambda_i \mid \lambda_i) + P_1(\lambda_j \mid \lambda_i) + P_0(\lambda_j \mid \lambda_i) = 1$, $i, j = 1, 2; i \neq j$. Блочная матрица инфинитезимальных характеристик процесса $\lambda(t)$ при этом примет вид

$$D = \begin{Vmatrix} -\lambda_1 & \lambda_1 P_0(\lambda_2 \mid \lambda_1) \\ \lambda_2 P_0(\lambda_1 \mid \lambda_2) & -\lambda_2 \end{Vmatrix} \begin{Vmatrix} \lambda_1 P_1(\lambda_1 \mid \lambda_1) & \lambda_1 P_1(\lambda_2 \mid \lambda_1) \\ \lambda_2 P_1(\lambda_1 \mid \lambda_2) & \lambda_2 P_1(\lambda_2 \mid \lambda_2) \end{Vmatrix} = \|D_0 \mid D_1\|.$$

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.

Заметим, что в определении MAP-потока событий в явном виде не оговаривается, в каком состоянии процесса $\lambda(t)$ наступает событие потока при переходе процесса $\lambda(t)$ из i-го состояния в j-е ($i,j = 1, 2; i \neq j$). Данное обстоятельство при последующих аналитических выкладках является несущественным, так как наступление события и переход процесса из i-го состояния в j-е ($i, j = 1, 2; i \neq j$) происходят мгновенно.

После каждого зарегистрированного в момент времени t_k события наступает время фиксированной длительности T (мертвое время), в течение которого другие события исходного MAP-потока недоступны наблюдению. По окончании мертвого времени первое наступившее событие снова создает период мертвого времени длительности T и т.д. Вариант возникающей ситуации приведен на рис. 1, где t_1, t_2,… – моменты наступления событий в наблюдаемом потоке; *1* и *2* – состояния случайного процесса $\lambda(t)$; штриховкой обозначены длительности мертвого времени; черными кружками обозначены события MAP-потока, недоступные наблюдению.

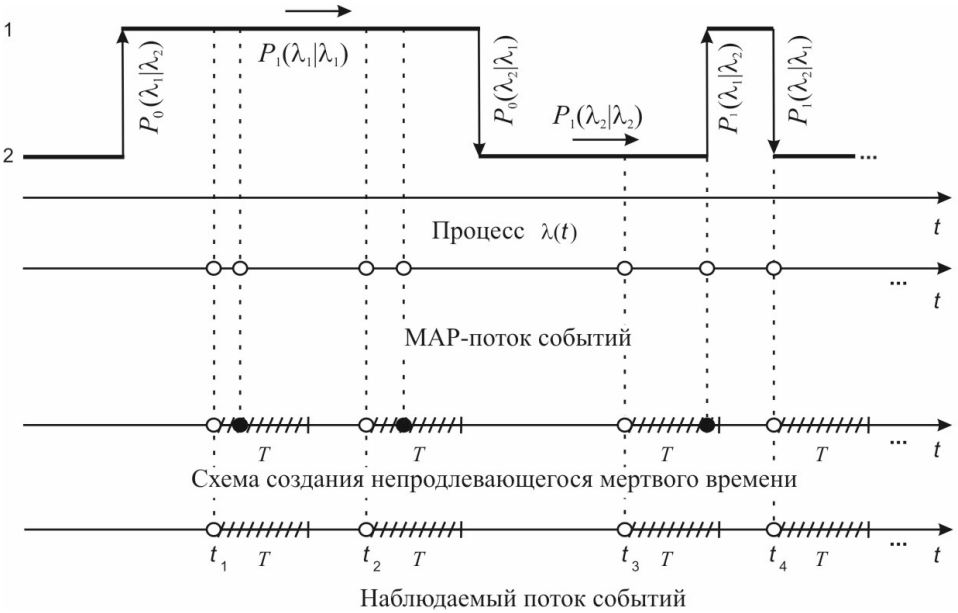


Рис. 1. Формирование наблюдаемого потока событий

Процесс $\lambda(t)$ является принципиально ненаблюдаемым (скрытый марковский процесс), а наблюдаемыми являются только временные моменты наступления событий t_1, t_2,… наблюдаемого потока. Рассматривается установившийся (стационарный) режим функционирования наблюдаемого потока событий, поэтому переходными процессами на интервале наблюдения (t_0, t), где t_0 – начало наблюдений, t – окончание наблюдений, пренебрегаем. Необходимо в момент окончания наблюдений (в момент времени t) осуществить методом максимального правдоподобия и методом моментов оценку $\hat{T}$ значений длительности мертвого времени и произвести сравнение качества полученных оценок.

2. Оценка длительности мертвого времени методом максимального правдоподобия

Обозначим $\tau_k = t_{k+1} - t_k$ (k = 1,2,…) значение длительности k-го интервала между соседними событиями наблюдаемого потока ($\tau_k > 0$). Так как рассматривается стационарный режим, то плотность вероятностей значений длительности k-го интервала $p_T(\tau_k) = p_T(\tau)$ ($\tau \geq 0$) для любого k (индекс T подчеркивает, что плотность вероятностей зависит от длительности мертвого времени). Здесь и далее ситуация, когда $\tau = 0$, означает доопределение изучаемых функций в граничной точке. В силу сказанного момент времени t_k без потери общности можно положить равным нулю или, что то же самое, момент наступления события наблюдаемого потока есть $\tau = 0$. Тогда [1] плотность вероятности примет вид

$$p_T(\tau) = \begin{cases} 0, 0 \leq \tau \leq T, \\ \dfrac{z_1}{z_2 - z_1}\left[z_2 - \dfrac{1}{\beta_1 + \beta_2} f(T)\right] e^{-z_1(\tau - T)} - \dfrac{z_2}{z_2 - z_1}\left[z_1 - \dfrac{1}{\beta_1 + \beta_2} f(T)\right] e^{-z_2(\tau - T)}, \tau \geq T, \end{cases}$$

$$f(T) = \lambda_1\lambda_2 A + \left\{\lambda_1\left[1 - P_0(\lambda_2 \mid \lambda_1)\right] - \lambda_2\left[1 - P_0(\lambda_1 \mid \lambda_2)\right]\right\} \frac{\beta_1 P_1 - \beta_2 P_2}{F(T)} e^{-(\beta_1 + \beta_2)T},$$

$$F(T) = 1 - P_0(\lambda_1 \mid \lambda_2) P_0(\lambda_2 \mid \lambda_1) - P e^{-(\beta_1 + \beta_2)T};$$

$$\beta_1 = \lambda_1\left[1 - P_1(\lambda_1 \mid \lambda_1)\right], \ \beta_2 = \lambda_2\left[1 - P_1(\lambda_2 \mid \lambda_2)\right]; \qquad (1)$$

$$A = P_1 + P_2, \ P_1 = P_1(\lambda_1 \mid \lambda_2) + P_1(\lambda_1 \mid \lambda_1) P_0(\lambda_1 \mid \lambda_2), \ P_2 = P_1(\lambda_2 \mid \lambda_1) + P_1(\lambda_2 \mid \lambda_2) P_0(\lambda_2 \mid \lambda_1),$$

$$p = \frac{P_1(\lambda_1 \mid \lambda_1) P_1(\lambda_2 \mid \lambda_2) - P_1(\lambda_1 \mid \lambda_2) P_1(\lambda_2 \mid \lambda_1)}{1 - P_0(\lambda_1 \mid \lambda_2) P_0(\lambda_2 \mid \lambda_1)};$$

$$z_{1,2} = \frac{1}{2}\left[\lambda_1 + \lambda_2 \mp \sqrt{(\lambda_1 - \lambda_2)^2 + 4\lambda_1\lambda_2 \mathrm{P}_0(\lambda_2 \mid \lambda_1) \mathrm{P}_0(\lambda_1 \mid \lambda_2)}\right], 0 < z_1 < z_2.$$

В (1) функция $F(T)>0$ для любых T ($0 \leq T \leq \tau$).

Пусть $\tau_1 = t_2 - t_1$, $\tau_2 = t_3 - t_2$,…, $\tau_k = t_{k+1} - t_k$ – последовательность измеренных (в результате наблюдения за потоком в течение интервала наблюдения $(0,t)$) значений длительности интервалов между соседними событиями потока. Упорядочим величины $\tau_1 \ldots \tau_k$ по возрастанию: $\tau_{\min} = \tau^{(1)} < \tau^{(2)} < \ldots < \tau^{(k)}$. В силу предпосылок наблюдаемый поток обладает марковским свойством, если его эволюцию рассматривать начиная с момента наступления события (с момента t_k, k = 1,2,…). Тогда функция правдоподобия, с учетом (1) [29], запишется в виде

$$L(\lambda_i, P_1(\lambda_i \mid \lambda_i), P_1(\lambda_i \mid \lambda_j), P_0(\lambda_i \mid \lambda_j), T \mid \tau^{(1)}, \ldots, \tau^{(k)}) = 0, \ 0 \leq \tau_{\min} < T;$$

$$L(\lambda_i, P_1(\lambda_i \mid \lambda_i), P_1(\lambda_i \mid \lambda_j), P_0(\lambda_i \mid \lambda_j), T \mid \tau^{(1)}, \ldots, \tau^{(k)}) = \prod_{j=1}^{k} p_T(\tau^{(j)}), \ \tau_{\min} \geq T.$$

Поскольку поставленная задача заключается в построении оценки $\hat{T}$ значения длительности мертвого времени (в предположении, что остальные параметры потока λ_i, $P_1(\lambda_i|\lambda_i)$, $\mathrm{P}_1(\lambda_i|\lambda_j)$, $\mathrm{P}_0(\lambda_i|\lambda_j)$, i, j = 1,2, $i \neq j$, известны), то согласно методу максимального правдоподобия ее реализация есть решение оптимизационной задачи:

$$L(T\mid\tau^{(1)},\ldots,\tau^{(k)})=\prod_{j=1}^{k}p_T(\tau^{(j)})=\prod_{j=1}^{k}\{\frac{z_1}{z_2-z_1}\left[z_2-\frac{1}{\beta_1+\beta_2}f(T)\right]e^{-z_1(\tau^{(j)}-T)}-$$
$$-\frac{z_2}{z_2-z_1}\left[z_1-\frac{1}{\beta_1+\beta_2}f(T)\right]e^{-z_2(\tau^{(j)}-T)}\}\Rightarrow\max_T,\ 0\le T\le\tau_{\min},\tau_{\min}>0, \qquad (2)$$

где $f(T)$, β_1, β_2, z_1, z_2 определены в (1).

Значение T, при котором (2) достигает своего глобального максимума, есть $\hat{T}_{МП}$ – оценка максимального правдоподобия значения длительности мертвого времени.

Строго аналитическое решение оптимизационной задачи (2) приводит к следующему результату: при любых значениях параметров MAP-потока событий оценка $\hat{T}_{МП}=\tau_{\min}$.

3. Оценка длительности мертвого времени методом моментов

В [1] показано, что MAP-поток событий, функционирующий в условиях непродлевающегося мертвого времени, в общем случае является коррелированным. Только в частных случаях поток становится рекуррентным.

Пусть $(t_k,\ t_{k+1})$, $(t_{k+1},\ t_{k+2})$ – два смежных интервала в наблюдаемом потоке с соответствующими значениями длительностей: $\tau_k=t_{k+1}-t_k$, $\tau_{k+1}=t_{k+2}-t_{k+1}$; их расположение на временной оси, в силу стационарности потока, произвольно. Тогда можно положить $k=1$ и рассмотреть соседние интервалы $(t_1,\ t_2)$, $(t_2,\ t_3)$ с соответствующими значениями длительностей: $\tau_1=t_2-t_1$, $\tau_2=t_3-t_2$; $\tau_1\ge 0$, $\tau_2\ge 0$. При этом $\tau_1=0$ соответствует моменту t_1 наступления события наблюдаемого потока; $\tau_2=0$ соответствует моменту t_2 наступления следующего события наблюдаемого потока. Соответствующая совместная плотность вероятности при этом есть $p_T(\tau_1,\ \tau_2)$, $\tau_1\ge 0$, $\tau_2\ge 0$ [1]:

$$p_T(\tau_1,\tau_2)=0,\ 0\le\tau_1\le T,\ 0\le\tau_1\le T;\ p_T(\tau_1,\tau_2)=p_T(\tau_1)p_T(\tau_2)+e^{-(\beta_1+\beta_2)T}C_T\times$$
$$\times\left(z_1e^{-z_1(\tau_1-T)}-z_2e^{-z_2(\tau_1-T)}\right)\left(z_1e^{-z_1(\tau_2-T)}-z_2e^{-z_2(\tau_2-T)}\right),\tau_1\ge T,\tau_2\ge T;$$

$$C_T=\frac{p}{\left[(z_1-z_2)(\beta_1+\beta_2)\right]^2}\left[1-P_0(\lambda_1|\lambda_2)P_0(\lambda_2|\lambda_1)\right]\{\lambda_1\left[1-P_0(\lambda_2|\lambda_1)\right]-\lambda_2\left[1-P_0(\lambda_1|\lambda_2)\right]\}\times$$
$$\times\frac{\beta_1\left[P_1(\lambda_1|\lambda_2)+P_1(\lambda_1|\lambda_1)P_0(\lambda_1|\lambda_2)\right]-\beta_2\left[P_1(\lambda_2|\lambda_1)+P_1(\lambda_2|\lambda_2)P_0(\lambda_2|\lambda_1)\right]}{\left\{1-P_0(\lambda_1|\lambda_2)P_0(\lambda_2|\lambda_1)-\left[P_1(\lambda_1|\lambda_1)P_1(\lambda_2|\lambda_2)-P_1(\lambda_1|\lambda_2)P_1(\lambda_2|\lambda_1)\right]e^{-(\beta_1+\beta_2)T}\right\}^2}\times \qquad (3)$$
$$\times\{z_1z_2-\left[2z_1z_2-(\beta_1+\beta_2)(\lambda_1+\lambda_2)\right]e^{-(\beta_1+\beta_2)T}+$$
$$+\left[z_1z_2-(\beta_1+\beta_2)\left(\lambda_1P_1(\lambda_1|\lambda_2)+\lambda_2P_1(\lambda_2|\lambda_2)\right)\right]e^{-2(\beta_1+\beta_2)T}\},$$

$z_1,z_2,\beta_1,\beta_2,p,p_T(\tau_1),p_T(\tau_2)$ для $\tau=\tau_1$ и $\tau=\tau_2$ определены в (1).

Теоретическая ковариация значений τ_1 и τ_2 запишется в виде

$$\mathrm{cov}(\tau_1,\tau_2)=\int_T^\infty\int_T^\infty\tau_1\tau_2p_T(\tau_1,\tau_2)d\tau_1d\tau_2-\left[\int_T^\infty\tau p_T(\tau)d\tau\right]^2. \qquad (4)$$

Подставляя (1),(3) в (4), находим явный вид теоретической ковариации:

$$\mathrm{cov}(\tau_1,\tau_2)=\left(\frac{z_2-z_1}{z_1z_2}\right)^2e^{-(\beta_1+\beta_2)T}C_T, \qquad (5)$$

где C_T определена в (3).

Пусть за время наблюдения MAP-потока событий в течении временного интервала (t_0,t) реализовалось n интервалов (t_k,t_{k+1}) длительности τ_k, $k=\overline{1,n}$. Введем статистику:

$$\widehat{\mathrm{cov}}(\tau_1,\tau_2)=\frac{1}{n-1}\sum_{k=1}^{n-1}\tau_k\tau_{k+1}-\left(\frac{1}{n}\sum_{k=1}^{n}\tau_k\right)^2, \tag{6}$$

где $\widehat{\mathrm{cov}}(\tau_1,\tau_2)$ является оценкой теоретической ковариации (5). Тогда согласно методу моментов [29] уравнение моментов, учитывающее коррелированность MAP-потока событий, запишется в виде

$$\left(\frac{z_2-z_1}{z_1z_2}\right)^2 e^{-(\beta_1+\beta_2)T}C_T=\widehat{\mathrm{cov}}(\tau_1,\tau_2). \tag{7}$$

Подставляя в (7) выражение C_T из (3), вводя новую переменную $x=e^{-(\beta_1+\beta_2)T}$ и проделывая при этом необходимые преобразования, находим (7) в виде

$$\begin{gathered}
ax^3+bx^2+cx+d=0,\\
a=h\{z_1z_2-(\beta_1+\beta_2)[\lambda_1P_1(\lambda_1\mid\lambda_1)+\lambda_2P_1(\lambda_2\mid\lambda_2)]\},\\
b=-\{\widehat{\mathrm{cov}}(\tau_1,\tau_2)p^2+h[2z_1z_2-(\beta_1+\beta_2)(\lambda_1+\lambda_2)]\},\\
c=[z_1z_2h+2p\,\widehat{\mathrm{cov}}(\tau_1,\tau_2)],\ d=-\widehat{\mathrm{cov}}(\tau_1,\tau_2),\\
h=p[z_1z_2(\beta_1+\beta_2)]^{-2}\{\beta_1[P_1(\lambda_1\mid\lambda_2)+P_1(\lambda_1\mid\lambda_1)P_0(\lambda_1\mid\lambda_2)]-\\
-\beta_2[P_1(\lambda_2\mid\lambda_1)+P_1(\lambda_2\mid\lambda_2)P_0(\lambda_2\mid\lambda_1)]\}\times\{\lambda_1[1-P_0(\lambda_2\mid\lambda_1)]-\lambda_2[1-P_0(\lambda_1\mid\lambda_2)]\}.
\end{gathered} \tag{8}$$

Решение уравнения (8) определит три корня x_i , i = 1, 2, 3, которые, в свою очередь, дают три оценки длительности мертвого времени:

$$\hat{T}_{MM}{}^{(i)}=-\frac{1}{\beta_1+\beta_2}\ln(x_i),\ i=1,2,3.$$

Алгоритм определения единственной оценки $\hat{T}_{MM}$ следующий:

1) для определенного набора параметров λ_1, λ_2, $P_1(\lambda_i|\,\lambda_i)$, $P_1(\lambda_i|\,\lambda_j)$, $P_0(\lambda_i|\,\lambda_j)$, i, j = 1, 2; $i \neq j$, T, n – количество наблюдаемых событий, осуществляется имитационное моделирование наблюдаемого потока;

2) результатом работы имитационной модели является оценка теоретической ковариации (6);

3) решается уравнение (8), т.е. находятся корни x_1, x_2, x_3;

4) если все корни комплексные, то $\hat{T}_{MM}=\tau_{\min}$;

5.1) если положительный вещественный корень один – x_1, тогда:

а) если $\hat{T}_{MM}{}^{(1)}\le 0$, то $\hat{T}_{MM}=\tau_{\min}$; б) если $\hat{T}_{MM}{}^{(1)}>0$, то: б.1) если $\hat{T}_{MM}{}^{(1)}\le\tau_{\min}$, то $\hat{T}_{MM}=\hat{T}_{MM}{}^{(1)}$; б.2) если $\hat{T}_{MM}{}^{(1)}>\tau_{\min}$, то $\hat{T}_{MM}=\tau_{\min}$;

5.2) если положительных вещественных корня два – x_1, x_2 ($\hat{T}_{MM}{}^{(1)}\le\hat{T}_{MM}{}^{(2)}$), тогда: а) если $\hat{T}_{MM}{}^{(1)}\le\hat{T}_{MM}{}^{(2)}\le 0$, то $\hat{T}_{MM}=\tau_{\min}$; б) если $\hat{T}_{MM}{}^{(1)}\le 0<\hat{T}_{MM}{}^{(2)}$, то: б.1) если $\hat{T}_{MM}{}^{(2)}\le\tau_{\min}$, то $\hat{T}_{MM}=\hat{T}_{MM}{}^{(2)}$; б.2) если $\hat{T}_{MM}{}^{(2)}>\tau_{\min}$, то $\hat{T}_{MM}=\tau_{\min}$; в) если $0<\hat{T}_{MM}{}^{(1)}\le\hat{T}_{MM}{}^{(2)}$, то: в.1) если $\tau_{\min}\le\hat{T}_{MM}{}^{(1)}\le\hat{T}_{MM}{}^{(2)}$, то $\hat{T}_{MM}=\tau_{\min}$; в.2) если $\hat{T}_{MM}{}^{(1)}\le\tau_{\min}\le\hat{T}_{MM}{}^{(2)}$, то $\hat{T}_{MM}=\hat{T}_{MM}{}^{(1)}$; в.3) если $\hat{T}_{MM}{}^{(1)}\le\hat{T}_{MM}{}^{(2)}\le\tau_{\min}$, то $\hat{T}_{MM}=\hat{T}_{MM}{}^{(2)}$;

5.3) если положительных вещественных корня три – x_1, x_2 , x_3 ($\hat{T}_{MM}{}^{(1)}\le\hat{T}_{MM}{}^{(2)}\le\hat{T}_{MM}{}^{(3)}$), тогда: а) если $\hat{T}_{MM}{}^{(1)}\le\hat{T}_{MM}{}^{(2)}\le\hat{T}_{MM}{}^{(3)}\le 0$, то $\hat{T}_{MM}=\tau_{\min}$; б) если $\hat{T}_{MM}{}^{(1)}\le\hat{T}_{MM}{}^{(2)}\le 0<\hat{T}_{MM}{}^{(3)}$, то: б.1) если $\hat{T}_{MM}{}^{(3)}\le\tau_{\min}$, то $\hat{T}_{MM}=\hat{T}_{MM}{}^{(3)}$; б.2) если $\hat{T}_{MM}{}^{(3)}>\tau_{\min}$, то $\hat{T}_{MM}=\tau_{\min}$; в) если $\hat{T}_{MM}{}^{(1)}\le 0<\hat{T}_{MM}{}^{(2)}\le\hat{T}_{MM}{}^{(3)}$, то: в.1) если $\tau_{\min}\le\hat{T}_{MM}{}^{(2)}\le\hat{T}_{MM}{}^{(3)}$, то $\hat{T}_{MM}=\tau_{\min}$; в.2) если $\hat{T}_{MM}{}^{(2)}\le\tau_{\min}\le\hat{T}_{MM}{}^{(3)}$, то $\hat{T}_{MM}=\hat{T}_{MM}{}^{(2)}$; в.3) если $\hat{T}_{MM}{}^{(2)}\le\hat{T}_{MM}{}^{(3)}\le\tau_{\min}$, то $\hat{T}_{MM}=\hat{T}_{MM}{}^{(3)}$; г) если $0<\hat{T}_{MM}{}^{(1)}\le\hat{T}_{MM}{}^{(2)}\le\hat{T}_{MM}{}^{(3)}$, то: г.1) если $\tau_{\min}\le\hat{T}_{MM}{}^{(1)}\le$ $\le\hat{T}_{MM}{}^{(2)}\le\hat{T}_{MM}{}^{(3)}$, то $\hat{T}_{MM}=\tau_{\min}$; г.2) если $\hat{T}_{MM}{}^{(1)}\le\tau_{\min}\le\hat{T}_{MM}{}^{(2)}\le\hat{T}_{MM}{}^{(3)}$, то $\hat{T}_{MM}=\hat{T}_{MM}{}^{(1)}$; г.3) если

$\hat{T}_{MM}^{(1)} \le \hat{T}_{MM}^{(2)} \le \tau_{\min} \le \hat{T}_{MM}^{(3)}$, то $\hat{T}_{MM} = \hat{T}_{MM}^{(2)}$; г.4) если $\hat{T}_{MM}^{(1)} \le \hat{T}_{MM}^{(2)} \le \hat{T}_{MM}^{(3)} \le \tau_{\min}$, то $\hat{T}_{MM} = \hat{T}_{MM}^{(3)}$.

В результате работы алгоритма находится единственная ММ-оценка $\hat{T}_{MM}$ значения длительности мертвого времени.

4. Численное сравнение МП- и ММ-оценок

Для получения численных результатов разработан алгоритм вычисления МП- и ММ-оценок. Программа расчета реализована на языке программирования Borland C++, Builder 6. Первый этап расчета предполагает имитационное моделирование (при заданных параметрах λ_1, λ_2, $P_1(\lambda_i|\lambda_i)$, $P_1(\lambda_i|\lambda_j)$, $P_0(\lambda_i|\lambda_j)$, $i, j = 1, 2$; $i \neq j$, T, n) наблюдаемого потока событий. Описание алгоритма имитационного моделирования здесь не приводится, так как никаких принципиальных трудностей он не содержит. Результатом работы имитационной модели является набор значений длительности временных интервалов τ_1, τ_2,..., τ_n (n = 2,3,…). На втором этапе расчета вычисляются МП- и ММ-оценки: 1) на основе полученной на первом этапе выборки τ_1, τ_2,..., τ_n (n = 2,3,…) находится МП-оценка $\hat{T}_{MП} = \min \tau_k (k = \overline{1,n})$; 2) вычисляется значение оценки теоретической ковариации (6); 3) решается уравнение (8) и находится единственная ММ-оценка $\hat{T}_{MM}$; 4) вычисляются величины $\Delta\hat{T}_{MM} = (\hat{T}_{MM} - T)^2$, $\Delta\hat{T}_{MП} = (\hat{T}_{MП} - T)^2$, где T – истинное значение длительности мертвого времени, заданное на первом этапе расчета.

Для сравнения методов по качеству получаемых оценок проведен статистический эксперимент, состоящий из следующих этапов: 1) для заданного набора параметров λ_1, λ_2, $P_1(\lambda_i|\lambda_i)$, $P_1(\lambda_i|\lambda_j)$, $P_0(\lambda_i|\lambda_j)$, $i, j = 1, 2$; $i \neq j$, T, n осуществляется моделирование наблюдаемого потока (отдельный j-й эксперимент j = 1,2,…,N); 2) вычисляются значения оценок $\hat{T}_{MM}^{(j)}$, $\hat{T}_{MП}^{(j)}$ и значения $\Delta\hat{T}_{MM}^{(j)}$, $\Delta\hat{T}_{MП}^{(j)}$; 3) осуществляется повторение N ($j = \overline{1,N}$) раз шагов 1–2.

Результатом работы приведенного выше алгоритма являются две выборки: $(\Delta\hat{T}_{MM}^{(1)},...,\Delta\hat{T}_{MM}^{(N)})$ и $(\Delta\hat{T}_{MП}^{(1)},...,\Delta\hat{T}_{MП}^{(N)})$, на основе которых вычисляются выборочные вариации $\hat{V}_{MП} = \frac{1}{N}\sum_{j=1}^{N}\Delta\hat{T}_{MП}^{(j)}$ и $\hat{V}_{MM} = \frac{1}{N}\sum_{j=1}^{N}\Delta\hat{T}_{MM}^{(j)}$.

Сравнивая значения выборочных вариаций, определяется, какая из полученных оценок лучше: если $\hat{V}_{MП} > \hat{V}_{MM}$, то ММ-оценка лучше, если $\hat{V}_{MП} < \hat{V}_{MM}$, то лучше МП-оценка.

Результаты статистического эксперимента приведены в табл. 1–8. В первой строке указано количество наблюдаемых событий потока n (длительность имитационного моделирования). Во второй и третьей строках таблиц для каждого n приведены соответствующие численные значения $\hat{V}_{MП}$ и $\hat{V}_{MM}$. В четвертой строке таблиц для каждого n приведены численные значения разности $\hat{V}_{MM} - \hat{V}_{MП}$. Численные результаты для всех таблиц получены для N = 100.

Таблица 1

Результаты статистического эксперимента ($\lambda_1 = 1$, $\lambda_2 = 0{,}9$, $P_1(\lambda_1|\lambda_1) = 0{,}3$, $P_1(\lambda_2|\lambda_1) = 0{,}3$, $P_0(\lambda_2|\lambda_1) = 0{,}4$, $P_1(\lambda_2|\lambda_2) = 0{,}4$, $P_1(\lambda_1|\lambda_2) = 0{,}3$, $P_0(\lambda_1|\lambda_2) = 0{,}3$, $T = 2{,}75$)

n	**10**	**20**	**30**	**40**	**50**
$\hat{V}_{МП}$	0,1517824	0,0326041	0,0075076	0,0018225	0,0002025
$\hat{V}_{ММ}$	0,0349281	0,0056169	0,0011664	0,0010816	0,0001296
$\hat{V}_{ММ} - \hat{V}_{МП}$	–0,1168543	–0,0269872	–0,0063412	–0,0007409	–0,0000729

Таблица 2

Результаты статистического эксперимента ($\lambda_1 = 5$, $\lambda_2 = 1$, $P_1(\lambda_1|\,\lambda_1) = 0{,}3$, $P_1(\lambda_2|\,\lambda_1) = 0{,}3$, $P_0(\lambda_2|\,\lambda_1) = 0{,}4$, $P_1(\lambda_2|\,\lambda_2) = 0{,}4$, $P_1(\lambda_1|\,\lambda_2) = 0{,}3$, $P_0(\lambda_1|\,\lambda_2) = 03$, $T = 0{,}4$)

n	10	20	30	40	50
$\hat{V}_{\text{МП}}$	0,0827554	0,0134249	0,0073008	0,0053361	0,0015154
$\hat{V}_{\text{ММ}}$	0,0382048	0,0051302	0,0048444	0,0034151	0,0006989
$\hat{V}_{\text{ММ}} - \hat{V}_{\text{МП}}$	–0,0445506	–0,0082947	–0,0024564	–0,001921	–0,0008165

Таблица 3

Результаты статистического эксперимента ($\lambda_1 = 2$, $\lambda_2 = 1{,}9$, $P_1(\lambda_1|\,\lambda_1) = 0{,}5$, $P_1(\lambda_2|\,\lambda_1) = 0{,}4$, $P_0(\lambda_2|\,\lambda_1) = 0{,}1$, $P_1(\lambda_2|\,\lambda_2) = 0{,}3$, $P_1(\lambda_1|\,\lambda_2) = 0{,}3$, $P_0(\lambda_1|\,\lambda_2) = 0{,}4$, $T = 1$)

n	10	20	30	40	50
$\hat{V}_{\text{МП}}$	0,1444804	0,0172391	0,0044775	0,0038770	0,0030555
$\hat{V}_{\text{ММ}}$	0,0335241	0,0128594	0,0041494	0,0019210	0,0016461
$\hat{V}_{\text{ММ}} - \hat{V}_{\text{МП}}$	–0,1109563	–0,0043797	–0,0003281	–0,0019560	–0,0014094

Таблица 4

Результаты статистического эксперимента ($\lambda_1 = 0{,}5$, $\lambda_2 = 0{,}1$, $P_1(\lambda_1|\,\lambda_1) = 0{,}7$, $P_1(\lambda_2|\,\lambda_1) = 0{,}2$, $P_0(\lambda_2|\,\lambda_1) = 0{,}1$, $P_1(\lambda_2|\,\lambda_2) = 0{,}6$, $P_1(\lambda_1|\,\lambda_2) = 0{,}3$, $P_0(\lambda_1|\,\lambda_2) = 0{,}1$, $T = 3$)

n	10	20	30	40	50
$\hat{V}_{\text{МП}}$	0,9629230	0,5469985	0,2391044	0,1526460	0,0983072
$\hat{V}_{\text{ММ}}$	0,0915849	0,0651088	0,0299428	0,0154135	0,0028359
$\hat{V}_{\text{ММ}} - \hat{V}_{\text{МП}}$	–0,8713381	–0,4818898	–0,2091616	–0,1372325	–0,0954714

Таблица 5

Результаты статистического эксперимента ($\lambda_1 = 1$, $\lambda_2 = 0{,}9$, $P_1(\lambda_1|\,\lambda_1) = 0{,}3$, $P_1(\lambda_2|\,\lambda_1) = 0{,}3$, $P_0(\lambda_2|\,\lambda_1) = 0{,}4$, $P_1(\lambda_2|\,\lambda_2) = 0{,}4$, $P_1(\lambda_1|\,\lambda_2) = 0{,}3$, $P_0(\lambda_1|\,\lambda_2) = 0{,}3$, $T = 2{,}75$)

n	200	300	600	800	1000
$\hat{V}_{\text{МП}}$	0,0004096	0,0001347	0,0000441	0,0000074	0,0000003
$\hat{V}_{\text{ММ}}$	0,0001505	0,0002692	0,0002601	0,0001849	0,0000838
$\hat{V}_{\text{ММ}} - \hat{V}_{\text{МП}}$	–0,0002591	0,0001345	0,0002160	0,0001775	0,0000835

Таблица 6

Результаты статистического эксперимента ($\lambda_1 = 5$, $\lambda_2 = 1$, $P_1(\lambda_1|\,\lambda_1) = 0{,}3$, $P_1(\lambda_2|\,\lambda_1) = 0{,}3$, $P_0(\lambda_2|\,\lambda_1) = 0{,}4$, $P_1(\lambda_2|\,\lambda_2) = 0{,}4$, $P_1(\lambda_1|\,\lambda_2) = 0{,}3$, $P_0(\lambda_1|\,\lambda_2) = 0{,}3$, $T = 0{,}4$)

n	200	400	600	800	1000
$\hat{V}_{\text{МП}}$	0,0001340	0,0000291	0,0000152	0,0000050	0,0000071
$\hat{V}_{\text{ММ}}$	0,0001391	0,0002228	0,0004602	0,0002382	0,0002200
$\hat{V}_{\text{ММ}} - \hat{V}_{\text{МП}}$	0,0000051	0,0001937	0,000445	0,0002335	0,0002129

Таблица 7

Результаты статистического эксперимента ($\lambda_1 = 2$, $\lambda_2 = 1{,}9$, $P_1(\lambda_1|\,\lambda_1) = 0{,}5$, $P_1(\lambda_2|\,\lambda_1) = 0{,}4$, $P_0(\lambda_2|\,\lambda_1) = 0{,}1$, $P_1(\lambda_2|\,\lambda_2) = 0{,}3$, $P_1(\lambda_1|\,\lambda_2) = 0{,}3$, $P_0(\lambda_1|\,\lambda_2) = 0{,}4$, $T = 1$)

n	200	400	600	800	1000
$\hat{V}_{\text{МП}}$	0,0001600	0,0000548	0,0000100	0,0000076	0,0000021
$\hat{V}_{\text{ММ}}$	0,0010568	0,0006839	0,0000605	0,0000392	0,0000201
$\hat{V}_{\text{ММ}} - \hat{V}_{\text{МП}}$	0,0008968	0,0006292	0,0000505	0,0000316	0,0000180

Таблица 8

Результаты статистического эксперимента ($\lambda_1 = 0{,}5$, $\lambda_2 = 0{,}1$, $P_1(\lambda_1|\,\lambda_1) = 0{,}7$, $P_1(\lambda_2|\,\lambda_1) = 0{,}2$, $P_0(\lambda_2|\,\lambda_1) = 0{,}1$, $P_1(\lambda_2|\,\lambda_2) = 0{,}6$, $P_1(\lambda_1|\,\lambda_2) = 0{,}3$, $P_0(\lambda_1|\,\lambda_2) = 0{,}1$, $T = 3$)

n	200	400	600	800	1000
$\hat{V}_{\text{МП}}$	0,0084623	0,0013386	0,0008519	0,0004733	0,0000847
$\hat{V}_{\text{ММ}}$	0,0021580	0,0015700	0,0019321	0,0013225	0,0014161
$\hat{V}_{\text{ММ}} - \hat{V}_{\text{МП}}$	–0,0063043	0,0002314	0,0010802	0,0008492	0,0013314

Анализ приведенных численных результатов показывает, что при малом времени (n = 10, 20, 30, 40, 50) наблюдения за потоком ММ-оценка лучше МП-оценки (табл. 1–4), что волне объяснимо, так как при малом времени наблюдения оценка $\hat{T}_{\text{МП}}$ может быть достаточно сильно смещенной относительно T. При большом времени наблюдения за потоком МП-оценка лучше ММ-оценки (табл. 4–8), что является естественным, так как при большом времени наблюдения смещение оценки $\hat{T}_{\text{МП}}$ относительно T уменьшается.

Заключение

Результаты проведенного исследования оценок длительности мертвого времени T, полученных методом моментов и методом максимального правдоподобия, указывают на то, что при малом времени наблюдения лучшую оценку дает метод моментов, при достаточно большом времени наблюдения лучшую оценку дает метод максимального правдоподобия. Границу применения той или иной оценки при заданных параметрах можно определить только путем имитационного моделирования.

ЛИТЕРАТУРА

1. *Gortsev A.M. and Solov'ev A.A.* Joint Probability Density of Interarrival Interval of a Flow of Physical Events with Unextendable Dead Time Period // Russian Physics Journal. 2014. V. 57, No. 7. P. 973–983.
2. *Gortsev A.M., Nezhelskaya L.A., Soloviev A.A.* Optimal State Estimation in MAP Event Flows with Unextendable Dead Time // Automation and Remote Control. 2012. V. 73, No. 8. P. 1316−1326.
3. *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.
4. *Gortsev A.M., Nezhel'skaya L.A.* Estimation of the dead time period and intensities of the synchronous double stochastic event flow // Radiotekhnika. 2004. No. 10. P. 8–16.
5. Vasi*l'eva 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. 64, No. 12. P. 1890–1898.
6. *Gortsev A.M., Nezhel'skaya L.A.* Estimation of the dead-time period and parameters of a semi-synchronous double-stochastic stream of events // Measurement Techniques. 2003. V. 46, No. 6. P. 536–545.
7. *Gortsev A.M., Parshina M.E.* Estimation of parameters of an alternate stream of events in "dead" time conditions // Russian Physics Journal. 1999. V. 42, No. 4. P. 373–378.
8. *Gortsev A.M., Klimov I.S.* Estimation of the non-observability period and intensity of Poisson event flow // Radiotekhnika. 1996. No. 2. P. 8–11.
9. *Gortsev A.M., Klimov I.S.* An estimate for intensity of Poisson flow of events under the condition of its partial missing // Radiotekhnika. 1991. No. 12. P. 3–7.
10. *Горцев А.М., Назаров А.А., Терпугов А.Ф.* Управление и адаптация в системах массового обслуживания. Томск : Изд-во Том. ун-та, 1978. 208 с.
11. *Bushlanov I.V., Gortsev A.V.* Optimal estimation of the states of a synchronous double stochastic flow of events // Automation and Remote Control. 2004. V. 65, No. 9. P. 1389–1399.
12. *Bushlanov I.V., Gortsev A.V.* Optimal estimation of the states of a synchronous double stochastic flow of events // Avtomatika i Telemekhanika. 2004. No. 9. P. 40–51.
13. *Gortsev A.M., Shmyrin I.S.* Optimal estimation of states of a double stochastic flow of events in the presence of measurement errors of time instants // Automation and Remote Control. 1999. V. 60, Part 1. P. 41–51.
14. *Gortsev A.M., Nezhel'skaya L.A., Shevchenko T.I.* Estimation of the states of an MC-stream of events in the presence of measurement errors // Russian Physics Journal. 1993. V. 36, No. 12. P. 1153–1167.
15. G*ortsev A.M., Nezhelskaya L.A.* An asynchronous double stochastic flow with initiation of superfluous events // Discrete Mathematics and Applications. 2011. V. 21, No. 3. 2011. P. 283–290.

16. *Bushlanov I.V., Gortsev A.M., Nezhel'skaya L.A.* Estimating parameters of the synchronous twofold-stochastic flow of events // Automation and Remote Control. 2008. V. 69, No. 9. P. 1517–1533.
17. *Vasil'eva L.A., Gortsev A.M.* Dead-time interval estimation of incompletely observable asynchronous bistochastic flow of events // Avtomatika i Telemekhanika. 2003. No. 12. P. 69–79.
18. *Vasil'eva L.A., Gortsev A.M.* Parameter estimation of a doubly stochastic flow of events under incomplete observability // Avtomatika i Telemekhanika. 2002. No. 3. P. 179–184.
19. *Vasil'eva 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.
20. *Gortsev A.M., Shmyrin I.S.* Optimal estimate of the parameters of a twice stochastic poisson stream of events with errors in measuring times the events occur // Russian Physics Journal. 1999. V. 42, No. 4. P. 385–393.
21. *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, No. 1. P. 56–63.
22. *Gortsev A.M., Nezhel'skaya L.A.* Estimation of the parameters of a synchro-alternating Poisson event flow by the method of moments // Radiotekhnika. 1995. V. 40, No. 7–8. P. 6–10.
23. *Gortsev A.M., Klimov I.S.* Estimation of the parameters of an alternating Poisson stream of events // Telecommunications and Radio Engineering (English translation of Elektrosvyaz and Radiotekhnika). 1993. V. 48, No. 10. P. 40–45.
24. *Gortsev A.M., Klimov I.S.* Estimation of intensity of Poisson stream of events for conditions under which it is partially unobservable // Telecommunications and Radio Engineering (English translation of Elektrosvyaz and Radiotekhnika). 1992. V. 47, No. 1. P. 33–38.
25. *Апанасович В.В., Коляда А.А., Чернявский А.Ф.* Статистический анализ случайных потоков в физическом эксперименте. Минск : Университетское, 1988. 254 с.
26. *Горцев А.М., Калягин А.А., Нежельская Л.А.* Оценка максимального правдоподобия длительности мёртвого времени в обобщённом полусинхронном потоке // Вестник ТГУ. Управление, вычислительная техника и информатика. 2015. № 1 (30). С. 27–37.
27. *Горцев А.М., Леонова М.А., Нежельская Л.А.* Сравнение МП- и ММ-оценок длительности мёртвого времени в обобщенном асинхронном потоке событий // Вестник ТГУ. Управление, вычислительная техника и информатика. 2013. № 4 (25). С. 32–42.
28. *Леонова М.А., Нежельская Л.А.* Оценка максимального правдоподобия длительности мертвого времени в обобщенном асинхронном потоке событий // Вестник Томского государственного университета. Управление, вычислительная техника и информатика. 2013. № 2(23). С. 54–63.
29. *Шуленин В.П.* Математическая статистика. Томск: Изд-во НТЛ, 2012. Ч. 1. 540 с.

Соловьев Александр Александрович. E-mail:sizal19@mail.ru
Горцев Александр Михайлович, д-р техн. наук, профессор. E-mail:gam@mail.fpmk.tsu.ru
Томский государственный университет

Поступила в редакцию 12 августа 2015 г.

Gortzev Alexandr M., Solovev Alexandr A. (Tomsk State University, Russian Federation).
Comparison ML- and MM- estimates fixed duration of the dead time MAP-flow of events
Keywords: MAP-flow fixed dead time; the method of maximum likelihood; maximum likelihood estimation; method of moments; the evaluation method of moments.

DOI: 10.17223/19988605/33/2

The Markovian arrival process (MAP-flow) with the intensity representing a piecewise-constant random process $\lambda(t)$ with two states is considered: $\lambda(t) = \lambda_1$ or $\lambda(t) = \lambda_2$ ($\lambda_1 > \lambda_2$). The dwell time of the process $\lambda(t)$ in the *i*-th state is random variable with exponential distribution function $F_i = 1 - e^{\lambda_i t}$, $i = 1, 2$. At the moment of termination of the *i*-th state of the process $\lambda(t)$ the following situations are possible, each of which proceeds instantly: 1) process $\lambda(t)$ transits from the *i*-th state to the *i*-th state, and an event of flow in the *i*-th state occurs; the joint probability of this situation is $P(\lambda_i \to \lambda_i, 1) = P_1(\lambda_i \mid \lambda_i)$ $i = 1, 2$; 2) process $\lambda(t)$ transits from the *i*-th state to the *j*-th state, and an event of the flow occurs; the joint probability of the situation is $P(\lambda_i \to \lambda_j, 1) = P_1(\lambda_j \mid \lambda_i)$, $i, j = 1, 2$; $i \neq j$; 3) process $\lambda(t)$ transits from the *i*-th state to the *j*-th state, and no event of the flow occurs; the joint probability of the situation is $P(\lambda_i \to \lambda_j, 0) = P_0(\lambda_j \mid \lambda_i)$, $i, j = 1,2$; $i \neq j$. Thus, $P_1(\lambda_i \mid \lambda_i) + P_1(\lambda_j \mid \lambda_i) + P_0(\lambda_j \mid \lambda_i) = 1$, $i, j = 1, 2$; $i \neq j$

$$\mathrm{D} = \left\| \begin{matrix} -\lambda_1 & \lambda_1 P_0(\lambda_2 \mid \lambda_1) \\ \lambda_2 P_0(\lambda_1 \mid \lambda_2) & -\lambda_2 \end{matrix} \middle| \begin{matrix} \lambda_1 P_1(\lambda_1 \mid \lambda_1) & \lambda_1 P_1(\lambda_2 \mid \lambda_1) \\ \lambda_2 P_1(\lambda_1 \mid \lambda_2) & \lambda_2 P_1(\lambda_2 \mid \lambda_2) \end{matrix} \right\| = \left\| \mathrm{D}_0 \mid \mathrm{D}_1 \right\|.$$

The elements of the matrix D_1 are intensities of transitions from state with occurrence of an event. Non-diagonal elements of the matrix D_0 are the intensities of transitions from state to state, but no event occurs. Diagonal elements of the matrix D_0 are the intensities of the exit of the process $\lambda(t)$ from the states taken with the opposite sign. Under these assumptions, $\lambda(t)$ is the hidden Markov process.

A comparison of the estimates of fixed dead time obtained (observations obtained as a result of a simulation model of the flow of events MAP) by the method of maximum likelihood and the method of moments.

REFERENCES

1. Gortsev, A.M. & Solov'ev, A.A. (2014) Joint Probability Density of Interarrival Interval of a Flow of Physical Events with Unextendable Dead Time Period. *Russian Physics Journal*. 57(7). pp. 973-983. DOI: 10.1007/s11182-014-0333-4
2. Gortsev, A.M., Nezhelskaya, L.A. & Soloviev, A.A. (2012) Optimal State Estimation in MAP Event Flows with Unextendable Dead Time. *Automation and Remote Control*. 73(8). pp. 1316-1326. DOI: 10.1134/S000511791208005X
3. Gortsev, A.M. & Nissenbaum, O.V. (2005) Estimation of the dead time period and parameters of an asynchronous alternative flow of events with unextendable dead time period. *Russian Physics Journal*. 48(10). pp. 1039-1054. DOI: 10.1007/s11182-006-0023-y
4. Gortsev, A.M. & Nezhel'skaya, L.A. (2004) Estimation of the dead time period and intensities of the synchronous double stochastic event flow. *Radiotekhnika*. 10. pp. 8-16.
5. Vasil'eva, L.A. & Gortsev, A.M. (2003) Estimation of the dead time of an asynchronous double stochastic flow of events under incomplete observability. *Automation and Remote Control*. 64(12). pp. 1890-1898. DOI: 10.1023/B:AURC.0000008427.99676.df
6. Gortsev, A.M. & Nezhel'skaya, L.A. (2003) Estimation of the dead-time period and parameters of a semi-synchronous double-stochastic stream of events. *Measurement Techniques*. 46(6). pp. 536-545. DOI: 10.1023/A:1025499509015
7. Gortsev, A.M. & Parshina, M.E. (1999) Estimation of parameters of an alternate stream of events in "dead" time conditions. *Russian Physics Journal*. 42(4). pp. 373-378.
8. Gortsev, A.M. & Klimov, I.S. (1996) Estimation of the non-observability period and intensity of Poisson event flow. *Radiotekhnika*. 2. pp. 8-11.
9. Gortsev, A.M. & Klimov, I.S. (1991) An estimate for intensity of Poisson flow of events under the condition of its partial missing. *Radiotekhnika*. 12. pp. 3-7.
10. Gortsev, A.M., Nazarov, A.A. & Terpugov, A.F. (1978) *Upravlenie i adaptatsiya v sistemakh massovogo obsluzhivaniya* [Control and adaptation in queueing systems]. Tomsk: Tomsk State University.
11. Bushlanov, I.V. & Gortsev, A.V. (2004) Optimal estimation of the states of a synchronous double stochastic flow of events. *Automation and Remote Control*. 65(9). pp. 1389-1399. DOI: 10.1023/B:AURC.0000041418.09187.63
12. Bushlanov, I.V. & Gortsev, A.V. (2004) Optimal estimation of the states of a synchronous double stochastic flow of events. *Avtomatika i Telemekhanika*. 9. pp. 40-51. DOI: 10.1023/B:AURC.0000041418.09187.63
13. Gortsev, A.M. & Shmyrin, I.S. (1999) Optimal estimation of states of a double stochastic flow of events in the presence of measurement errors of time instants. *Automation and Remote Control*. 60(1). pp. 41-51.
14. Gortsev, A.M., Nezhel'skaya, L.A. & Shevchenko, T.I. (1993) Estimation of the states of an MC-stream of events in the presence of measurement errors. *Russian Physics Journal*. 36(12). pp. 1153-1167.
15. Gortsev, A.M. & Nezhelskaya, L.A. (2011) An asynchronous double stochastic flow with initiation of superfluous events. *Discrete Mathematics and Applications*. 21(3). pp. 283-290. DOI: 10.4213/dm1141
16. Bushlanov, I.V., Gortsev, A.M. & Nezhel'skaya, L.A. (2008) Estimating parameters of the synchronous twofold-stochastic flow of events. *Automation and Remote Control*. 69 (9). pp. 1517-1533. DOI: 10.1134/S0005117908090075
17. Vasil'eva, L.A. & Gortsev, A.M. (2003) Dead-time interval estimation of incompletely observable asynchronous bistochastic flow of events. *Avtomatika i Telemekhanika*. 12. pp. 69-79.
18. Vasil'eva, L.A. & Gortsev, A.M. (2002) Parameter estimation of a doubly stochastic flow of events under incomplete observability. *Avtomatika i Telemekhanika*. 3. pp. 179-184.
19. Vasil'eva, L.A. & Gortsev, A.M. (2002) Estimation of parameters of a double-stochastic flow of events under conditions of its incomplete observability. *Automation and Remote Control*. 63(3). pp. 511-515.
20. Gortsev, A.M. & Shmyrin, I.S. (1999) Optimal estimate of the parameters of a twice stochastic poisson stream of events with errors in measuring times the events occur. *Russian Physics Journal*. 42(4). pp. 385-393. DOI: 10.1007/BF02509674
21. Gortsev, A.M. & Nezhel'skaya, L.A. (1996) Estimate of parameters of synchronously alternating Poisson stream of events by the moment method. *Telecommunications and Radio Engineering*. 50(1). pp. 56-63.
22. Gortsev, A.M. & Nezhel'skaya, L.A. (1995) Estimation of the parameters of a synchro-alternating Poisson event flow by the method of moments. *Radiotekhnika*. 40(7-8). pp. 6-10.
23. Gortsev, A.M. & Klimov, I.S. (1993) Estimation of the parameters of an alternating Poisson stream of events. *Telecommunications and Radio Engineering*. 48(10). pp. 40-45.
24. Gortsev, A.M. & Klimov, I.S.(1992) Estimation of intensity of Poisson stream of events for conditions under which it is partially unobservable. *Telecommunications and Radio Engineering*. 47(10). pp. 33-38.
25. Apanasovich, V.V., Kolyada, A.A. & Chernyavsky, A.F. (1988) *Statisticheskiy analiz sluchaynykh potokov v fizicheskom eksperimente* [Statistical analysis of stochastic flows in physical experiment]. Minsk: Universitetskoe.
26. Gortsev, A.M., Kalyagin, A.A. & Nezhelskaya, L.A. (2015) Maximum likelihood estimation of dead time at a generalized semysynchronous flow of events. *Vestnik Tomskogo gosudarstvennogo universiteta. Upravlenie, vychislitel'naya tekhnika i informatika – Tomsk State University Journal of Control and Computer Science*. 1(14). pp. 13-21. (In Russian).
27. Gortsev, A.M., Leonova, M.A., Nezhelskaya, L.A. (2013) The comparison of maximum likelihood estimation and method of moments estimation of dead time value in a generalized asynchronous flow of events. *Vestnik Tomskogo gosudarstvennogo universiteta. Upravlenie, vychislitel'naya tekhnika i informatika – Tomsk State University Journal of Control and Computer Science*. 4(25). pp. 32-42. (In Russian).
28. Leonova, M.A., Nezhelskaya, L.A. (2013) Maximum likelihood estimation of dead time value at a generalized asynchronous flow of events. *Vestnik Tomskogo gosudarstvennogo universiteta. Upravlenie, vychislitel'naya tekhnika i informatika – Tomsk State University Journal of Control and Computer Science*. 2(23). pp. 54-63. (In Russian).
29. Shulenin, V.P. (2012) *Matematicheskaya statistika* [Mathematical Statistics]. Part 1. Tomsk: NTL.

ВЕСТНИК ТОМСКОГО ГОСУДАРСТВЕННОГО УНИВЕРСИТЕТА

2015 Управление, вычислительная техника и информатика № 4 (33)

УДК 519.21
DOI: 10.17223/19988605/33/3

А.Р. Ерёмина, Ю.В. Малинковский

ИНВАРИАНТНОСТЬ СТАЦИОНАРНОГО РАСПРЕДЕЛЕНИЯ СЕТЕЙ МАССОВОГО ОБСЛУЖИВАНИЯ С МНОГОРЕЖИМНЫМИ СТРАТЕГИЯМИ И ОТРИЦАТЕЛЬНЫМИ ЗАЯВКАМИ

Рассматривается открытая сеть массового обслуживания с многорежимными стратегиями, положительными и отрицательными заявками нескольких типов. Время обслуживания заявок имеет показательное распределение, количество работы по переключению режима прибора в узле – произвольное распределение. Устанавливается инвариантность стационарного распределения вероятностей состояний сети относительно функциональной формы распределений величин работ, требующихся на переключение режимов работы приборов.

Ключевые слова: сеть массового обслуживания; инвариантность; многорежимное обслуживание; отрицательные заявки; LCFS PR.

Сети массового обслуживания используются в качестве адекватных моделей производственных и транспортно-логистических сетей, сетей связи и передачи данных, информационных и компьютерных сетей.

Сети джексоновского типа с отрицательными заявками (так называемые «*G*-сети») были впервые рассмотрены Э. Геленбе в [1, 2]. В них наряду с обычными, «положительными» заявками в сеть поступают «отрицательные» заявки, которые не требуют обслуживания и ведут себя как сигналы (отрицательная заявка уменьшает количество положительных заявок в узле на одну и не оказывает никакого влияния на узел, если в нем нет положительных заявок). Э. Геленбе установил, что для таких сетей стационарное распределение также имеет форму произведения, а уравнение трафика (нелинейное) имеет единственное положительное решение. После этого многими авторами рассматривались различные обобщения *G*-сетей. В частности, в [3–5] исследовались сети с многорежимными стратегиями, отрицательными заявками и информационными сигналами. В этих работах полагалось, что длительности обслуживания заявок и времена пребывания приборов в режимах имеют показательные распределения. Однако в сетевых моделях, описывающих реальные объекты в экономике, финансах, технике и т.п., указанные распределения чаще всего отличаются от показательного. Так, например, любые технические средства в силу естественного износа или нарушения условий эксплуатации могут полностью прекращать функционирование, требовать ремонта (замены) или продолжать работать с меньшей производительностью. Поэтому большую важность для практических приложений представляет изучение сетей массового обслуживания с многорежимными стратегиями, в которых обслуживающие приборы в узлах могут работать в нескольких режимах, соответствующих различной производительности. Подобные сети позволяют моделировать ситуации, когда обслуживающие приборы частично надёжны, что особенно актуально для реальных сетей, в которых любые технические средства в силу физического износа, нарушения условий эксплуатации и т.д. могут выходить из строя полностью или частично, работая при этом с различной производительностью. Стандартный случайный процесс, описывающий такие сети, не является марковским, что затрудняет их исследование и нахождение стационарного распределения.

В [6] была исследована сеть, в которой некоторые узлы доступны для отрицательных заявок и длительности обслуживания имеют показательное распределение, а оставшиеся узлы недоступны для отрицательных заявок и длительности обслуживания имеют произвольную функцию распределения. Была доказана инвариантность стационарного распределения вероятностей состояний указанной сети по отношению к функциям распределения длительностей обслуживания заявок в узлах, недоступных для отрицательных заявок.

В [7] рассматривалась сеть с многорежимным обслуживанием и отрицательными заявками одного типа, когда обслуживание имело временную интерпретацию (т.е. скорость обслуживания равна 1). Была установлена инвариантность стационарного распределения относительно функций распределения величин работ, требующихся для переключения режимов функционирования приборов в узлах при фиксированных математических ожиданиях.

В данной статье результат, полученный в [7], распространен на случай, когда в сеть с многорежимным обслуживанием поступают положительные и отрицательные заявки различных типов.

1. Описание сети

Рассматривается открытая сеть массового обслуживания, состоящая из N однолинейных узлов, в которую поступают два независимых простейших потока положительных и отрицательных заявок с интенсивностями λ^+ и λ^- соответственно. Каждая заявка входного потока положительных заявок независимо от других заявок направляется в l-й узел и становится заявкой типа u, $u = \overline{1, M}$, с вероятностью $p^+_{0(l,u)}$ ($\sum_{l=1}^{N}\sum_{u=1}^{M} p^+_{0(l,u)} = 1$). Каждая заявка входного потока отрицательных заявок независимо от других заявок направляется в l-й узел и становится заявкой типа u с вероятностью $p^-_{0(l,u)}$ ($\sum_{l=1}^{N}\sum_{u=1}^{M} p^-_{0(l,u)} = 1$). Отрицательная заявка типа u, поступающая в узел сети, обслуживания не требует. Она уменьшает число заявок типа u в этом узле на единицу, если в очереди данного узла есть заявки типа u, и не оказывает никакого влияния на состояние узла в противном случае.

После обслуживания в l-м узле положительная заявка типа u независимо от других заявок мгновенно с вероятностью $p^+_{(l,u)(k,v)}$ направляется в k-й узел как положительная заявка типа v, с вероятностью $p^-_{(l,u)(k,v)}$ направляется в k-й узел как отрицательная заявка типа v, а с вероятностью $p_{(l,u)0}$ покидает сеть ($\sum_{k=1}^{N}\sum_{v=1}^{M}(p^+_{(l,u)(k,v)} + p^-_{(l,u)(k,v)}) + p_{(l,u)0} = 1; l,k = \overline{1,N}; u,v = \overline{1,M}$).

Заявки в узлах обслуживаются согласно дисциплине LCFS PR (заявка, поступающая в l-й узел, вытесняет заявку с прибора и начинает обслуживаться, а вытесненная заявка становится первой в очереди на обслуживание). Таким образом, поступающие в узел заявки имеют абсолютный приоритет. Время обслуживания заявки типа u, находящейся в l-м узле, имеет показательное распределение с параметром μ_{lu}.

В каждом из N узлов находится единственный прибор, который может работать в $r_l + 1$ режимах $0, 1, ..., r_l$, $l = \overline{1, N}$.

Состояние сети в момент времени t характеризуется вектором $x(t) = (x_1(t), x_2(t), ..., x_N(t))$, где состояние l-го узла в момент времени t описывается вектором $x_l(t) = (\overline{x}_l(t), j_l(t)) = (x_{l1}(t), x_{l2}(t), ..., x_{l,n(l)}(t), j_l(t))$; $x_{l1}(t)$ – тип заявки, стоящей последней в очереди на обслуживание в l-м узле в момент времени t; $x_{l2}(t)$ – тип заявки, стоящей предпоследней в очереди на обслуживание в l-м узле в момент времени t и т.д.; $x_{l,n(l)-1}(t)$ – тип заявки, стоящей первой в очереди на облуживание в l-м узле в момент времени t; $x_{l,n(l)}(t)$ – тип заявки, находящейся на облуживании в l-м узле в момент времени t; $j_l(t)$ – режим, в котором работает l-й узел в момент времени t. Процесс $x(t)$ обладает пространством состояний

$$X = X_1 \times X_2 \times ... \times X_N,$$

где $X_l = \{(0, j_l), (x_{l1}, j_l), (x_{l1}, x_{l2}, j_l), (x_{l1}, x_{l2}, x_{l3}, j_l), ...: x_{lk} = \overline{1, M}, k = 1,2,...; j_l = \overline{0, r_l}\}$.

В качестве основного режима работы обслуживающего прибора полагается режим работы 0. Переключение происходит только на соседние режимы. Во время переключения прибора с одного режима на другой число заявок в узле не меняется.

Количество работы, необходимое для переключения прибора l-го узла из основного режима работы в режим 1, является случайной величиной с произвольной функцией распределения $\Phi_l(0,\tilde{u})$ и математическим ожиданием $\eta_l(0)$. При этом если в момент времени t в узле находится $n(l)$ заявок (состояние узла $(\bar{x}_l, j_l)$), то указанная работа выполняется со скоростью $\nu_l(\bar{x}_l,0)$.

Для состояний x_l, у которых $1 \le j_l \le r_l - 1$, количество работы, необходимое для изменения режима j_l, также является случайной величиной с произвольной функцией распределения $\Phi_l(j_l,\tilde{u})$ и математическим ожиданием $\eta_l(j_l)$.

Если в момент времени t состояние узла $(\bar{x}_l, j_l)$, то выполнение работы происходит со скоростью $\nu_l(x_l) + \varphi_l(x_l)$, при этом с вероятностью $\dfrac{\nu_l(x_l)}{\nu_l(x_l) + \varphi_l(x_l)}$ прибор l-го узла переходит в режим $j_l + 1$, а с вероятностью $\dfrac{\varphi_l(x_l)}{\nu_l(x_l) + \varphi_l(x_l)}$ прибор l-го узла переходит в режим $j_l - 1$.

Аналогично количество работы, необходимое для выхода прибора l-го узла из режима работы r_l в $r_l - 1$, имеет произвольную функцию распределения $\Phi_l(r_l,\tilde{u})$ и математическое ожидание $\eta_l(r_l)$. При этом если в момент времени t состояние узла $(\bar{x}_l, j_l)$, то указанная работа выполняется со скоростью $\varphi_l(\bar{x}_l, r_l)$.

Математические ожидания всех вышеперечисленных случайных величин конечны, т.е. $\eta_l(j_l) < +\infty,\ j_l = \overline{0,\ r_l},\ l = \overline{1, N}.$

Для сокращения выкладок вводятся операторы:

$$T_u^+(\bar{x}_l) = T_u^+(x_{l1},...,x_{l,n(l)}) = (x_{l1},...,x_{l,n(l)},u),$$

$$T^-(\bar{x}_l) = T^-(x_{l1},...,x_{l,n(l)}) = (x_{l1},...,x_{l,n(l)-1}),$$

$$T_{(l,u)}^+(x) = T_{(l,u)}^+(x_1,...,x_N) = (\tilde{x}_1,...,\tilde{x}_N),\ \text{где}\ \tilde{x}_k = x_k\ \text{при}\ k \ne l,\ \tilde{x}_l = (T_u^+(\bar{x}_l), j_l),$$

$$T_l^-(x) = T_l^-(x_1,...,x_N) = (\tilde{x}_1,...,\tilde{x}_N),\ \text{где}\ \tilde{x}_k = x_k\ \text{при}\ k \ne l,\ \tilde{x}_l = (T^-(\bar{x}_l), j_l),$$

$$R_l^{j_l+1}(x) = R_l^{j_l+1}(x_1,...,x_N) = (\tilde{x}_1,...,\tilde{x}_N),\ \text{где}\ \tilde{x}_k = x_k\ \text{при}\ k \ne l,\ \tilde{x}_l = (\bar{x}_l, j_l + 1),$$

$$R_l^{j_l-1}(x) = R_l^{j_l-1}(x_1,...,x_N) = (\tilde{x}_1,...,\tilde{x}_N),\ \text{где}\ \tilde{x}_k = x_k\ \text{при}\ k \ne l,\ \tilde{x}_l = (\bar{x}_l, j_l - 1).$$

Полагается, что все величины μ_l, $\nu_l(x_l)$, $\varphi_l(x_l)$ строго больше 0, а матрица $\left(p_{(l,u)(k,v)}^+\right)$, $u, v = \overline{1, M}$, $l, k = \overline{0, N}$, неприводима.

Обозначим через $\alpha_{(l,u)}^+$ среднюю интенсивность поступления положительных заявок типа u в l-й узел, а через $\alpha_{(l,u)}^-$ – среднюю интенсивность поступления отрицательных заявок типа u в l-й узел, $u = \overline{1, M}$.

Рассмотрим систему нелинейных относительно $\alpha_{(l,u)}^+$ и $\alpha_{(l,u)}^-$ уравнений:

$$\alpha_{(l,u)}^+ = \lambda^+ p_{0(l,u)}^+ + \sum_{(k,v)} \frac{\alpha_{(k,v)}^+ \mu_{(k,v)}}{\mu_{(k,v)} + \alpha_{(k,v)}^-} p_{(k,v)(l,u)}^+, \tag{1}$$

$$\alpha_{(l,u)}^- = \lambda^- p_{0(l,u)}^- + \sum_{(k,v)} \frac{\alpha_{(k,v)}^+ \mu_{(k,v)}}{\mu_{(k,v)} + \alpha_{(k,v)}^-} p_{(k,v)(l,u)}^-, \tag{2}$$

$$l,\ k = \overline{1, N};\ u,\ v = \overline{1, M}.$$

Система уравнений (1)–(2) имеет решение

$$\left(\alpha^{+}_{(1,u)},\alpha^{+}_{(2,u)},...,\alpha^{+}_{(N,u)};\alpha^{-}_{(1,u)},\alpha^{-}_{(2,u)},...,\alpha^{-}_{(N,u)}\right),\ u=\overline{1,M}.$$

Полагается, что оно существует, причём $\alpha^{+}_{(1,u)}>0, l=\overline{1,N}.$

Через $\psi_{lj_l(t)}(t)$ обозначается количество работы, которое осталось выполнить с момента t для переключения режима обслуживания на соседний режим в l-м узле, если обслуживающий прибор работает в режиме j_l, $\psi(t)=\left(\psi_{1,j_1(t)}(t),\psi_{2,j_2(t)}(t),...,\psi_{N,j_N(t)}(t)\right)$.

В силу вышесказанного

$$\frac{d\psi_{l,j_l(t)}(t)}{dt}=-\left(\nu_l(x_l)I_{(j_l\neq r_l)}+\varphi_l(x_l)I_{(j_l\neq 0)}\right).$$

В общем случае процесс $x(t)$ не является марковским, поэтому рассматривается кусочно-линейный марковский процесс $\zeta(t)=\left(x(t),\psi(t)\right)$, полученный добавлением к $x(t)$ непрерывной компоненты $\psi(t)$.

Под $P=\{P(x)\}$ понимается стационарное распределение вероятностей состояний процесса $x(t)$, $P(x)=\lim\limits_{t\to\infty}P\{x(t)=x\}$.

Функции

$$F(x,z)=F(x,z_{1,j_1},z_{2,j_2},...,z_{N,j_N})=$$

$$=\lim_{t\to\infty}P\left\{x(t)=x;\psi_{1,j_1}(t)<z_{1,j_1},\psi_{2,j_2}(t)<z_{2,j_2};...;\psi_{N,j_N}(t)<z_{N,j_N}\right\}$$

называются стационарными функциями распределения вероятностей состояний кусочно-линейного марковского процесса $\zeta(t)$.

2. Марковский случай

Пусть длительности пребывания в режимах имеют показательное распределение, т.е. для l-го узла $\Phi_l(x_l,\tilde{u})=1-\exp\{-(\nu_l(\bar{x}_l,j_l)+\varphi_l(\bar{x}_l,j_l))\tilde{u}\}$, $(\tilde{u}>0)$. Тогда $x(t)$ – однородный марковский процесс с непрерывным временем и не более чем счётным фазовым пространством состояний $X=X_1\times X_2\times...\times X_N$, где $X_l=\{(0,j_l),\ (x_{l1},j_l),\ (x_{l1},x_{l2},j_l),\ ...: x_{lk}=\overline{1,M}, k=1,2,...;\ j_l=\overline{0,r_l}\}$.

В [8] установлено, что при выполнении условий

$$\nu_l(\bar{x}_l,j_l-1)\phi_l(T^{-}(\bar{x}_l),j_l)=\nu_l(T^{-}(\bar{x}_l),j_l-1)\varphi_l(\bar{x}_l,j_l), \tag{3}$$

$$\sum_{x\in X}q_x\prod_{l=1}^{N}\left[\prod_{s=1}^{n(l)}\left(\frac{\alpha^{+}_{(l,x_{ls})}}{\mu_{(l,x_{ls})}+\alpha^{-}_{(l,x_{ls})}}\right)\prod_{k=1}^{j_l}\frac{\nu_l(\bar{x}_l,k-1)}{\varphi_l(\bar{x}_l,k)}\right]<+\infty, \tag{4}$$

$$q_x=\lambda^{+}+\lambda^{-}+\sum_{l=1}^{N}\sum_{u=1}^{M}\mu_{(l,u)}+\sum_{l=1}^{n}\left[\nu_l(\bar{x}_l,j_l)+\varphi_l(\bar{x}_l,j_l)\right]$$

процесс $x(t)$ эргодичен. При этом финальное стационарное распределение имеет мультипликативную форму, в которой множители характеризуют отдельные узлы

$$P(x)=p_1(x_1)p_2(x_2)\times...\times p_N(x_N),$$

где

$$p_l(\bar{x}_l,j_l)=\prod_{s=1}^{n(l)}\left(\frac{\alpha^{+}_{(l,x_{ls})}}{\mu_{(l,x_{ls})}+\alpha^{-}_{(l,x_{ls})}}\right)\prod_{k=1}^{j_l}\frac{\nu_l(\bar{x}_l,k-1)}{\varphi_l(\bar{x}_l,k)}p_l(0,0),\ (\bar{x}_l,j_l)\in X_i, \tag{5}$$

а $\left(\alpha^{+}_{(l,u)},\alpha^{-}_{(l,u)}\right)$, $l=\overline{1,N}$, – положительное решение уравнений трафика (1)–(2).

3. Основной результат

Пусть теперь количество работы, необходимое для переключения режима прибора l-го узла на один из соседних режимов, имеет произвольную функцию распределения $\Phi_l(j_l, \tilde{u})$, когда состояние l-го узла есть $(\bar{x}_l, j_l)$. Тогда имеет место следующая теорема.

Теорема. Если выполнены условия (3)–(4), то процесс $\zeta(t)$ эргодичен, при этом стационарные функции распределения вероятностей состояний $F(x,z)$ определяются по формулам

$$F(x,z) = \prod_{l=1}^{N} p_l(x_l)\eta_l^{-1}(j_l) \int_0^{z_{l,j_l}} (1-\Phi_l(j_l,\tilde{u}))d\tilde{u},\ x \in X, \tag{6}$$

где

$$p_l(\bar{x}_l, j_l) = \prod_{s=1}^{n(l)} \left(\frac{\alpha^+_{(l,x_{ls})}}{\mu_{(l,x_{ls})} + \alpha^-_{(l,x_{ls})}} \right) \frac{\eta(j_l)}{\eta(0)} \prod_{k=1}^{j_l} \frac{\nu_l(0,k-1)}{\varphi_l(0,k)} p_l(0,0), \tag{7}$$

$$p_l(0,0) = \eta(0) \left(1 - \prod_{s=1}^{n(l)} \frac{\alpha^+_{(l,x_{ls})}}{\mu_{(l,x_{ls})} + \alpha^-_{(l,x_{ls})}} \right) \left(\sum_{j_l=0}^{r_l} \prod_{k=1}^{j_l} \frac{\nu_l(0,k-1)}{\varphi_l(0_l,k)} \right)^{-1}, \tag{8}$$

а $\left(\alpha^+_{(l,u)}, \alpha^-_{(l,u)}\right)$, $l = \overline{1,N}$, – положительное решение уравнений трафика (1)–(2).

Доказательство. Пусть выполнены условия (3)–(4), т.е. в случае, когда $x(t)$ – марковский процесс, существует стационарное эргодическое распределение $x(t)$, тогда, по-видимому, и в общем случае при выполнении условий (3)–(4) существует стационарное эргодическое распределение $\zeta(t)$, так как $\zeta(t)$ получается из $x(t)$ добавлением непрерывных компонент. Строгое доказательство этого факта может быть проведено, если учесть, что процесс $\zeta(t)$ является регенерирующим. Функционирование сети схематично можно представить как чередование периодов, когда сеть находится в состоянии «0» (в каждом узле сети нет заявок, прибор работает в нулевом режиме), и периодов занятости сети (в противном случае). Далее доказательство сводится к применению предельной теоремы Смита для регенерирующих процессов [9. С. 41], при этом учитывается, что среднее время обслуживания заявки равно среднему времени обслуживания заявки в марковском случае.

Обозначим $\vartheta_l(\bar{x}_l, j_l) = \nu_l(x_l) I_{(j_l \neq r_l)} + \varphi_l(x_l) I_{(j_l \neq 0)}$, $n(l) > 0$, $j_l = \overline{0, r_l}$, $l = \overline{1,N}$.

Для $F(x,z)$ справедлива следующая система разностно-дифференциальных уравнений:

$$\left(\lambda^+ + \lambda^- + \sum_{l=1}^{N} \sum_{u=1}^{M} \mu_{lu} \right) F(x,z) = \tag{9}$$

$$= \sum_{l=1}^{N} \vartheta_l(\bar{x}_l, j_l) \left(\frac{\partial F(x,z)}{\partial z_{l,j_l}} - \left(\frac{\partial F(x,z)}{\partial z_{l,j_l}} \right)_{z_{l,j_l}=0} \right) + \sum_{l=1}^{N} \sum_{u=1}^{M} \mu_{lu} p_{(l,u)0} F(T^+_{(l,u)}(x), z) +$$

$$+ \lambda^+ \sum_{l=1}^{N} p^+_{0(l,x_{l,n(l)})} F(T^-_l(x), z) + \lambda^- \sum_{l=1}^{N} \sum_{u=1}^{M} p^-_{0(l,u)} F(T^+_{(l,u)}(x), z) +$$

$$+ \sum_{l=1}^{N} \sum_{s=1, s\neq l}^{N} \sum_{u=1}^{M} \mu_{su} p^+_{(s,u)(l,x_{l,n(l)})} F(T^+_{(s,u)}(T^-_l(x)), z) +$$

$$+ \sum_{l=1}^{N} \sum_{u=1}^{M} \mu_{lu} p^+_{(l,u)(l,x_{l,n(l)})} F(T^+_{(l,u)}(T^-_l(x)), z) +$$

$$+ \sum_{l=1}^{N} \sum_{s=1, s\neq l}^{N} \sum_{u=1}^{M} \sum_{v=1}^{M} \mu_{su} p^-_{(s,u)(l,v)} F(T^+_{(s,u)}(T^+_{(l,v)}(x)), z) +$$

$$+\sum_{l=1}^{N}\sum_{u=1}^{M}\sum_{v=1}^{M}\mu_{lu}\,p^{-}_{(l,u)(l,v)}F(T^{+}_{(l,u)}(T^{+}_{(l,v)}(x)),z)+$$

$$+\sum_{l=1}^{N}v_l(\bar{x}_l,j_l-1)\Phi_l\left(j_l,z_{l,j_l}\right)\left(\frac{\partial F(R_l^{j_l-1}(x),z)}{\partial z_{l,j_l-1}}\right)_{z_{l,j_l-1}=0}+$$

$$+\sum_{l=1}^{N}\varphi_l(\bar{x}_l,j_l+1)\Phi_l\left(j_l,z_{l,j_l}\right)\left(\frac{\partial F(R_l^{j_l+1}(x),z)}{\partial z_{l,j_l+1}}\right)_{z_{l,j_l+1}=0},\ x\in X.$$

Для данных уравнений предполагается, что если аргумент функции $F(x,z)$ не принадлежит фазовому пространству, т.е. если $x\notin X$, то $F(x,z)=0$.

Полученная система разбивается на уравнения локального баланса следующим образом:

$$\left(\lambda^{+}+\lambda^{-}+\sum_{l=1}^{N}\sum_{u=1}^{M}\mu_{lu}\right)F(x,z)= \tag{10}$$

$$=\sum_{l=1}^{N}\sum_{u=1}^{M}\mu_{lu}\,p_{(l,u)0}F(T^{+}_{(l,u)}(x),z)+$$

$$+\lambda^{+}\sum_{l=1}^{N}p^{+}_{0(l,x_{l,n(l)})}F(T_l^{-}(x),z)+\lambda^{-}\sum_{l=1}^{N}\sum_{u=1}^{M}p^{-}_{0(l,u)}F(T^{+}_{(l,u)}(x),z)+$$

$$+\sum_{l=1}^{N}\sum_{s=1,s\neq l}^{N}\sum_{u=1}^{M}\mu_{su}\,p^{+}_{(s,u)(l,x_{l,n(l)})}F(T^{+}_{(s,u)}(T_l^{-}(x)),z)+$$

$$+\sum_{l=1}^{N}\sum_{u=1}^{M}\mu_{lu}\,p^{+}_{(l,u)(l,x_{l,n(l)})}F(T^{+}_{(l,u)}(T_l^{-}(x)),z)+$$

$$+\sum_{l=1}^{N}\sum_{s=1,s\neq l}^{N}\sum_{u=1}^{M}\sum_{v=1}^{M}\mu_{su}\,p^{-}_{(s,u)(l,v)}F(T^{+}_{(s,u)}(T^{+}_{(l,v)}(x)),z)+$$

$$+\sum_{l=1}^{N}\sum_{u=1}^{M}\sum_{v=1}^{M}\mu_{lu}\,p^{-}_{(l,u)(l,v)}F(T^{+}_{(l,u)}(T^{+}_{(l,v)}(x)),z),$$

$$\vartheta_l(\bar{x}_l,j_l)\left(\frac{\partial F(x,z)}{\partial z_{l,j_l}}-\left(\frac{\partial F(x,z)}{\partial z_{l,j_l}}\right)_{z_{l,j_l}=0}\right)= \tag{11}$$

$$=v_l(\bar{x}_l,j_l-1)\Phi_l\left(j_l,z_{l,j_l}\right)\left(\frac{\partial F(R_l^{j_l-1}(x),z)}{\partial z_{l,j_l-1}}\right)_{z_{l,j_l-1}=0}+$$

$$+\varphi_l(\bar{x}_l,j_l+1)\Phi_l\left(j_l,z_{l,j_l}\right)\left(\frac{\partial F(R_l^{j_l+1}(x),z)}{\partial z_{l,j_l+1}}\right)_{z_{l,j_l+1}=0}.$$

Функции распределения вероятностей $F(x,z)$, определённые формулами (6)–(8), являются решениями уравнений (10)–(11), а следовательно, и уравнений (9).

Действительно, подставим (6) в (10), приведём подобные слагаемые и, учитывая уравнения трафика (1)–(2), получим тождество. Подставляя (6) в (11) и учитывая (3), также получим тождество.

Теорема доказана.

Из теоремы с учётом равенства $P(x)=F(x,+\infty)$ имеем:

Следствие. Если выполняются соотношения (3)–(4), то процесс $x(t)$ эргодичен, а его стационарное распределение $\{P(x),x\in X\}$ не зависит от функционального вида распределения $\Phi_l\left(j_l,\tilde{u}\right)$ и имеет вид

$$P(x)=p_1(x_1)p_2(x_2)\times\ldots\times p_N(x_N),$$

где $p_l(x_l)$ определяются по формулам (7)–(8).

Заключение

В настоящей работе установлены условия независимости стационарного распределения вероятностей состояний открытых сетей с многорежимными стратегиями обслуживания, положительными и отрицательными заявками разных типов от вида законов распределения величин работ, требующихся для переключения режимов функционирования приборов в узлах, когда дисциплиной обслуживания является LCFS PR (абсолютный приоритет поступающего требования с дообслуживанием). При этом установлено, что стационарное распределение сети имеет форму произведения.

ЛИТЕРАТУРА

1. *Gelenbe E.* Random Neural Networks with Negative and Positive Signals and Product Form Solution // Neural Computation. 1989. V. 1. P. 502–510.
2. *Gelenbe E., Muntz R.R.* Probabilistic Models of Computer Systems. Part I: Exact Results // Acta Inform. 1976. № 7. P. 35–60.
3. *Малинковский Ю.В., Нуеман А.Ю.* Инвариантная мера марковского сетевого процесса с многорежимными стратегиями // Известия Гомельского государственного университета имени Ф. Скорины. 2002. № 6(15). С. 183–188.
4. *Нуеман А.Ю.* Открытые сети с многорежимными стратегиями обслуживания и отрицательными заявками // Вестник Томского государственного университета. 2002. № 1 (1). С. 90–93.
5. *Нуеман А.Ю.* Сети массового обслуживания с ненадежными приборами и отрицательными заявками // Новые математические методы и компьютерные технологии в проектировании, производстве и научных исследованиях : материалы V Респ. науч. конф. студентов и аспирантов. Гомель, 18–20 марта 2002 г. / Гомельский гос. ун-т им. Ф. Скорины; редкол.: Д.Г. Лин [и др.]. Гомель, 2002. С. 179–180.
6. *Бочаров П.П., Печинкин А.В.* Теория массового обслуживания : учеб. М. : РУДН, 1995. 529 с.
7. *Старовойтов А.Н.* Сети с многорежимным обслуживанием, отрицательными заявками и произвольным временем пребывания в режимах // Известия Гомельского государственного университета им. Ф. Скорины. 2007. № 6(45). С. 193–198.
8. *Летунович Ю.Е.* Стационарное распределение состояний открытой неоднородной сети с многорежимными стратегиями и немедленным обслуживанием // Современные информационные компьютерные технологии : сб. науч. ст. : в 2 ч. / ГрГУ им. Я. Купалы. Гродно, 2008. С. 97–99. Ч. 2.
9. *Ивницкий В.А.* Теория сетей массового обслуживания. М. : Физматлит, 2004. 772 с.

Ерёмина Александра Рафаэловна, канд. физ.-мат. наук. E-mail: a.eremina@grsu.by
Малинковский Юрий Владимирович, д-р физ.-мат. наук, профессор. E-mail: Malinkovsky@gsu.by
Гомельский государственный университет имени Франциска Скорины,
Республика Беларусь

Поступила в редакцию 23 сентября 2015 г.

Eryomina Alexandra R., Malinkovskiy Yury V. (Yanka Kupala State University of Grodno, Francisk Skorina Gomel State University, Republic of Belarus).

Invariance of the stationary distribution of queuing networks with multimode strategies and negative demands.

Keywords: queueing network; invariance; multimode strategies; LCFS PR.

DOI: 10.17223/19988605/33/3

The open queueing network with two independent simplest incoming flows of positive and negative demands with λ^+ and λ^- corresponding intensities is considered in article.

Every demand of incoming flow of positive demands goes to *l*-unit independently of other demands and becomes the *u*-type demand, $u = \overline{1,M}$, with $p^+_{0(l,u)}$ probability ($\sum_{l=1}^{N}\sum_{u=1}^{M} p^+_{0(l,u)} = 1$). Every demand of incoming flow of negative demands goes to *l*-unit independently of other demands and becomes the *u*-type demand with $p^-_{0(l,u)}$ probability ($\sum_{l=1}^{N}\sum_{u=1}^{M} p^-_{0(l,u)} = 1$). Negative *u*-type demand, entering in *l*-unit, does not require the service. It reduces the number of *u*-type demands in this unit by one if the queue of the unit has *u*-type demands, and has no effect on the state of the unit otherwise.

After serving in *l*-unit *u*-type positive demand goes to *k*-unit instantly and independently of other demands as *v*-type positive demand with $p^+_{(l,u)(k,v)}$ probability and as *v*-type negative demand with $p^-_{(l,u)(k,v)}$ probability, and with $p_{(l,u)0}$ probability this *u*-type positive demand goes away from the network ($\sum_{k=1}^{N}\sum_{v=1}^{M}(p^+_{(l,u)(k,v)} + p^-_{(l,u)(k,v)}) + p_{(l,u)0} = 1; l,k = \overline{1,N}; u,v = \overline{1,M}$).

The single line units can work in some strategies. Every strategy responds to different degree of serviceability. The dispatching rule of demands by device of *l*-unit is LCFS PR $(l = \overline{1,N})$.

The network condition at moment t is characterized by vector $x(t) = (x_1(t), x_2(t), ..., x_N(t))$, where condition of l-unit at moment t is vector $x_l(t) = (\overline{x}_l(t), j_l(t)) = (x_{l1}(t), x_{l2}(t), ..., x_{l,n(l)}(t), j_l(t))$, $x_{l1}(t)$ is the type of demand, which stands on the final position in l-unit at moment t, $x_{l2}(t)$ is the type of demand, which stands on the penultimate position in l-unit at moment t, $x_{l,n(l)-1}(t)$ is the type of demand, which stands on the first position in l-unit at moment t, $x_{l,n(l)}(t)$ is the type of demand, which is servicing by device in l-unit at moment t, $n(l)$ is total amount of demands in l-unit, $j_l(t)$ is the number of device strategy in l-unit at moment t, $l = \overline{1, N}$.

Servicing time of u-type demand, located in l-unit at moment t, has the exponential distribution with μ_{lu} parameter.

The basic device strategy is strategy 0. Switching occurs only on the neighboring strategies. During switching the device from one strategy to another one the number of demands in the unit does not change.

For x_l conditions, where $0 \le j_l \le r_l$, the quantity of work, which is necessary for switching of device work strategy j_l, is random variable with arbitrary distribution function $\Phi_l(j_l, \tilde{u})$ and expectation value $\eta_l(j_l)$.

The piecewise-linear Markovian process $\zeta(t) = \big(x(t), \psi(t)\big)$ is considered. This process is obtained by addition to $x(t)$ continuous component $\psi(t) = \big(\psi_{1,j_1(t)}(t), \psi_{2,j_2(t)}(t), ..., \psi_{N,j_N(t)}(t)\big)$, where $\psi_{lj_l(t)}(t)$ is quantity of work, which lefts to execute from moment t for switching of device work strategy to neighboring strategy in l-unit, if device works with j_l strategy.

Suppose that $P = \{P(x)\}$ is stationary distribution of state probabilities of $x(t)$ process; $F(x, z)$ are stationary distribution functions of state probabilities of piecewise-linear Markovian process $\zeta(t)$:

$$F(x, z) = F(x, z_{1,j_1}, z_{2,j_2}, ..., z_{N,j_N}) =$$

$$= \lim_{t \to \infty} P\left\{x(t) = x; \psi_{1,j_1}(t) < z_{1,j_1}, \psi_{2,j_2}(t) < z_{2,j_2}; ...; \psi_{N,j_N}(t) < z_{N,j_N}\right\}.$$

It was proved that stationary distribution of network state probabilities is invariant in relation to functional form of distributions of work's quantities, which are necessary for switching of device work strategies.

REFERENCES

1. Gelenbe, E. (1989) Random Neural Networks with Negative and Positive Signals and Product Form Solution. *Neural Computation*. 1. pp. 502-510.
2. Gelenbe, E. & Muntz, R.R. (1976) Probabilistic Models of Computer Systems. Part I: Exact Results. *Acta Inform*. 7. pp. 35-60.
3. Malinkovskiy, Yu.V. & Nueman, A.Yu. (2002) Invariantnaya mera markovskogo setevogo protsessa s mnogorezhimnymi strategiyami [The invariant measure of the Markov network process with multimode strategies]. *Izvestiya Gomel'skogo gosudarstvennogo universiteta imeni F. Skoriny*. 6(15). pp. 183-188.
4. Nueman, A.Yu. (2002) Otkrytye seti s mnogorezhimnymi strategiyami obsluzhivaniya i otritsatel'nymi zayavkami [Open networks with with multimode service strategies and negative demands]. *Vestnik Tomskogo gosudarstvennogo universiteta – Tomsk State University Journal*. 1(1). pp. 90-93.
5. Nueman, A.Yu. (2002) [Queuing network with unreliable devices and negative demands]. *Novye matematicheskie metody i komp'yuternye tekhnologii v proektirovanii, proizvodstve i nauchnykh issledovaniyakh* [New mathematical methods and computer technology in design, production and research]. Proc. of the 5th Rep. Conf. of students and post-graduates. Gomel. 18th to 20th March 2002. Gomel: Gomel State University. pp. 179-180.
6. Bocharov P.P. & Pechinkin, A.V. (1995) *Teoriya massovogo obsluzhivaniya* [Queuing theory]. Moscow: RUDN.
7. Starovoytov, A.N. (2007) Seti s mnogorezhimnym obsluzhivaniem, otritsatel'nymi zayavkami i proizvol'nym vremenem prebyvaniya v rezhimakh [Networks with the multimode service, negative demands and arbitrary time of staying in conditions]. *Izvestiya Gomel'skogo gosudarstvennogo universiteta im. F. Skoriny*. 6 (45). pp. 193-198.
8. Letunovich, Yu.E. (2008) Statsionarnoe raspredelenie sostoyaniy otkrytoy neodnorodnoy seti s mnogorezhimnymi strategiyami i nemedlennym obsluzhivaniem [Stationary distribution of states of open heterogeneous network with multimode strategies and immediate service]. In: *Sovremennye informatsionnye komp'yuternye tekhnologii* [Modern information computer technologies]. Grodno: GrSU im. Ya. Kupaly. pp. 97-99.
9. Ivnitskiy, V.A. (2004) *Teoriya setey massovogo obsluzhivaniya* [Theory of queuing networks]. Moscow: Fizmatlit.

ВЕСТНИК ТОМСКОГО ГОСУДАРСТВЕННОГО УНИВЕРСИТЕТА
2015 Управление, вычислительная техника и информатика № 4 (33)

УДК 519.95
DOI: 10.17223/19988605/33/4

Н.А. Игнатьев

ВЫЧИСЛЕНИЕ ОБОБЩЕННЫХ ОЦЕНОК И ИЕРАРХИЧЕСКАЯ ГРУППИРОВКА ПРИЗНАКОВ

Рассматривается процесс формирования нового признакового пространства, размерность которого меньше исходного. Предлагается последовательный отбор непересекающихся подмножеств разнотипных признаков в описании объектов и нелинейное отображения их на числовую ось. При отборе используется правило иерархической группировки для попарного объединения признаков. Решение принимается по значениям степени размытости результатов отображения объектов классов на числовой оси.
Ключевые слова: обобщённые оценки; иерархическая группировка; логические закономерности; отступ.

Обобщённые оценки – это агрегированные (комбинированные) показатели, которые в [1] использовались для отображения отношений между объектами двух классов в разнотипном признаковом пространстве на числовую ось. Было разработано два метода вычисления оценок: стохастический и детерминистический. Критерием для выбора параметров алгоритма стохастического метода служила максимальная разность (отступ) между линейными проекциями двух объектов из разных классов. Из минимального значения на числовой оси одного класса вычиталось максимальное значение другого класса. Одним из применений метода было отображение описаний (визуализация) объектов [2] на плоскость.

В алгоритме детерминистического метода применялось разбиение на интервалы доминирования значений количественных признаков объектов одного из двух классов. При вычислении обобщённой оценки объекта использовались значения функций принадлежности к интервалам доминирования для количественных признаков и частоты встречаемости градаций для номинальных признаков.

Переход к однотипным шкалам измерений и поэтапное сокращение размерности признакового пространства посредством вычисления обобщённых оценок объектов описан в [3]. На первом этапе обобщённая оценка объекта по номинальным признакам интерпретировалась как значение нового (латентного) количественного признака. На втором этапе вычисление оценки производилось по расширенному множеству количественных признаков.

Результаты вычислительного эксперимента в [3] по выборке данных GERMAN из [4] показали, что обобщающая способность решающих правил на основе обобщённых оценок выше, чем у известного метода LDA [5].

Потребность во введении латентных признаков возникает при поиске спрямляющего пространства, в котором объекты из разных классов были бы линейно разделимы. В методе опорных векторов SVM [6] нелинейность разделяющей поверхности достигается за счёт использования ядерных функций, поиск параметров дискриминантных функций производится путём максимизации отступа между объектами двух классов в новом (спрямляющем) признаковом пространстве.

В данном исследовании предлагается правило для агломеративной иерархической группировки разнотипных признаков с целью нелинейного отображения их значений в описании объектов на числовую ось. Результаты нелинейного отображения рассматриваются как значения обобщённых оценок (новых признаков) в описании объектов. Предложены критерии, на основе которых определяются число обобщённых оценок (непересекающихся групп), количество исходных признаков, входящих в группу, и их состав.

Решающие правила по значениям каждого нового признака в описании объектов образуют совокупность базовых алгоритмов. Базовый алгоритм может рассматриваться как самостоятельный классификатор либо использоваться в композиции с другими алгоритмами.

Вычисление обобщенных оценок с помощью иерархической агломеративной группировки целесообразно по нескольким причинам:

– обобщённые оценки образуют новое признаковое пространство, размеры которого меньше исходного;

– решается проблема использования алгоритмов классификации, реализация которых была неэффективна из-за большой размерности признакового пространства либо возможна при одном типе шкал измерений;

– в процессе группировки происходит последовательный отбор информативных наборов признаков;

– нелинейное отображение описаний объектов на числовую ось по определяемым комбинациям признаков является средством обнаружения устойчивых логических закономерностей (новых знаний) в хранилищах данных.

1. Обобщенные оценки объектов на базе иерархической группировки признаков

Рассматривается множество из T допустимых объектов, разбитое на 2 непересекающихся подмножества (класса). Представители классов K_1, K_2 заданы через выборку (подмножество T) объектов $E_0 = \{S_1,\ldots,S_m\}$, $E_0 = K_1 \cup K_2$. Объекты выборки описываются с помощью n разнотипных признаков $X(n) = (x_1,\ldots,x_n)$, множество допустимых значений ξ из которых измеряются в интервальных шкалах, $n - \xi$ – в номинальной.

На E_0 задано правило последовательного разбиения набора $X(n)$ на непересекающиеся подмножества $X_1(k_1),\ldots, X_\tau(k_\tau)$, $\tau \geq 1$, $k_1+\ldots+k_\tau \leq n$. Требуется для каждого $X_i(k_i)$ определить алгоритм A_i (распознающий оператор в терминологии алгебраического подхода к распознавания образов Ю.И. Журавлёва [7]) для отображения значений признаков из $X_i(k_i)$ в описании объекта $S_j \in E_0$, $j = 1,\ldots,m$, в значение (обобщённую оценку) на числовой оси.

Обозначим множество номеров количественных и номинальных признаков соответственно как I и J. Процесс последовательного вычисления значений обобщённых оценок (новых признаков) реализуется алгоритмом иерархической агломеративной группировки по описываемому ниже правилу. Для идентификации признаков в описании объектов на p-м шаге $0 \leq p < n$ иерархической группировки будем использовать $\{x_i^p\}_{i \in (I \cup J)}$.

В процессе группировки и формирования обобщённых оценок состав элементов и мощность множеств I и J, $|I| + |J| \leq n$ будут изменятся. В зависимости от шкал измерений признаков, объединяемых в группы, используются различные способы вычисления их параметров для отображения на числовую ось. Для количественных признаков это производится следующим образом.

Упорядоченное множество значений признака $x_j^p, j \in I, p \geq 0,$ объектов из E_0 разделим на два интервала $[c_1^{ip}, c_2^{ip}], (c_2^{ip}, c_3^{ip}]$, каждый из которых рассматривается как градация номинального признака. Критерий для определения границы c_2^{ip} основывается на проверке гипотезы (утверждения) о том, что каждый из двух интервалов содержит значения количественного признака объектов только одного класса.

Пусть u_i^1, u_i^2 – количество значений признака x_j^p, $j \in I,$ класса K_i, $i = 1,2$, соответственно в интервалах $[c_1^{ip}, c_2^{ip}], (c_2^{ip}, c_3^{ip}]$; $|K_i| > 1$, ν – порядковый номер элемента упорядоченной по возрастанию последовательности $r_{j_1},\ldots,r_{j_\nu},\ldots,r_{j_m}$ значений x_j^p из E_0, определяющий границы интервалов как $c_1^{jp} = r_{j_1}$, $c_2^{jp} = r_{j_\nu}$ $c_3^{jp} = r_{j_m}$. Критерий

$$\left(\frac{\sum_{i=1}^{2} u_i^1 \left(u_i^1 - 1\right) + u_i^2 \left(u_i^2 - 1\right)}{\sum_{i=1}^{2} |K_i| \left(|K_i| - 1\right)} \right) \left(\frac{\sum_{d=1}^{2} \sum_{i=1}^{2} u_i^d \left(|K_{3-i}| - u_{3-i}^d\right)}{2|K_1||K_2|} \right) \rightarrow \max_{c_1^{jp} < c_2^{jp} < c_3^{jp}} \qquad (1)$$

позволяет оценивать значение границы между интервалами $[c_1^{ip}, c_2^{ip}]$,$(c_2^{ip}, c_3^{ip}]$. Выражение в левых скобках (1) представляет внутриклассовое сходство, в правых – межклассовое различие.

Экстремум критерия (1) используется в качестве веса w_j^p $(0 \le w_j^p \le 1)$ признака x_j^p. При $w_j^p = 1$ значения признака x_j^p у объектов из классов K_1 и K_2 не пересекаются между собой.

При включении в группу номинального признака с целью вычисления обобщённой оценки объектов требуется определить значение его веса и вкладов каждой из градаций.

Обозначим через π число градаций признака x_r^p, $r \in J$, $p = 0$, g_{dr}^t – количество значений t-й $(1 \le t \le \pi)$ градации r-го признака в описании объектов класса K_d, l_{dr} – число градаций r-го признака в K_d, $d = 1,2$.

Различие по r-му признаку между классами K_1 и K_2 определяется как величина

$$\lambda_r = 1 - \frac{\sum_{t=1}^{\pi} g_{1r}^t g_{2r}^t}{|K_1||K_2|}. \tag{2}$$

Степень однородности (мера внутриклассового сходства) β_r значений градаций r-го признака по классам K_1, K_2 вычисляется по формулам

$$D_{dr} = \begin{cases} \left(|K_d| - l_{dr} + 1\right)\left(|K_d| - l_{dr}\right), \pi > 2, \\ |K_d|\left(|K_d| - 1\right), \pi \le 2; \end{cases}$$

$$\beta_r = \begin{cases} \dfrac{\sum_{t=1}^{\pi} g_{1r}^t \left(g_{1r}^t - 1\right) + g_{2r}^t \left(g_{2r}^t - 1\right)}{D_{1r} + D_{2r}}, D_{1r} + D_{2r} > 0, \\ 0, \ D_{1r} + D_{2r} = 0. \end{cases} \tag{3}$$

С помощью (2), (3) вес номинального признака с $r \in J$ определяется как

$$\nu_r = \lambda_r \beta_r.$$

Очевидно, что множество чисел, идентифицирующих π градаций номинального признака, всегда можно взаимно однозначно отобразить в множество $\{1,\ldots,\pi\}$. С учётом такого отображения для объекта $S = (a_i,\ldots,a_n)$ вклад признака $a_i = j$, $i \in J$, $j \in \{1,\ldots,\pi\}$ в обобщённую оценку определяется величиной

$$\mu_i(j) = \nu_i \left(\frac{\alpha_{ij}^1}{|K_1|} - \frac{\alpha_{ij}^2}{|K_2|} \right),$$

где $\alpha_{ij}^1, \alpha_{ij}^2$ – количество значений j-й градации i-го признака соответственно в классах K_1 и K_2, ν_i – вес i-го признака.

Значение обобщённой оценки b_{rij}^p объекта $S_r = \{a_{ru}^p\}_{u \in (I \cup J)}$, $S_r \in E_0$, по паре x_i^p, x_j^p, $0 \le p < n$, $i, j \in (I \cup J), i \in j$, вычисляется как

$$b_{rij}^p = \begin{cases} \mu_i\left(a_{ri}^p\right) + \mu_j\left(a_{rj}^p\right), i, j \in J, \\ \mu_i\left(a_{ri}^p\right) + t_j w_j^p \left(a_{rj}^p - c_2^{jp}\right) / \left(c_3^{jp} - c_1^{jp}\right), i \in J, j \in I, t_j \in \{-1,1\}, \\ \eta_{ij}\left(t_i w_i^p \left(a_{ri}^p - c_2^{ip}\right) / \left(c_3^{ip} - c_1^{ip}\right) + t_j w_j^p \left(a_{rj}^p - c_2^{jp}\right) / \left(c_3^{jp} - c_1^{jp}\right)\right) + \\ + \left(1 - \eta_{ij}\right) t_{ij} w_{ij}^p \left(a_{ri}^p a_{rj}^p - c_2^{ijp}\right) / \left(c_3^{ijp} - c_1^{ijp}\right), \\ i, j \in I, t_{ij}, t_i, t_j \in \{-1,1\}, \eta_{ij} \in [0,1], \end{cases} \tag{4}$$

где w_i^p, w_j^p, w_{ij}^p – веса признаков, определяемые по (1) соответственно по множеству значений признаков x_i^p, x_j^p и их произведения $x_i^p x_j^p$, значения $t_{ij}, t_i, t_j \in \{-1,1\}$, $\eta_{ij} \in [0,1]$ выбираются по экстремуму функционала

$$\varphi(p,i,j)=\frac{\min\limits_{S_r\in K_1} b_{rij}^p-\max\limits_{S_r\in K_2} b_{rij}^p}{\max\limits_{S_r\in E_0} b_{rij}^p-\min\limits_{S_r\in E_0} b_{rij}^p}=\max_{t_{ij},t_i,t_j\in\{-1,1\},\eta_{ij}\in[0,1]}. \tag{5}$$

Значение (5) интерпретируется как отступ между объектами классов K_1 и K_2.

Обозначим через $\left\{z_{ij}^p\right\}_{i,j\in(I\cup J)}$, $p\ge 0$, квадратную матрицу размера $(n-p)\times(n-p)$, значение элемента z_{ij}^p которой определяется как

$$z_{ij}^p=\begin{cases}0, i=j,\\ \text{значению}\quad (1)\quad \text{по}\quad \left\{b_{rij}^p\right\}_{r=1}^m, i\ne j,\end{cases} \tag{6}$$

через G_η, $\eta>0$, – подмножество номеров признаков из $X(n)$. Пошаговая реализация алгоритма итеративной группировки будет такой:

1-й шаг: $\eta=1$, $G_\eta=\varnothing$, $p=0$, $\lambda c=0$;

2-й шаг: Вычислить значения элементов матрицы $\left\{z_{ij}^p\right\}_{i,j\in(I\cup J)}$ по (6);

3-й шаг: Вычислить $\lambda n=\max\limits_{u,v\in(I\cup J)} z_{uv}^p$. Выделить $\Omega=\left\{(s,t),s,t\in I\cup J\middle| z_{st}^p=\lambda n \quad and \quad s<t\right\}$. Определить пару $\{i,j\}, i<j$, как

$$\{i,j\}=\begin{cases}\Omega,|\Omega|=1,\\ \{s,t\},(s,t)\in\Omega \quad \text{and} \quad \varphi(p,s,t)>\max\limits_{(u,v)\in\Omega\setminus(s,t)}\varphi(p,u,v);\end{cases}$$

4-й шаг: Если $G_\eta=\varnothing$, то $G_\eta=\{i,j\}$, $Margin=\varphi(p,i,j)$, идти 8;

5-й шаг: Если $G_\eta\cap\{i,j\}=\varnothing$, то идти 7;

6-й шаг: Если $\lambda n>\lambda c$ или $\lambda n>\lambda c$ и $Margin<\varphi(p,i,j)$, то $G_\eta=G_\eta\cup\{i,j\}$, $Margin=\varphi(p,i,j)$, идти 8;

7-й шаг: $\eta=\eta+1$, $G_\eta=\varnothing$. Идти 4;

8-й шаг: $p=p+1$, $I\cup J=(I\cup J)\setminus\max(i,j)$, $I=I\cup\min(i,j)$, $k=\min(i,j)$, $\lambda c=\lambda n$. Заменить значения признаков в описании объекта $S_r=\left\{a_{ru}^{p-1}\right\}_{u\in(I\cup J)}$, $r=1,\ldots,m$ на

$$a_{ru}^p=\begin{cases}a_{ru}^{p-1}, u\in(I\cup J)\setminus k,\\ b_{rij}^p, u=k;\end{cases}$$

9-й шаг: Определить значение

$$z_{uv}^p=\begin{cases}z_{uv}^{p1}, u\in I\setminus\{k\}, v\in I,\\ \text{значению}\quad (1)\quad \text{по}\quad \left\{a_{rv}^p\right\}_{r=1}^m, u=k, v\in I.\end{cases}$$

Если $n-p>1$, то идти 3;

10-й шаг: Конец.

Через конечное число рекурсивных обращений к описанному выше алгоритму все исходные признаки сводятся к одной нелинейной оценке. По практическим соображениям ограничение на число обобщённых оценок для конкретных выборок данных может определяться по результатам вычислительного эксперимента либо исходя из дополнительных критериев выбора.

Рассмотрим пример классификатора на базе обобщённых оценок (4). Пусть $\left\{a_{ir}^p\right\}_{i=1}^m$, $p<n$, $r\in I$ – множество значений обобщённой оценки (признака), вычисленной по (4), и по критерию (1) эти значения разбиты на интервалы $[c_1,c_2],(c_2,c_3]$. Для решающего правила нужно выбрать порог, равный

$$w_0=\frac{c_2+z}{2}, \tag{7}$$

где z – ближайшее к c_2 значение из интервала $(c_2,c_3]$. Анализ результатов использования порога (7) в дискриминантных функциях приводится в [3].

3. Вычислительный эксперимент

В качестве материала для эксперимента была взята выборка данных из [8], описывающая челюсти 30 собак (класс K_1) и 12 волков (класс K_2) по следующим 6 количественным признакам:

x_1 – (CBL) основная длина;

x_2 – (LUJ) длина верхней челюсти;

x_3 – (WID) ширина верхней челюсти;

x_4 – (LUC) длина верхнего карнивора;

x_5 – (LFM) длина первого верхнего моляра;

x_6 – (WFM) ширина первого верхнего моляра.

Порядок синтеза значений обобщённых оценок (латентных признаков) из комбинаций признаков по критерию (1) и отступов между объектами классов по (5) приведён в табл. 1.

Таблица 1

Порядок синтеза обобщённых оценок объектов

Комбинация признаков	Значение критерия (1)	Отступ между классами (5)
X_1, x_4	1,0000	0,0403
x_1, x_4, x_5	1,0000	0,1060
x_1, x_4, x_5, x_3	1,0000	0,1233
x_1, x_4, x_5, x_3, x_2	1,0000	0,1674
x_1, x_4, x_5, x_3, x_2, x_6	1,0000	0,1778

Аналитический вид решающего правила по значениям обобщённой оценки, полученной при синтезе признаков x_1 и x_4 (табл. 1) с учётом (7), будет выглядеть так:

$$d(x) = 0{,}4(-0{,}0037(x_1-221)-0{,}09538(x_4-22{,}5))+0{,}0001(x_1x_4-5130)+0{,}01971.$$

Судя по результатам табл. 1, все комбинации исходных признаков попадают в одну группу, попарное объединение признаков в комбинацию удовлетворяет такому свойству, как монотонность по значениям отступа между объектами классов. Теоретическое обоснование выполнения монотонности при синтезе комбинаций признаков на произвольной двухклассовой обучающей выборке требует отдельного рассмотрения. Возможным вариантом решения проблемы монотонности является обнаружение и исключение из выборки шумовых объектов.

Для демонстрации того, что различные признаки в составе обобщённых оценок (4) компенсируют недостатки друг друга, воспользуемся табл. 2 [1]. Таблица содержит значения границ интервалов $[c_1^i, c_2^i], (c_2^i, c_3^i]$ и экстремумы критерия (1) для признаков $\{x_i\}$, $i = 1,\ldots,6$.

Таблица 2

Результаты оптимизации по критерию (1)

Признак	c_1^i	c_2^i	c_3^i	w_i
x_1	129,000	221,000	255,000	0,378
x_2	64,000	114,000	126,000	0,389
x_3	52,000	76,000	95,000	0,288
x_4	16,700	22,500	26,500	0,897
x_5	11,200	14,700	16,800	0,625
x_6	13,000	18,300	27,000	0,800

Согласно [1] точность классификации по линейным дискриминантным функциям (дискриминанта Фишера в том числе) напрямую зависит от использования признаков x_4 и x_6, имеющих наибольшие значения весов (табл. 2), равных соответственно 0,897 и 0,800. Доказано, что корректное разделение линейных проекций объектов обучения на классы с единичным значением критерия (1) возможно лишь на наборах $\{x_1, x_2, x_3, x_4, x_5, x_6\}$ и $\{x_1, x_2, x_4, x_5, x_6\}$. Наилучший результат (см. табл. 1) в смысле разделимости по (1) и отступа между классами (5) по различным парам исходных признаков был достигнут при использовании нелинейного отображения значений из $\{x_1, x_4\}$ в описании объектов на числовую ось.

С помощью критерия (1) преобразуем количественные признаки в номинальные. Каждой градации номинального признака поставим в соответствие один из непересекающихся интервалов, полученный по (1). В табл. 3 приведены результаты группировки для случая, когда все признаки номинальные,

в табл. 4 представлено два подмножества: $\{x_1, x_2\}$ – количественных и $\{x_3, x_4, x_5, x_6\}$ – номинальных признаков.

Таблица 3

Группировка по номинальным признакам

№ группы	Состав группы	Значение критерия (1)
1	x_2, x_4, x_5, x_6	0,8965
2	x_1x_3	0,3781

Таблица 4

Группировка по разнотипным признакам

№ группы	Состав группы	Значение критерия (1)
1	x_1, x_2, x_4, x_5	1,0000
2	x_3,$x6$	0,2884

Анализ результатов из табл. 1, 2 и 3 по критерию (1) показывает, что преобразование значений признаков из количественных (сильных) шкал измерений в значения номинальной (слабой) шкалы приводят к снижению точности решающих правил с порогом (7) на базе обобщённых оценок.

Для проверки процедурой кросс-валидации обобщающей способности решающих правил с порогом (7) на базе нелинейных обобщённых оценок с максимальным отступом между классами использовалось разделение выборки на обучение и контроль в соотношении 9:1. Результаты проверки следующие: точность на обучении – 100%, на контроле – 98%.

Заключение

Процесс вычисления обобщённых оценок сводится к формированию нового признакового пространства для описания допустимых объектов в задачах распознавания образов. Практическое применение этих оценок позволяет:

– находить устойчивые логические закономерности в базах (хранилищах) данных, не прибегая к перебору всевозможных вариантов;

– использовать их для реализации дискриминантных функций, решающих списков, решающих деревьев, алгоритмов вычисления оценок.

Теоретический и практический интерес представляет оценка границ допустимых значений латентных признаков на основе обобщённых оценок.

ЛИТЕРАТУРА

1. *Игнатьев Н.А.* Вычисление обобщённых показателей и интеллектуальный анализ данных // Автоматика и телемеханика. 2011. № 5. С. 183–190.
2. *Игнатьев Н.А.* О конструировании признакового пространства для поиска логических закономерностей в задачах распознавания образов // Вычислительные технологии. 2012. Т. 17, № 4. С. 56–62.
3. *Игнатьев Н.А., Нуржонов Ш.Ю.* Выбор параметров регуляризации для повышения обобщающей способности дискриминантных функций // Узбекистон Республикаси Курол Кучлари академиясининг хабарлари. 2014. № 1 (14). С. 81–87.
4. *Asuncion A., Newman D.J.* UCI Machine Learning Repository // University of California, Irvine. 2007. URL: www.ics.uci.edu/mlearn/MLRepository.html
5. URL: http://www.mathworks.com/help/stats/discriminant-analysis.html
6. *Потапов А.С.* Технологии искусственного интеллекта. СПб. : СПбГУ ИТМО, 2010. 218 с.
7. *Журавлёв Ю.И.* Об алгебраическом подходе к решению задач распознавания или классификации // Проблемы кибернетики. М. : Наука, 1978. Т. 33. С. 5–68.
8. *Жамбю М.* Иерархический кластер-анализ и соответствия. М. : Финансы и статистика, 1988. 342 с.

Игнатьев Николай Александрович, д-р физ.-мат. наук, профессор. E-mail: ignatev@rambler.ru
Национальный университет Узбекистана

Поступила в редакцию 22 сентября 2015 г.

Ignat'ev Nikolai A. (National University of Uzbekistan. Republic of Uzbekistan).
Computation generalized estimates of objects and hierarchical clustering of features
Keywords: generalized estimates; hierarchical clustering; logical regularity; margin.

DOI: 10.17223/19988605/33/4

We consider the set acceptable objects in T, broken into two disjoint subsets (classes). Representatives of the classes K_1, K_2 are given by the sample objects $E_0 = \{S_1,\ldots,S_m\}$, $E_0 = K_1 \cup K_2$. Objects of the sample are described by n heterogeneous features $X(n) = (x_1,\ldots,x_n)$ the set of acceptable values of ξ which measurements are on the interval scale and n - ξ are on the nominal.

Given the rule is a sequence of partitions set $X(n)$ into disjoint subsets $X_1(k_1),\ldots,X_\tau(k_\tau)$, $\tau \geq 1$, $k_1+\ldots+k_\tau \leq n$ at E_0. Required for each $X_i(k_i)$ algorithm to determine A_i nonlinear display feature values $X_i(k_i)$ of the object in the description $S_j \in E_0$, $j=1,\ldots,m$, a value (generalized estimation) on the real axis.

For the identification generalized estimates (new features) in object descriptions on p-th step $0 \leq p < n$ hierarchical clustering is used $\{x_i^p\}_{i \in (I \cup J)}$, where I and J, respectively, the set of indices of nominal and quantitative features. Making a decision to merge pairs of features is based on the interval analysis of the results of non-linear displaying them on the real axis and the margin between classes.

The ordered set feature values of objects in E0 is divided into two intervals $[c_1^{jp}, c_2^{jp}],(c_2^{jp}, c_3^{jp}]$. The criterion for determining the borders c_2^{jp} based on hypothesis testing (assertion) that each of the two intervals contains the values of a quantitative feature of only one class objects.

Let u_i^1, u_i^2 – the number of characteristic values $x_j^p, j \in I$ class K_i, $i = 1,2$ respectively, in the intervals $[c_1^{jp}, c_2^{jp}],(c_2^{jp}, c_3^{jp}]$, $|K_i| > 1$, ν – the ordinal number of the element ordered ascending sequence of x_j^p values of E_0, determining interval limits as $r_{j_1},\ldots,r_{j_\nu},\ldots,r_{j_m}$. Criterion

$$\left(\frac{\sum_{i=1}^{2} u_i^1\left(u_i^1-1\right)+u_i^2\left(u_i^2-1\right)}{\sum_{i=1}^{2}|K_i|\left(|K_i|-1\right)}\right)\left(\frac{\sum_{d=1}^{2}\sum_{i=1}^{2} u_i^d\left(|K_{3-i}|-u_{3-i}^d\right)}{2|K_1||K_2|}\right) \to \max_{c_1^{jp}<c_2^{jp}<c_3^{jp}} \quad (1)$$

allows to assess the meaning borders between interval $[c_1^{jp}, c_2^{jp}],(c_2^{jp}, c_3^{jp}]$. The extremum of the criterion (1) is used as a weight w_j^p ($0 \leq w_j^p \leq 1$) of feature x_j^p and for a decision by the rule of hierarchical clustering.

If $\forall x_t \in X(n) \setminus \bigcup_{d=1}^{i} X_d(k_d)$ $\left|\bigcup_{d=1}^{i} X_d(k_d)\right| < n$ nonlinear mapping on the real axis of objects E_0 on $X_i(k_i) \cup \{x_t\}$ value (1) less than or equal to the analogical value (with a less margin between classes) on $X_i(k_i) \cup \{x_t\}$ that is formed new group for the synthesis of the generalized estimation.

Calculation of generalized estimates using hierarchical clustering advisable for several reasons:

– generalized estimates form a new feature space whose dimensions are smaller than the original;

– solves the problem of the use of classification algorithms, the implementation of which was inefficient due to the large dimension of feature space, is possible at any single type measurement scales;

– in the process of clustering occurs consistent selection of informative feature sets;

– nonlinear mapping object description to real axis defined by a combination of features is a means of detection stable patterns of logic (new knowledge) in data warehouses.

REFERENCES

1. Ignat'ev, N.A. (2011) Computing generalized parameters and data mining. *Automation and Remote Control*. 72 (5). pp. 183-190. DOI: 10.1134/S0005117911050146
2. Ignat'ev, N.A. (2012) On the construction of the feature space for finding logical regularities in pattern recognition problems. *Vychislitel'nye tekhnologii – Computational Technologies*. 17(4). pp. 56-62. (In Russian).
3. Ignat'ev, N.A. & Nurzhonov, Sh.Yu. (2014) Vybor parametrov regulyarizatsii dlya povysheniya obobshchayushchey sposobnosti diskriminantnykh funktsiy [The choice of the regularization parameters for improving the generalization ability of the discriminant functions]. *Uzbekiston Respublikasi Kurol Kuchlari akademiyasining khabarlari – Bulletin of the Academy of the Armed Forces of the Republic of Uzbekistan*. 1(14). pp. 81-87.
4. Asuncion, A. & Newman, D.J. (2007) *UCI Machine Learning Repository*. University of California, Irvine. [Online] Available from: www.ics.uci.edu/mlearn/MLRepository.html
5. Mathworks. (n.d.) *Discriminant Analysis*. [Online] Available from: http://www.mathworks.com/help/stats/discriminant-analysis.html
6. Potapov, A.S. (2010) *Tekhnologii iskusstvennogo intellekta* [Artificial intelligence technology]. St. Petersburg: SPbGU ITMO.
7. Zhuravlev, Yu.I. (1978) Ob algebraicheskom podkhode k resheniyu zadach raspoznavaniya ili klassifikatsii [On an algebraic approach to solving the problems of pattern recognition and classification]. In: Girevich., I.B. (ed.) *Problemy kibernetiki* [Problems of Cybernetics]. Vol. 33. Moscow: Nauka. pp. 5-68.
8. Jambu, M. (1988) *Ierarkhicheskiy klaster-analiz i sootvetstviya* [Hierarchical cluster analysis and compliance]. Moscow: Finansy i statistika.

УДК 336:51
DOI: 10.17223/19988605/33/5

G.A. Medvedev

THE NELSON-SIEGEL-SVENSSON YIELDS. PROBABILITY PROPERTIES AND ESTIMATION

Probabilty properties of the yield interest rates that are generated by model of Nelson – Siegel and Nelson – Siegel – Svensson are considered. The paper is directly related to [1]. It is shown that the model of Nelson – Siegel does not differ from the traditional two-factor model of affine yield, the volatility of which does not depend on the market state variables. Accordingly, the model of Nelson – Siegel – Svensson – from a four-factor model. These models generate the interest rates of yield to maturity and the forward yield with a normal distribution, for which the expectations and covariance matrices are found. To estimate the values of the rates of yield to maturity in the current time it is offered a recurrent procedure based on the use of the Kalman filter.
Keywords: term structure; yield curve; affine models; Nelson-Siegel-Svensson yield model; Kalman filter.

The Nelson – Sigel (NS) yield models belong to the affine family of models. The affine models are the factor models where some variables, for example a components of the state vector x, are considered as factors. In affine models of the term structure of interest rates it is assumed that the yield to maturity $y(\tau, x)$ on a zero coupon bond with price $P(\tau, x)$ is defined by relation

$$y(\tau, x) \equiv \frac{-\ln P(\tau, x)}{\tau} = \frac{x^{\mathrm{T}} B(\tau) - A(\tau)}{\tau},$$

where τ – term to maturity and $A(\tau)$, $B(\tau)$ are the functions of term structure. Usually as components of the state vector x some market indexes are taken and then functions of term structure $A(\tau)$ and $B(\tau)$ are determined so to avoid an arbitrage opportunities. In their yield model C. Nelson and A. Sigel [2] do conversely: they set vector $B(\tau)$ and then determine vector x so to fit the yield $y(\tau, x)$ to the market observations. Vector $B(\tau)$ that set by Nelson and Sigel is

$$B(\tau) = \begin{pmatrix} \dfrac{1-e^{-\gamma\tau}}{\gamma} \\ \dfrac{1-e^{-\gamma\tau}}{\gamma} - \tau e^{-\gamma\tau} \end{pmatrix} \qquad (1)$$

with parameter $\gamma > 0$. $A(\tau) = -\tau L(\tau)$ where $L(\tau)$ is known as a *level* factor. Vector x has components S_t and C_t: S_t – a *slope* factor and C_t – a *curvature* factor. Factors L, S and C are named too *latent* factors.

Thus in the NS yield models the yield curve $y(\tau) \equiv y(\tau \mid L, S, C)$ and the forward curve $f(\tau) \equiv f(\tau \mid L, S, C)$ are determined by following relations

$$y(\tau) = L(\tau) + S_t \frac{1-e^{-\gamma\tau}}{\gamma\tau} + C_t \left(\frac{1-e^{-\gamma\tau}}{\gamma\tau} - e^{-\gamma\tau} \right), \qquad (2)$$

$$f(\tau) = (\tau L(\tau))' + S_t \exp(-\gamma\tau) + C_t \gamma\tau \exp(-\gamma\tau).$$

L. Svensson [3] offered to extend the NS model in order to increase the flexibility and improve the fit models to empirical data. He introduced the additional latent factor but with parameter $\delta \neq \gamma$, $\delta > 0$. Then vectors $B(\tau)$ and x will have four components:

$$B(\tau)=\begin{pmatrix} \dfrac{1-e^{-\gamma\tau}}{\gamma} \\ \dfrac{1-e^{-\gamma\tau}}{\gamma}-\tau e^{-\gamma\tau} \\ \dfrac{1-e^{-\delta\tau}}{\delta} \\ \dfrac{1-e^{-\delta\tau}}{\delta}-\tau e^{-\delta\tau} \end{pmatrix},\ x=\begin{pmatrix} S_t \\ C_t \\ G_t \\ H_t \end{pmatrix}. \tag{3}$$

Note that in Svensson's proposal $G_t = 0$ (more exactly: G_t is absent) and vectors $B(\tau)$ and x have only three components. However as it is shown [1] in this case the no-arbitrage conditions are not fulfilled. Therefore we will assume that G_t can be not zero.

So that in the Nelson – Sigel – Svensson (NSS) model the yield curve $y(\tau) \equiv y(\tau \mid L, S, C, G, H)$ and the forward curve $f(\tau) \equiv f(\tau \mid L, S, C, G, H)$ are determined by following relations

$$y(\tau) = L(\tau) + S_t\frac{1-e^{-\gamma\tau}}{\gamma\tau} + C_t\left(\frac{1-e^{-\gamma\tau}}{\gamma\tau}-e^{-\gamma\tau}\right) + G_t\frac{1-e^{-\delta\tau}}{\delta\tau} + H_t\left(\frac{1-e^{-\delta\tau}}{\delta\tau}-e^{-\delta\tau}\right),$$

$$f(\tau) = (\tau L(\tau))' + S_t \exp(-\gamma\tau) + C_t\,\gamma\tau \exp(-\gamma\tau) + G_t \exp(-\delta\tau) + H_t\,\delta\tau \exp(-\delta\tau).$$

1. Probability properties

It was shown [1] that the yield model is the no-arbitrage Nelson – Sigel (NS) model or the no-arbitrage Nelson – Sigel – Svensson (NSS) model only when the vector of a state variables of financial market $X(t) = (X_1, X_2, \ldots, X_n)^{\mathrm{T}}$ follows to homogeneous for time Markov process generated by the stochastic differential equation with linear function of drift $\mu(x)$ over x, and the volatility matrix $\sigma(x)$ not depended on x.

Under these conditions this equation can be written down in a form

$$dX(t) = [-KX(t) + K\theta]dt + \sigma W(t),\quad X(t_0) = X_0, \tag{4}$$

where K – $(n\times n)$-matrix, θ – n-vector, σ – $(n\times m)$-matrix and $W(t) = \{W(t);\ t \ge t_0\}$ – m-dimensional Wiener process with components, which are independent scalar Wiener processes. As well as any differential equation (10) should be supplied by the entry condition fixing a vector $X(t)$ at some initial time t_0. It is possible to show [4] that the equation (4) can be solved analytically in the following form

$$X(t) = \Theta(t_0,t)\left[X_0 + \int_{t_0}^{t}\Theta^{-1}(t_0,s)K\theta ds + \int_{t_0}^{t}\Theta^{-1}(t_0,s)\sigma dW(s)\right], \tag{5}$$

where $\Theta(t_0,t)$ – $(n\times n)$ fundamental matrix of solutions, satisfying to the initial condition $\Theta(t_0,t_0) = I$ and to the homogeneous matrix differential equation

$$d\,\Theta(t_0,t) = -K\Theta(t_0,t)dt, \tag{6}$$

which can be considered as n the vector differential equations. Note that to solve this equation with respect $\Theta(t_0,t)$ in an explicit form also it is possible. As the equation (4) is the equation with constant coefficients the fundamental matrix of solutions $\Theta(t_0,t)$ will depend not on two arguments, but only from one: $\Theta(t_0,t) = \Theta(t-t_0)$. To stop receiving cumbersome expressions, we present the solution of equations (4)–(6) for NS model, when $n = m = 2$. In this case matrix K is determined for NS model by the formula

$$K = \begin{pmatrix} \gamma & -\gamma \\ 0 & \gamma \end{pmatrix} \tag{7}$$

and the solution of the equation (6) has a form:

$$\Theta(t-t_0) = \begin{pmatrix} e^{-\gamma(t-t_0)} & \gamma(t-t_0)e^{-\gamma(t-t_0)} \\ 0 & e^{-\gamma(t-t_0)} \end{pmatrix},\quad \Theta^{-1}(t-t_0) = \begin{pmatrix} e^{\gamma(t-t_0)} & -\gamma(t-t_0)e^{\gamma(t-t_0)} \\ 0 & e^{\gamma(t-t_0)} \end{pmatrix}.$$

Therefore expression (5) takes the form

$$X(t) = \begin{pmatrix} X_1(t) \\ X_2(t) \end{pmatrix} = \begin{pmatrix} e^{-\gamma(t-t_0)}[X_1(t_0) + \gamma(t-t_0)X_2(t_0)] \\ e^{-\gamma(t-t_0)}X_2(t_0) \end{pmatrix} +$$

$$+ \begin{pmatrix} (1-e^{-(t-t_0)\gamma})\theta_1 - \gamma(t-t_0)e^{-(t-t_0)\gamma}\theta_2 \\ (1-e^{-(t-t_0)\gamma})\theta_2 \end{pmatrix} + \xi(t). \quad (8)$$

Here $\xi(t)$ – a two-dimensional normally distributed random vector process with the components having zero mean and covariance matrix

$$\Omega(t-t_0) = E\,[(X(t) - E\,[X(t)])\,(X(t) - E\,[X(t)])^{\mathrm{T}}] = E\,[\xi(t)\,\xi(t)^{\mathrm{T}}].$$

Generally the matrix $\Omega(t - t_0)$ is calculated easy, but its elements $[\Omega(t - t_0)]_{jk}$ are quite cumbersome. Therefore we present here only their analytical expressions for the simplified case where the off-diagonal elements of the volatility matrix σ are zero ($\sigma_{12} = \sigma_{21} = 0$, σ_{ki} – matrix elements σ).

$$[\Omega(t-t_0)]_{11} = \frac{1-e^{-2\gamma(t-t_0)}}{4\gamma}(2\sigma_{11}^2 + \sigma_{22}^2) - \frac{e^{-2\gamma(t-t_0)}}{4\gamma}2\gamma(t-t_0)(1+\gamma(t-t_0))\sigma_{22}^2,$$

$$[\Omega(t-t_0)]_{12} = [\Omega(t-t_0)]_{21} = \frac{1-e^{-2\gamma(t-t_0)} - 2\gamma(t-t_0)e^{-2\gamma(t-t_0)}}{4\gamma}\sigma_{22}^2, \quad (9)$$

$$[\Omega(t-t_0)]_{22} = \frac{1-e^{-2\gamma(t-t_0)}}{2\gamma}\sigma_{22}^2.$$

Expression (8) is composed of three terms: the first defines the time-decreasing dependence on the initial state X_0, the second approaches over time to a steady mean θ, and the third term is random process $\xi(t)$ with zero mean and tending eventually to its limiting form covariance matrix $\Omega(t - t_0)$. Limiting matrix $\Omega(t - t_0)$ when $t - t_0 \to \infty$ looks as follows:

$$\Omega(\infty) = \frac{1}{4\gamma}\begin{pmatrix} 2\sigma_{11}^2 + 2\sigma_{12}^2 + 2\sigma_{11}\sigma_{21} + 2\sigma_{12}\sigma_{22} + \sigma_{21}^2 + \sigma_{22}^2 & 2\sigma_{11}\sigma_{21} + 2\sigma_{12}\sigma_{22} + \sigma_{21}^2 + \sigma_{22}^2 \\ 2\sigma_{11}\sigma_{21} + 2\sigma_{12}\sigma_{22} + \sigma_{21}^2 + \sigma_{22}^2 & 2(\sigma_{21}^2 + \sigma_{22}^2) \end{pmatrix}.$$

Note that from properties of matrixes $\Omega(t - t_0)$ follows that components of a vector of state $X(t)$ – the correlated random processes. This correlation between vector components for some fixed time t. However, the serial correlation of the state vectors for different time points t and $t + s$ is important too.

$$\Omega(t, t+s) = E\,[(X(t) - E\,[X(t)])\,(X(t+s) - E\,[X(t+s)])^{\mathrm{T}}] =$$
$$= E\,[\xi(t)\,\xi(t+s)^{\mathrm{T}}] = Cov\,(X(t), X(t+s)).$$

As well as in the previous case, we present this matrix only in the case a diagonal volatility matrix σ. Besides, we will put that $t_0 = 0$.

$$[\Omega(t,t+s)]_{11} = \frac{e^{-\gamma s}}{4\gamma}((1-e^{-2\gamma t})(2\sigma_{11}^2 + (1+\gamma s)\sigma_{22}^2) - 2\gamma t e^{-2\gamma t}(1+(t+s))\sigma_{22}^2),$$

$$[\Omega(t,t+s)]_{12} = \frac{e^{-\gamma s}}{4\gamma}(1-e^{-2\gamma t} - 2\gamma t e^{-2\gamma t})\sigma_{22}^2,$$

$$[\Omega(t,t+s)]_{21} = \frac{e^{-\gamma s}}{4\gamma}((1-e^{-2\gamma t})(1+2\gamma s) - 2\gamma t e^{-2\gamma t})\sigma_{22}^2,$$

$$[\Omega(t,t+s)]_{22} = \frac{e^{-\gamma s}}{2\gamma}(1-e^{-2\gamma t})\sigma_{22}^2.$$

In a stationary case, i.e. at $t_0 \to -\infty$, covariance matrix $Cov\,(X(t), X(t+s))$ does not depend from t and looks like

$$Cov(X(t), X(t+s)) = \frac{e^{-\gamma s}}{4\gamma}\begin{pmatrix} 2\sigma_{11}^2 + (1+\gamma s)\sigma_{22}^2 & \sigma_{22}^2 \\ (1+\gamma s)\sigma_{22}^2 & 2\sigma_{22}^2 \end{pmatrix}.$$

Note that

$$Cov\,(X_2(t), X_2(t+s)) = \frac{\sigma_{22}^2}{2\gamma}e^{-\gamma s}.$$

Therefore the correlation function of components $X_2(t)$ depends only on one parameter γ, i.e.

$$\frac{Cov[X_2(t), X_2(t+s)]}{Var[X_2(t)]} = e^{-\gamma s}.$$

So if this correlation function can be estimated, then it is possible to estimate and parameter γ.

Because processes $X(t)$ and $\xi(t)$ have normal distribution the knowledge of mathematical expectations and covariance matrixes gives their comprehensive probabilistic description.

For NSS model $n = m = 4$ and in this case vector $B(\tau)$ has the form (3) and matrix K is determined by the formula

$$K = \begin{pmatrix} \gamma & -\gamma & 0 & 0 \\ 0 & \gamma & 0 & 0 \\ 0 & 0 & \delta & -\delta \\ 0 & 0 & 0 & \delta \end{pmatrix}.$$

The matrixes $\Theta(t - t_0)$, $\Omega(t - t_0)$, $\Omega(t, t + s)$ will be four-dimensional too. The four-dimensional process $X(t) = (X_1, X_2, X_3, X_4)^{\mathrm{T}}$ will be normal and will have the similar properties as such process of NS model. However the explicit expressions for this case will be considerable more bulky and here not cited.

It is shown [1] that in no-arbitrage affine models between the state variables of financial market $X(t)$ and latent factors of NS model $\{S_t, C_t\}$ there are an accordance

$$S_t = \phi_1 X_1(t) + \phi_2 X_2(t) = r(t), \quad C_t = \phi_1 X_2(t), \quad L_t = -\frac{A(\tau)}{\tau}\bigg|_{\tau = T - t}. \tag{10}$$

So that the factor of level L_t in the explicit form does not depend on current time, and depends only on term to maturity and is not random. A vector $\phi = (\phi_1 \ \phi_2)^{\mathrm{T}}$ is defined by following general limiting property

$$\lim_{\tau \to 0} y(\tau, r) = \lim_{\tau \to 0} \frac{x^{\mathrm{T}} B(\tau) - A(\tau)}{\tau} = x^{\mathrm{T}} \phi = r,$$

which is fulfiled uniformly over x. Therefore $\lim_{\tau \to 0} \frac{-A(\tau)}{\tau} = L(0) = 0$, and $B'(0) = \phi$. r is the short-term (riskless) interest rate. From this and from the expression (1) it is follows that $\phi = B'(0) = (1\ 0)$, i.e. $\phi_1 = 1$, $\phi_2 = 0$.

The slope and curvature factors $\{S_t, C_t\}$ are accurately the state variables $\{X_1(t), X_2(t)\}$. Hence all probability properties of state variables $\{X_1(t), X_2(t)\}$ are properties of latent factors $\{S_t, C_t\}$ also.

Thus for NS model the factor $L(\tau)$ is determinate and vector of factors $\{S_t, C_t\}$ are the normally distributed two-dimensional random variable with the vector of mathematical expectation $\theta = (\theta_1 \ \theta_2)^{\mathrm{T}}$ and the covariance matrix $\Omega(t - t_0)$ with elements that are computed by formulae (9). From this it follows that yield rate $y(\tau)$ and the forward rate $f(\tau)$ have mathematical expectations and variances

$$E[y(t, \tau)] = L(\tau) + \frac{1}{\tau}\theta^{\mathrm{T}} B(\tau), \ E[f(t, \tau)] = \frac{d[\tau L(\tau)]}{d\tau} + \theta^{\mathrm{T}} \frac{d[B(\tau)]}{d\tau}.$$

$$\mathrm{VAR}[Y(T, \tau)] = \frac{1}{\tau^2} B(\tau)^{\mathrm{T}} \Omega(T - T_0) B(\tau), \ \mathrm{VAR}[F(T, \tau)] = \frac{d[B(\tau)]}{d\tau}^{\mathrm{T}} \Omega(T - T_0) \frac{d[B(\tau)]}{d\tau}.$$

These expectations and variances can be computed in explicit form both for NS model and NSS model. Only for NS model the vectors and the matrixes are two-dimensional and for NSS model these vectors and matrixes are four-dimensional. Because of bulkiness these expressions here not cited.

2. Estimation

If to enter designations $\widetilde{S}_t = S_t - \theta_1$, $\widetilde{C}_t = C_t - \theta_2$ then representation (8) can be written down as well in more compact kind:

$$\begin{pmatrix} \widetilde{S}_t \\ \widetilde{C}_t \end{pmatrix} = \begin{pmatrix} \widetilde{S}_s + \gamma(t-s)\widetilde{C}_s \\ \widetilde{C}_s \end{pmatrix} e^{-\gamma(t-s)} + \zeta(t, s) = e^{-\gamma(t-s)} \begin{pmatrix} 1 & \gamma(t-s) \\ 0 & 1 \end{pmatrix} \begin{pmatrix} \widetilde{S}_s \\ \widetilde{C}_s \end{pmatrix} + \zeta(t, s),$$

From representations (8) and (9) follows that if the pair $\{S_s, C_s\}$ is known for $s < t$ then the least squares estimator of pair $\{S_t, C_t\}$ is expression

$$\begin{pmatrix} \hat{S}_t \\ \hat{C}_t \end{pmatrix} = e^{-\gamma(t-s)} \begin{pmatrix} 1 & \gamma(t-s) \\ 0 & 1 \end{pmatrix} \begin{pmatrix} S_s \\ C_s \end{pmatrix} + \begin{pmatrix} 1-e^{-(t-s)\gamma} & -\gamma(t-s)e^{-(t-s)\gamma} \\ 0 & 1-e^{-(t-s)\gamma} \end{pmatrix} \begin{pmatrix} \theta_1 \\ \theta_2 \end{pmatrix}. \quad (11)$$

This estimate is unbiased, and the error covariance matrix is $\Sigma(t-s) = E[\zeta(t, s)\, \zeta(t, s)^{\mathrm{T}}]$.

According to equalities (2) the yield $y_{\mathrm{NS}}(\tau, r)$ at the point time t is determined by expression

$$y_t(\tau) = L_t(\tau) + S_t \frac{1-e^{-\gamma\tau}}{\gamma\tau} + C_t \left(\frac{1-e^{-\gamma\tau}}{\gamma\tau} - e^{-\gamma\tau} \right) =$$

$$= L(\tau) + (b_1(\tau)\ \ b_2(\tau)) \times \begin{pmatrix} S_t \\ C_t \end{pmatrix} = L(\tau) + B(\tau)\, g(t). \quad (12)$$

Here it is designated $b_1(\tau) = \frac{B_1(\tau)}{\tau} = \frac{1-e^{-\gamma\tau}}{\gamma\tau}$, $b_2(\tau) = \frac{B_2(\tau)}{\tau} = \frac{1-e^{-\gamma\tau}}{\gamma\tau} - e^{-\gamma\tau}$, and also considered that the factor $L_t(\tau)$, as mentioned previously, is independent on current time t. Functions $b_1(\tau)$, $b_2(\tau)$ form a row vector $b(\tau) = (b_2(\tau), b_3(\tau))$, depending only on the τ and does not depend on t. Other vector, a vector-column of factors $g(t) = (S_t, C_t)^{\mathrm{T}}$, depends only from t and does not depend from τ.

Let quotes zero-coupon bonds are produced for n maturity terms τ_i, $i = 1, 2..., n$, for business dates t, $t = 1, 2..., N$. Assume also that the quotes $y_t(\tau_i)$ may contain some random errors ε_{it}. Then the structure of the sample can be described by the following relations

$$y_t(\tau_i) = L(\tau_i) + b(\tau_i)\, g(t) + \varepsilon_{it},\ i = 1, 2..., n,\ t = 1, 2..., N. \quad (13)$$

Let's form by means these equalities a n-vector Y_t from the quotations received in date t. Its structure will be represented by means of n-vector $\mathcal{L}$, $(n\times 2)$-matrix $\mathcal{B}$ composed from vectors-row $b(\tau_i)$ of functions of term structure and a n-vector of errors ε_t.

$$\mathcal{B} \equiv \begin{pmatrix} \frac{1-e^{-\gamma\tau_1}}{\gamma\tau_1} & \frac{1-e^{-\gamma\tau_1}}{\gamma\tau_1} - e^{-\gamma\tau_1} \\ \cdots & \cdots \\ \frac{1-e^{-\gamma\tau_n}}{\gamma\tau_n} & \frac{1-e^{-\gamma\tau_n}}{\gamma\tau_n} - e^{-\gamma\tau_n} \end{pmatrix},\ \mathcal{L} \equiv \begin{pmatrix} L(\tau_1) \\ L(\tau_2) \\ \cdots \\ L(\tau_n) \end{pmatrix},\ V \equiv E\,[\varepsilon_t \varepsilon_t^{\mathrm{T}}].$$

So that

$$Y_t = \mathcal{L} + \mathcal{B} g(t) + \varepsilon_t,\ t = 1, 2..., N. \quad (14)$$

From the resulting equation it is possible to estimate a vector $g(t)$ of factors as follows. We multiply equation (14) by the matrix $\mathcal{B}^{\mathrm{T}}$. In this case a multiplier at $g(t)$ becomes square (2×2)-matrix $\mathcal{B}^{\mathrm{T}}\mathcal{B}$, non-singular at $n > 3$. Then we find an inverse matrix $(\mathcal{B}^{\mathrm{T}}\mathcal{B})^{-1}$ and multiply on it at the left the modified equality (14) then the structure of a vector of factors $g(t)$ is easily found in form

$$g(t) = (\mathcal{B}^{\mathrm{T}}\mathcal{B})^{-1}\mathcal{B}^{\mathrm{T}}(Y_t - \mathcal{L}) - (\mathcal{B}^{\mathrm{T}}\mathcal{B})^{-1}\mathcal{B}^{\mathrm{T}}\varepsilon_t.$$

If random errors ε_{it} are independent for various indexes i and t, the vector of errors ε_t has a diagonal covariance matrix $\mathrm{Var}\,[\varepsilon_t] = \mathrm{Diag}\,[\sigma^2_{\varepsilon i} \mid i = 1, 2..., n]$ and a zero mathematical expectation. Therefore the least squares (LS) estimate of a vector of factors

$$\hat{g}(t) = (\mathcal{B}^{\mathrm{T}}\mathcal{B})^{-1}\mathcal{B}^{\mathrm{T}}(Y_t - \mathcal{L}) \quad (15)$$

is unbiased and has a covariance matrix

$$(\mathcal{B}^{\mathrm{T}}\mathcal{B})^{-1}\mathcal{B}^{\mathrm{T}} Y_t\, \mathrm{Var}\,[\varepsilon_t]\, Y_t^{\mathrm{T}} \mathcal{B} (\mathcal{B}^{\mathrm{T}}\mathcal{B})^{-1}.$$

If ε_{it} – normally distributed vector then this estimate will be also maximum likelihood (ML) estimate. Note that matrix $\mathcal{B}$ is known, as its rows are $b(\tau)$ with elements $b_1(\tau) = \frac{1-e^{-\gamma\tau}}{\gamma\tau}$, $b_2(\tau) = \frac{1-e^{-\gamma\tau}}{\gamma\tau} - e^{-\gamma\tau}$. As follows from (10), for determination $L(\tau)$ need the only function $A(\tau)$ with the opposite sign. Equality for function $A(\tau)$ actually is not the equation [1] when functions $B_1(\tau)$ and $B_2(\tau)$ are known and determined by expressions (1). If to take the

elementary case when a volatility matrix σ in the equation (9) is constant and diagonal, matrix K is set in the form of (7), and risk market prices λ – constants then the derivative of function $A(\tau)$ will be such [1]

$$-\frac{dA(\tau)}{d\tau} = \frac{d[\tau L(\tau)]}{d\tau} = (\gamma\theta_1 - \gamma\theta_2 - \sigma_1\lambda_1)\frac{1-e^{-\gamma\tau}}{\gamma} +$$
$$+ (\gamma\theta_2 - \sigma_2\lambda_2)\left(\frac{1-e^{-\gamma\tau}}{\gamma} - \tau e^{-\gamma\tau}\right) -$$
$$- \frac{\sigma_1^2}{2}\left(\frac{1-e^{-\gamma\tau}}{\gamma}\right)^2 - \frac{\sigma_2^2}{2}\left(\frac{1-e^{-\gamma\tau}}{\gamma} - \tau e^{-\gamma\tau}\right)^2. \quad (16)$$

Thus, matrix $\mathcal{B}$ in the ratio (15) for an estimation of a vector of factors $\hat{g}(t)$ can be specified analytically, however integration of expression (16) leads to bulky expression.

Note that in representation (12) $L_t(\tau) \equiv L(\tau)$ on the one hand does not depend on t, and on other – is determined analytically. Therefore this representation can be rewritten as

$$y_t(\tau) - L(\tau) = S_t\frac{1-e^{-\gamma\tau}}{\gamma\tau} + C_t\left(\frac{1-e^{-\gamma\tau}}{\gamma\tau} - e^{-\gamma\tau}\right) = (b_2(\tau), b_3(\tau)) \times \begin{pmatrix} S_t \\ C_t \end{pmatrix}.$$

Let $\tilde{y}_t(\tau) \equiv y_t(\tau) - L(\tau)$. Then the vector equation of observations (14) will be transformed to a form

$$\tilde{Y}_t = \mathcal{B}\, g(t) + \varepsilon_t, \quad t = 1, 2..., N. \quad (17)$$

The relations obtained allow us to construct the optimal mean-squares procedure of estimation the variables $\{S_t, C_t\}$, i.e. sequentially over time to estimate the yield curves in the form (12). This can be done by using a Kalman filter. Suppose that the yield to maturity (13) is listed over regular intervals Δ, so that at each time point $t = k\Delta$, $k = 0, 1, 2...$, is declared n returns $\{y_t(\tau_i), i = 1, 2..., n\}$.

Without stopping on the details of the output of the Kalman filter [5], we present the structure that adapted to our problem.

Let

$$F = e^{-\gamma\Delta}\begin{pmatrix} 1 & \gamma\Delta \\ 0 & 1 \end{pmatrix}.$$

Then the estimate (11) can be written down in a form

$$\begin{pmatrix} \hat{S}_t \\ \hat{C}_t \end{pmatrix}^- = F\begin{pmatrix} \hat{S}_{t-\Delta} \\ \hat{C}_{t-\Delta} \end{pmatrix} + (I - F)\begin{pmatrix} \bar{S} \\ \bar{C} \end{pmatrix},$$

where I – the identity matrix, and a minus in the top index means that the estimation is determined under the information which are available till the previous moment $t - \Delta$ inclusive. A covariance matrix P_t^- of this estimation satisfies to a relation

$$P_t^- = F\, P_{t-\Delta}\, F^{\mathrm{T}} + \Sigma.$$

Here availability of a minus in the top index P_t^- means the same, as in previous, and absence of such index in $P_{t-\Delta}$ means that the covariance matrix was determined under the information which are available till the moment, specified in the bottom index inclusive. Σ – a covariance matrix $E[\zeta(t, t-\Delta)\,\zeta(t, t-\Delta)^{\mathrm{T}}]$.

Suppose further that the vector equation of observations has the form (17)

Let's enter into matrix consideration

$$\mathcal{K}_t = P_t^-\,\mathcal{B}^{\mathrm{T}}\,(\mathcal{B}\,P_t^-\,\mathcal{B}^{\mathrm{T}} + V)^{-1},\ P_t = (I - \mathcal{K}_t\,\mathcal{B})\,P_t^-.$$

Matrix $\mathcal{K}_t$ determines the optimum compromise of contributions in an estimation of variables $\{S_t, C_t\}$: by observation Y_t arrived at the time point t and by estimates of these variables under the information, which are available till the time point $t - \Delta$ inclusive. The covariance matrix P_t is necessary for realization of the following iteration of algorithm at the point of time $t + \Delta$. Note that as matrixes F, Σ, H, V, $\mathcal{L}$ do not change with time and assumed to be known prior to the beginning of procedure of estimation, then sequence of matrixes P_t^-, $\mathcal{K}_t$, P_t can be calculated in advance prior to the beginning of recurrent procedure of estimation. The final estimate of variables $\{S_t, C_t\}$ at time t is calculated by the formula

$$\begin{pmatrix}\hat{S}_t\\ \hat{C}_t\end{pmatrix}=\begin{pmatrix}\hat{S}_t\\ \hat{C}_t\end{pmatrix}^{-}+\mathcal{K}_t\,\widetilde{Y}_t-\mathcal{K}_t\mathcal{B}\begin{pmatrix}\hat{S}_t\\ \hat{C}_t\end{pmatrix}^{-}=\mathcal{K}_t\widetilde{Y}_t+(I-\mathcal{K}_t\mathcal{B})\begin{pmatrix}\hat{S}_t\\ \hat{C}_t\end{pmatrix}^{-}=$$

$$=\mathcal{K}_t(Y_t-\mathcal{L})+(I-\mathcal{K}_t\mathcal{B})\left[F\begin{pmatrix}\hat{S}_{t-\Delta}\\ \hat{C}_{t-\Delta}\end{pmatrix}+(I-F)\begin{pmatrix}\overline{S}\\ \overline{C}\end{pmatrix}\right].$$

This formula allows recurrently during the receiving of observations Y_t compute the best mean-square estimate of factors $\{S_t, C_t\}$. Unfortunately, for realization of such procedure it is necessary to know all model parameters, i.e. γ, θ_1, θ_2, λ_1, λ_2, $\{\sigma_{ik}\}$, or to implement procedure of their estimation, which is a separate problem. An initial vector of estimations of factors $\{S_0, C_0\}$ can be obtained using (15) on the initial observation Y_0:

$$\begin{pmatrix}\hat{S}_0\\ \hat{C}_0\end{pmatrix}=(\mathcal{B}^{\mathrm{T}}\mathcal{B})^{-1}\mathcal{B}^{\mathrm{T}}(Y_0-\mathcal{L}).$$

Conclusion

The description of the Nelson – Siegel (NS) yield models and the Nelson – Siegel – Svensson (NSS) yield models as a traditional multi-dimensional affine yield models is presented. NS model it turns out the two-factor model, in the case of NSS model – the four-factor. This description differs from the representations of F. Diebold and G. Rudebusch [6] so that the dimension of the models is reduced by one, which simplifies the calculations. It was found that the Nelson – Siegel latent variables coincide with the state variables of the traditional models. The explicit representation of the probability distribution and the first two moments of these «latent variables» are obtained. It is computed the expectations and covariance matrices of interest rates of the yield to maturity and forward rates. It is formulated the procedure of latent variables recursion estimation based on the Kalman filter.

REFERENCES

1. *Medvedev G.A.* On the Nelson-Siegel-Svensson no-arbitrage yield curve models // Вестник Томского государственного университета. Управление, вычислительная техника и информатика. 2015. № 3 (32). С. 44–55.
2. *Nelson C.R., Siegel A.F.* Parsimonious modelling of yield curves // Journal of Business. 1987. V. 60. P. 473–489.
3. *Svensson L.* Estimating forward interest rates with the extended Nelson-Siegel method // Quarterly Review. Sveriges Riksbank. 1995. No. 3. P. 13–26.
4. *Kloeden E., Platen E.* Numerical solution of stochastic differential equations. Berlin : Springer-Verlag, 1992. 668 p.
5. *Welch G., Bishop G.* An Introduction to the Kalman Filter // UNC-Chapel Hill: TR 95-041, 2006. 16 p.
6. *Diebold F., Rudebusch G.* Yield Curve Modeling and Forecasting. Princeton: Princeton University Press, 2013. 159 p.

Medvedev Gennady Alekseevich, doctor of science, professor. E-mail: MedvedevGA@cosmostv.by
Belarusian State University, Republic of Belarus

Поступила в редакцию 16 июня 2015 г.

Медведев Геннадий А. (Белорусский государственный университет, Республика Беларусь)

Доходности Нельсона – Сигеля – Свенссона. Вероятностные свойства и оценивание.

Ключевые слова: временная структура; кривая доходности; аффинные модели; модель доходности Нельсона – Сигеля – Свенссона; фильтр Калмана.

DOI: 10.17223/19988605/33/5

Рассмотрены вероятностные свойства процентной ставки доходности, порождаемой моделями Нельсона – Сигеля (NS) и Нельсона – Сигеля – Свенссона (NSS). Описание моделей доходностей Нельсона – Сигеля и Нельсона – Сигеля – Свенссона представлено как изложение традиционных многомерных аффинных моделей доходности. Показано, что модель Нельсона – Сигеля практически не отличается от традиционной двухфакторной модели аффинной доходности с волатильностью, не зависящей от переменных состояния рынка, соответственно, модель Нельсона – Сигеля – Свенссона – от четырехфакторной такой модели. Такое описание отличается от известных представлений Ф. Диболда и Г. Радебуша тем, что размерность представленных в статье моделей уменьшается на единицу, что приводит к упрощению вычислений. Выяснено, что скрытые

переменные Нельсона – Сигеля полностью совпадают с переменными состояния традиционной модели. Эти модели порождают процентные ставки доходности до погашения и форвардной доходности с нормальным распределением, для которого найдены математические ожидания и ковариационные матрицы в явном виде. Для оценивания значений ставок доходности до погашения в текущем времени сформулирована рекуррентная процедура, основанная на применении фильтра Калмана.

REFERENCES

1. Medvedev G.A. (2015) On the Nelson-Siegel-Svensson no-arbitrage yield curve models. *Vestnik Tomskogo gosudarstvennogo universiteta. Upravlenie, vychislitel'naya tekhnika i informatika – Tomsk State University Journal of Control and Computer Science.* 3(32). pp. 44–55. (In Russian) DOI: 10.17223/19988605/32/5
2. Nelson, C.R. & Siegel, A.F. (1987) Parsimonious modelling of yield curves. *Journal of Business.* 60. pp. 473–489.
3. Svensson, L. (1995) Estimating forward interest rates with the extended Nelson-Siegel method. *Quarterly Review. Sveriges Riksbank,* 3. pp. 13–26. DOI: 10.3386/w4871
4. Kloeden, E. & Platen, E. (1992) *Numerical solution of stochastic differential equations*. Berlin: Springer-Verlag.
5. Welch, G. & Bishop, G. (2006) *An Introduction to the Kalman Filter*. UNC-Chapel Hill: TR 95-041.
6. Diebold, F. & Rudebusch, G. (2013) *Yield Curve Modeling and Forecasting*. Princeton: Princeton University Press.

ИНФОРМАТИКА И ПРОГРАММИРОВАНИЕ

УДК 004.056; 004.4
DOI: 10.17223/19988605/33/6

В.С. Оладько

ПРОГРАММНЫЙ КОМПЛЕКС ДЛЯ ОЦЕНКИ УРОВНЯ ЗАЩИЩЕННОСТИ СИСТЕМ ЭЛЕКТРОННОЙ КОММЕРЦИИ

Рассматриваются задача оценки уровня защищенности информации и ресурсов в системах электронной коммерции; основные модели и ресурсы систем электронной коммерции. Выделены типовые угрозы и механизмы защиты в системах электронной коммерции. Предложен подход к оценке уровня защищённости в системах электронной коммерции вида B2B и B2E. Разработан программный комплекс, реализующий предложенный подход. Проведены экспериментальные исследования.
Ключевые слова: электронный заместитель; риск; платежная система; угроза.

В настоящее время одной из наиболее активно развивающихся отраслей экономики и информационных технологий является электронная коммерция. По данным Synovate Comcon, в настоящий момент 58% пользователей Рунета прибегают к услугам электронной коммерции. При этом с каждым годом увеличивается число инцидентов, связанных с хищением денежных средств и нарушением информационной безопасности в системах электронной коммерции (СЭК); по данным [1], ущерб от подобных инцидентов колеблется в диапазоне от 250 тысяч до 60 млн руб. Следовательно, актуальным направлением являются исследования, связанные с анализом защищенности СЭК от угроз различной природы и синтезом наиболее эффективной системы защиты.

1. Проблемы безопасности системы электронной коммерции

В соответствии с [2] электронная коммерция – это форма коммерческой деятельности, осуществляемая полностью или частично в виртуальной среде, при которой информационные или транзакционные взаимодействия происходит на основе применения информационно-коммуникационных технологий. Анализ источников [2, 3] показывает, что наиболее востребована электронная коммерция в следующих отраслях:

– высокотехнологичное производство;
– финансовый сервис;
– розничная торговля;
– телекоммуникации;
– оптовая торговля;
– государственные услуги и закупки;
– транспорт.

В связи с этим в СЭК обрабатывается и храниться большое количество данных различной категории: платежные данные, электронные деньги, данные о транзакциях, а также персональные и идентификационные данные пользователей. В зависимости от способа функционирования и области использования СЭК реализуются с помощью следующих моделей, представленных в виде схемы на рис. 1. При этом наиболее распространёнными являются СЭК вида B2B и B2E, на долю которых, по данным [4], приходится 35 и 48% соответственно. Именно эти виды СЭК и будут рассматриваться автором в данной работе.

В связи с распределённой архитектурой СЭК, наличием подключения к глобальным сетям типа Интернет, круглосуточной доступностью сервисов и данных, а также большого количества пользователей-клиентов в процессе своего функционирования система подвергается ряду угроз и деструктивных воздействий как случайного характера, так и инициированных злоумышленником.

Основными источниками угроз СЭК являются:

1) непреднамеренные угрозы, вызванные стихийными бедствиями и техногенными катастрофами случайного характера, в ходе которых может быть нарушена непрерывность бизнес-процессов СЭК, доступность данных и сервисов, целостность данных;

2) ошибки пользователей и обслуживающего персонала СЭК;

3) преднамеренные действия злоумышленника, направленные на нарушение таких составляющих информационной безопасности, как доступность, целостность и конфиденциальность информации, циркулирующей в СЭК.

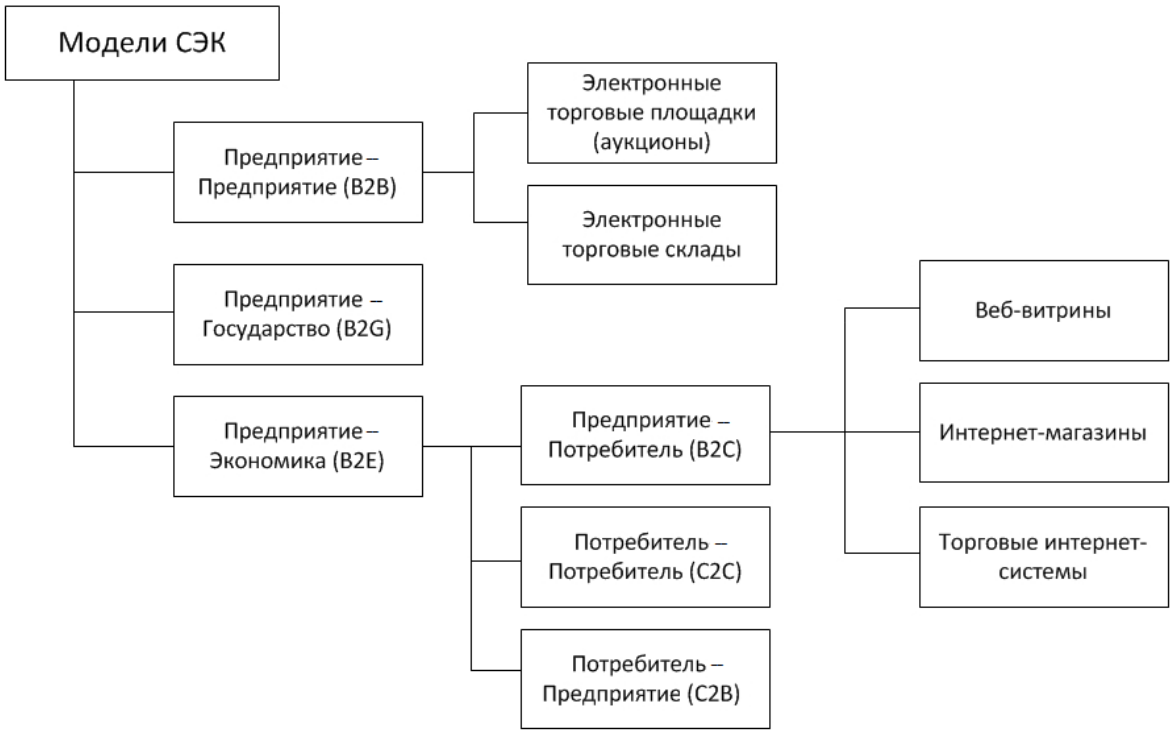


Рис. 1. Классификация моделей электронной коммерции

Согласно данным [5, 6] наибольшую опасность представляет собой класс угроз, связанных с действиями злоумышленников, поскольку именно этот класс несет наибольшие финансовые и репутационные риски. При этом главным объектом воздействий злоумышленника в СЭК являются финансовые средства, их электронные заместители, платежные данные и информация о транзакциях. По отношению к данным объектам злоумышленник преследует следующие цели:

– получение доступа к финансовым средствам и реквизитам для последующего использования;

– похищение финансовых средств и электронных заместителей;

– внедрение фальшивых финансовых средств;

– нарушение доступности сервисов СЭК и непрерывности бизнес-процессов;

– несанкционированный доступ к идентификационным данным пользователей СЭК;

– мошенничество, фишинг;

– несанкционированная модификация платежных данных, реквизитов и идентификационных данных.

Для эффективного противодействия большому числу угроз безопасности СЭК и обеспечения безопасности всех участников электронных платежей должны применяться различные средства и методы

защиты, правила применения и состав которых описываются в стандартах и рекомендациях регулирующих органов. Анализ литературных источников [7–11] показывает, что в качестве основных регулирующих органов в области безопасности платежных систем и СЭК в Российской Федерации выступают ЦБ РФ, ФСТЭК России и ФСБ России, которые, в соответствии с ФЗ № 161-ФЗ, образуют три уровня регулирования (рис. 2).

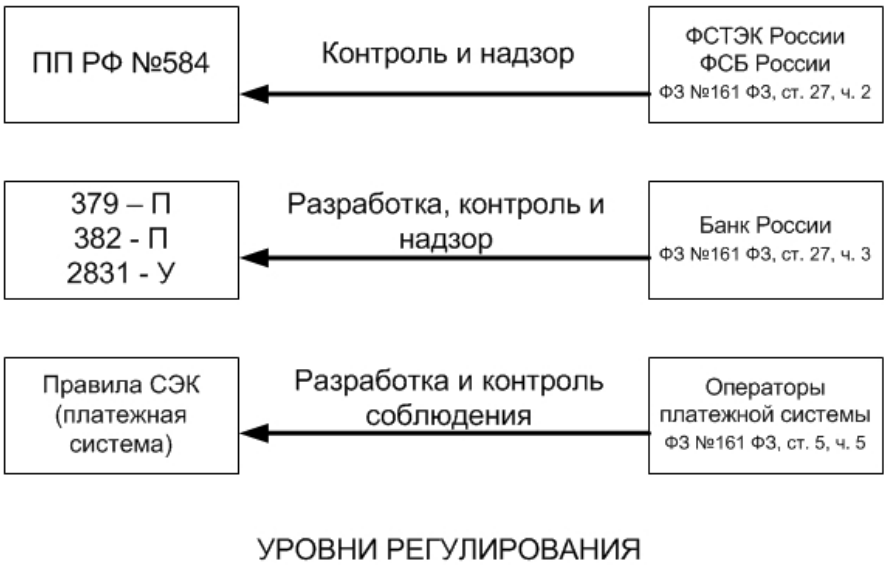


Рис. 2. Уровни регулирования безопасности в СЭК (область платежных систем)

В соответствии с требованиями регуляторов в СЭК должны применяться организационные и технические меры по защите информации. Данные меры по защите информации, с учетом использования в СЭК сетей общего пользования, должны предусматривать применение криптографических средств защиты информации, средств межсетевого экранирования, антивирусной защиты, обнаружения вторжений и анализа защищенности.

При этом средства анализа защищенности актуально применять на различных этапах жизненного цикла СЭК:

– на этапах проектирования и разработки системы защиты СЭК для выбора наиболее рационального и эффективного состава средств защиты;

– на этапе сопровождения с целью регулярного мониторинга и аудита безопасности СЭК и проверки соответствия системы защиты установленным регуляторами требованиям.

Таким образом, для исследования уровня защищенности СЭК и выбора наиболее рационального состава средств защиты информации в СЭК, позволяющих снизить риски от потенциальных угроз в пределах допустимого, актуально разработать процедуру оценки защищенности СЭК и автоматизировать ее с помощью программного комплекса.

2. Процедура оценки защищенности СЭК

Анализ литературных источников [5, 12, 13] показывает, что защищенность системы и риск нарушения информационной безопасности – два понятия, тесно связанных между собой. В частности, в работе [12] риск определяется как вероятный ущерб, который зависит от защищенности системы. Таким образом, риск – это экономический эквивалент защищенности системы от реализации возможных угроз. Следовательно, при проведении оценки защищенности СЭК целесообразно в качестве базового показателя использовать величину общего риска от реализации актуальных для СЭК угроз безопасности. А поскольку анализ рисков связан с оптимизационной задачей, заключающейся в поиске баланса между вложениями в систему защиты информации и возможным ущербом, то при оценке защищенности и выработке корректирующих рекомендаций будет решаться задача поиска наиболее рационального состава средств защиты, применение которого на практике позволит не только снизить общий уровень

риска, но и повысить защищенность СЭК. В данной работе будет использоваться качественный подход к оценке защищенности, в соответствии с которым защищенность СЭК будет оцениваться шкалой разбитой на три уровня: L = {низкий, средний, высокий}. Правила определения принадлежности защищенности СЭК к одному из трех уровней описаны автором в работе [14]. Таким образом, процедура оценки качественного уровня защищенности СЭК описывается в виде следующей последовательности шагов:

– составление модели СЭК с указанием объектов защиты, используемых средств защиты и пользователей СЭК;

– составление модели угроз и оценка рисков;

– оценка защищенности СЭК на основе данных об используемых средствах защиты и рисках от реализации актуальных угроз;

– формирование отчета о результатах оценки защищённости и выдача рекомендаций по реконфигурации системы защиты СЭК в случае необходимости.

На рис. 3 в нотации IDEF0 представлена разработанная функциональная модель оценки защищенности СЭК.

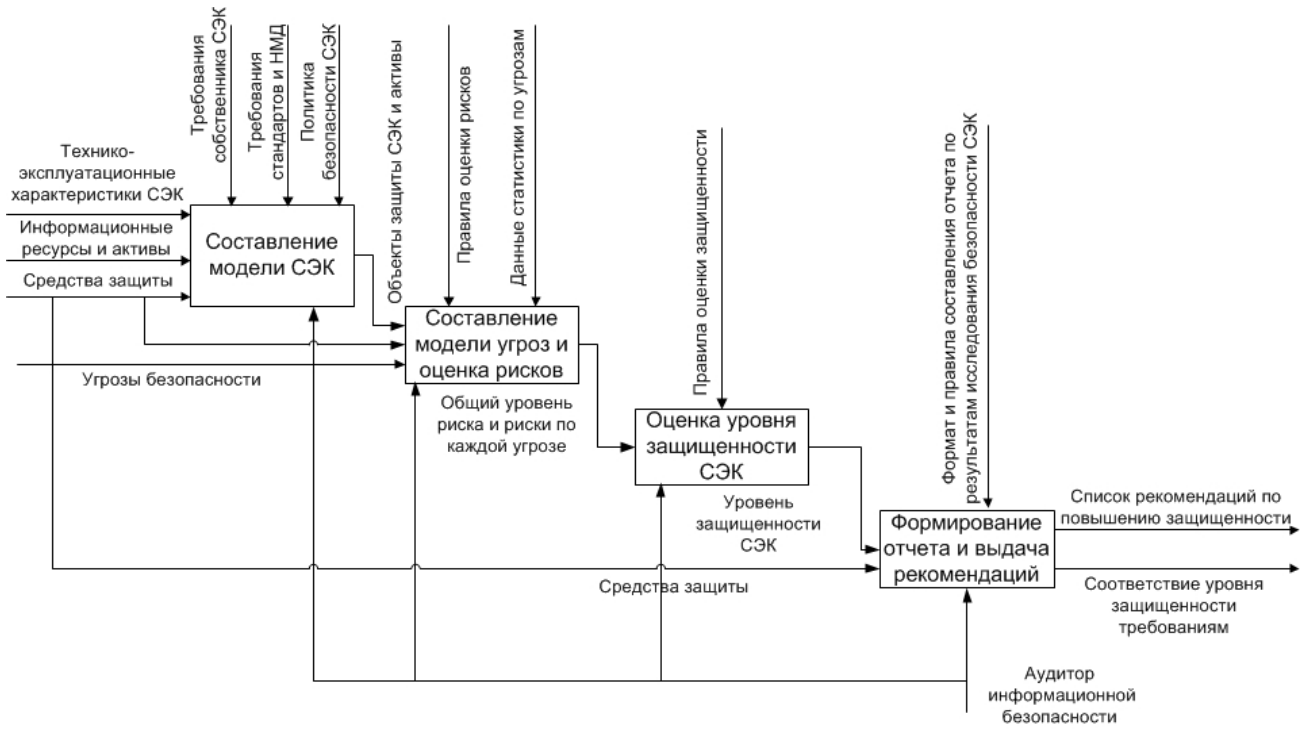


Рис. 3. Функциональная контекстная диаграмма оценки защищенности СЭК в нотации IDEF0

Входными данными являются:

– вид и технико-эксплуатационные характеристики СЭК;

– информационные ресурсы, активы СЭК и платежные данные;

– используемые средства и механизмы защиты;

– список угроз для СЭК;

– список злоумышленников;

– требования к допустимому уровню риска и уровню защищенности СЭК.

Выходными данными являются:

– оцененный уровень защищенности СЭК;

– соответствие или несоответствие оцененного уровня защищенности СЭК требованиям безопасности;

– список рекомендаций по повышению защищенности СЭК.

3. Программный комплекс оценки уровня защищенности СЭК

Для автоматизации описанных выше функций был разработан программный комплекс, архитектура которого представлена на рис. 4. В архитектуре схожие функции были объединены в отдельные модули с целью обеспечения их более эффективного выполнения.

Рис. 4. Архитектура программного комплекса оценки защищенности СЭК

Модуль сбора данных об СЭК предназначен для сбора информации о технико-эксплуатационных характеристиках СЭК, ее активах и объектах, которые подлежат защите; ввода требований к защищенности СЭК и уровню допустимого риска. Информация вводится ответственным за информационную безопасность в СЭК лицом.

Модуль выбора средств защиты предназначен для выбора из списка возможных средств и механизмов защиты, которые используются в исследуемой СЭК для обеспечения безопасности.

Модуль выбора угроз и оценки рисков СЭК предназначен для составления модели угроз для СЭК, установки для каждой угрозы вероятности реализации и потенциального ущерба и расчета риска по каждой угрозе и общего риска.

Модуль оценки защищенности предназначен для оценки уровня защищенности СЭК.

Модуль формирования рекомендаций и отчета обеспечивает взаимодействие между другими модулями и на основании данных об уровне защищенности СЭК вырабатывает комплекс мероприятий – рекомендаций по защите от каждой из актуальных угроз, определяет соответствие между рассчитанным уровнем защищенности и допустимым и выносит решение о защищенности или незащищённости СЭК. И если СЭК имеет низкую защищенность, то выводит список рекомендуемых средств и механизмов защиты. Отчет о результате оценки выводится на экран, а также может быть сохранен в формате .docx.

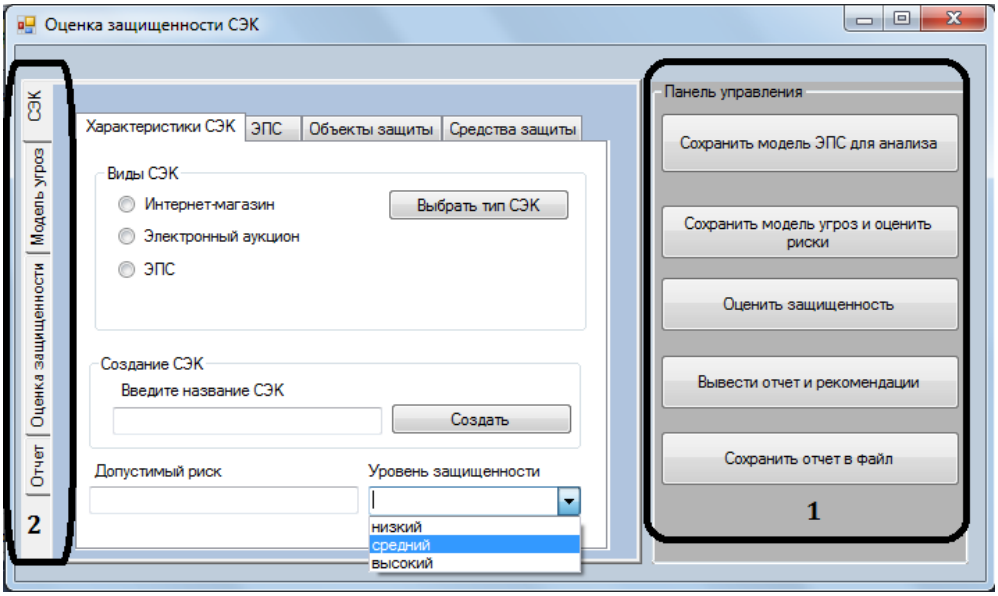


Рис. 5. Пользовательский интерфейс программного комплекса оценки защищенности СЭК (экранная копия)

Пользовательский интерфейс имеет графический вид и предназначен для ввода данных, вывода результатов и организации взаимодействия пользователя с программой. Состоит из блока управления программным комплексом (рис. 5, область *1*) и четырёх функциональных вкладок (рис. 5, область *2*), реализующих основные задачи оценки защищенности СЭК: составление модели СЭК, составление модели угроз для исследуемой СЭК, оценки защищенности СЭК и вывода отчета о результатах оценки защищенности на экран.

Разработанный программный комплекс предназначен для специалиста по информационной безопасности и/или другого ответственного за защиту информации в СЭК лица. Может использоваться в качестве вспомогательного средства при проведении внутреннего аудита информационной безопасности СЭК.

4. Экспериментальные исследования

Эксперименты направлены на исследование результатов оценки защищенности и возможности использования, предложенных моделью рекомендаций для повышения общего уровня защищенности СЭК и снижения рисков. В рамках данных экспериментов было проведено исследование защищенности 10 тестовых СЭК, результаты которых представлены в виде гистограммы на рис. 6.

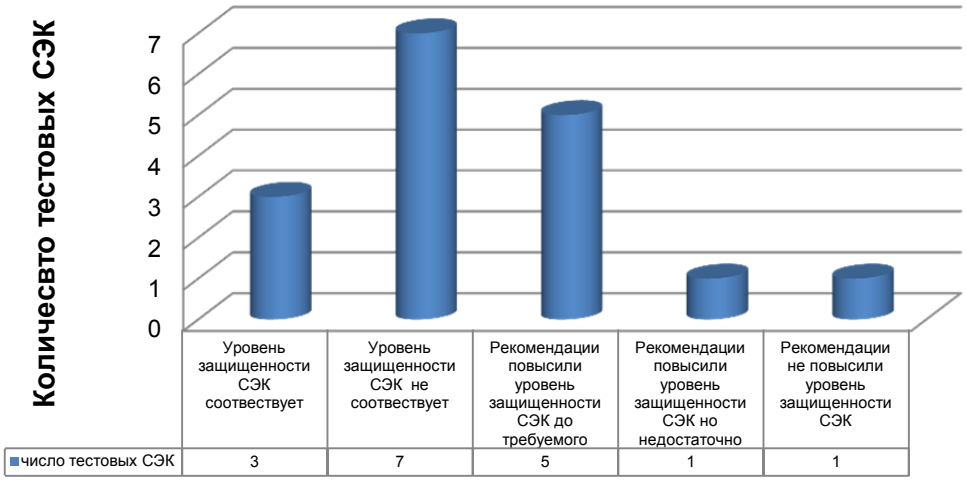


Рис. 6. Распределение результатов оценки защищенности для 10 тестовых СЭК различного вида.

Анализ полученных результатов позволяет сделать вывод, что в 86% случаях рекомендации, предложенные программным комплексом в ходе оценки, позволяют повысить уровень защищенности СЭК, при этом в 71% этого повышения достаточно для достижения требуемой защищенности. Таким образом, можно сделать вывод о возможности применения рекомендаций модели для повышения общей защищенности СЭК различного вида.

Заключение

Была разработана процедура оценки защищенности СЭК. В нотации IDEF0 описаны функции и этапы процедуры оценки защищенности СЭК, выделены входные и выходные данные. Описаны архитектура и пользовательский интерфейс программного комплекса, автоматизирующего предложенную процедуру. Разработанный программный комплекс может применяться в учебном процессе на лабораторном практикуме для обучения студентов направления «информационная безопасность» и на практике при проектировании систем электронной коммерции и в процессе функционирования для контроля над состоянием их безопасности.

ЛИТЕРАТУРА

1. *Абдеева З.Р.* Проблемы безопасности электронной коммерции в сети Интернет // Проблемы современной экономики. 2012. № 1. С. 172–175.
2. *Агафонова А.Н.* Информационный сервис в Интернет-экономике. М. : БИБЛИО-ГЛОБУС, 2014. 152 с.
3. *Ефимкин Н.В.* Электронная коммерция как современная форма торговли // Экономика и социум. 2015. № 2. URL: http://www.iupr.ru/domains_data/files/zurnal_15/Efimkin%20N.V.%20%28Informacionnye%20i%20kommunikativnye%20tehnologii%29.doc%20%281%29.pdf (дата обращения: 24.08.2015).
4. *Вольфсон М.* За прошлый год глобальный e-commerce вырос на 20% // E-business. URL: http://el-business.ucoz.ru/news/za_proshlyj_god_globalnyj_e_commerce_vyros_na_20/2015-03-14-167 (дата обращения: 01.04.2015).
5. *Яндыбаева Э.Э. Машкина И.В.* Оценка актуальности угроз информационной безопасности в информационной системе электронной торговой площадки // Безопасность информационных технологий. 2014. № 1. С. 41–44.
6. *Оладько В.С.* Модель действия злоумышленника в системах электронной коммерции // Международный научно-исследовательский журнал. 2015. № 7 (38). С. 83–85.
7. *Федеральный закон* от 27.06.2011 № 161-ФЗ (ред. от 29.12.2014) «О национальной платежной системе» (с изм. и доп., вступ. в силу с 01.03.2015) (27 июня 2011 г.) // Консультант Плюс. URL: http://www.consultant.ru/document/cons_doc_LAW_173643 (дата обращения: 15.03.2015).
8. *Максимов В.* № 161-ФЗ «О национальной платежной системе»: роли, правила, требования Банка России к защите информации, сроки исполнения, последствия // Андек. URL: http://www.andek.ru/ehkspertiza/banki/nacionalnaya-platezhnaya-sistema/ (дата обращения: 15.03.2015).
9. *Положение* Банка России от 09.06.2012 г. № 382-П: «Положение о требованиях к обеспечению защиты информации при осуществлении переводов денежных средств и о порядке осуществления Банком России контроля за соблюдением требований к обеспечению защиты информации при осуществлении переводов денежных средств» // Консультант Плюс. URL: http://www.consultant.ru/document/cons_doc_LAW_131473/ (дата обращения: 24.08.2015).
10. *Указание* Банка России от 09.06.2012 № 2831-У «Об отчетности по обеспечению защиты информации при осуществлении переводов денежных средств операторов платежных систем, операторов услуг платежной инфраструктуры, операторов по переводу денежных средств» // Консультант Плюс. URL: http://www.consultant.ru/document/cons_doc_LAW_131471/ (дата обращения: 23.08.2015).
11. *Положение* Банка России от 31.05.2012 г. № 379-П: «Положение о бесперебойности функционирования платежных систем и анализе рисков в платежных системах» // Консультант Плюс. URL: http://base.consultant.ru/cons/CGI/online.cgi?req=doc;base=LAW;n=167370 (дата обращения: 23.08.2015).
12. *Ерохин С.* Методика оценки рисков нарушения информационной безопасности // Информационная безопасность. URL: http://esstm.blogspot.com/2011/09/microsoft.html (дата обращения: 16.04.13).
13. *Банокин П.И., Вичугов В.Н.* Реализация программной системы для предотвращения внутренних утечек корпоративных данных // Вестник Томского государственного университета. Управление, вычислительная техника и информатика. 2014. № 1 (26). С. 84–88.
14. *Аткина В.С., Оладько А.Ю.* Модель оценки безопасности систем электронной коммерции // Вестник компьютерных и информационных технологий. 2015. № 3. С. 33–37.

Оладько Владлена Сергеевна, канд. техн. наук, доцент. E-mail: oladko.vs@yandex.ru
Волгоградский государственный университет

Поступила в редакцию 7 сентября 2015 г.

Oladko Vladlena S. (Volgograd State University, Russian Federation).

The program is assessing the level of security of e-commerce.

Keywords: electronic substituent; risk; payment system; threat.

DOI: 10.17223/19988605/33/6

This article deals with the problem of information security in data processing system of e-commerce. The author highlights the scope and the main models of e-commerce. Models type B2E and B2B are selected for the study of security. The main objects that are processed in the system of e-commerce are the payment information, finance, electronic substituents, transaction data, personal data, user's identification and authentication data.

The main causes of violations of data security and business continuity in the systems of electronic commerce are analyzed. These are the threat of unintentional and intentional. Unintentional threats include natural disasters and man-made disasters, user errors and software failures. Intentional threats relate to the actions of an insider or an external attacker or group of attackers. The main objects of the impact of the attacker's e-commerce systems are electronic substituents financial resources and personal data. Regulatory requirements to protect payment data and e-commerce systems from the threats of a different nature are allocated. It is concluded that in addition to the use of information security tools necessary to carry out regular monitoring of the security of e-commerce and data processed by it.

It is proposed to consider security as a measure of the quality associated with the value of the residual risk by using means of information protection. Security of e-commerce is divided into three levels and describes the scale of L = {low, medium, high}. It is designed and described how the qualitative assessment of the level of security of e-commerce. Its main steps are:

– drafting a model of e-commerce with an indication of the objects of protection, protection systems and users;

– drafting threat models and risk assessment;

– assessment of the security of e-commerce on the basis of data on the means used to protect and risks of the actual threats;

– forming a report on the evaluation of security and making recommendations on the reconfiguration of the system of protection in case of need.

Functional assessment model of security is composed. The modular architecture of software for the evaluation of the level of protection has been developed and described. The main function modules include:

– performance data collection of e-commerce and the input requirements of security and the level of acceptable risk;

– selection module remedies designed to select from a list of possible means and mechanisms of protection that are used in the system to ensure safety;

– selection module threat and risk assessment system for e-commerce for drawing up the threat model, the settings for each threat and probability of potential damage and the calculation of risk for each threat, and the overall risk;

– module security assessment is designed to evaluate the quality level of system security;

– generation unit and the recommendations of the report.

The graphical user interface is presented. The experimental results of test models of e-commerce systems are described. The author concluded that performance of software and applications as a tool for the assessment of the current security systems, e-commerce and make recommendations for its increase.

REFERENCES

1. Abdeeva, Z.R. (2012) Electronic commerce in the Internet: problems of security. *Problemy sovremennoy ekonomiki – Problems of modern economics*. 1(41). pp. 172-175. (In Russian).
2. Agafonova, A.N. (2014) *Informatsionnyy servis v Internet-ekonomike* [Information service in the Internet economy]. Moscow: BIBLIO-GLOBUS.
3. Efimkin, N.V. (2015) Elektronnaya kommertsiya kak sovremennaya forma torgovli [E-commerce as a modern form of trading]. *Ekonomika i sotsium*. 2. [Online] Available from: http://www.iupr.ru/domains_data/files/zurnal_15/Efimkin%20N.V.%20%28Informacionnye%20i%20kommunikativnye%20tehnologii%29.doc%20%281%29.pdf. (Accessed: 24th August 2015).
4. Wolfson, M. (2015) *Za proshlyy god global'nyy e-commerce vyros na 20%* [Over the past year the global e-commerce has grown by 20%]. [Online] Available from: http://el-business.ucoz.ru/news/za_proshlyj_god_globalnyj_e_commerce_vyros_na_20/2015-03-14-167. (Accessed: 1st April 2015).
5. Yandybaeva, E.E. & Mashkina, I.V. (2014) The Relevance Assessment of Information Security Threats in the Information System of Electronic Trading Platform. *Bezopasnost' informatsionnykh tekhnologiy*. 1. pp. 41-44. (In Russian).
6. Olad'ko, V.S. (2015) Model' deystviya zloumyshlennika v sistemakh elektronnoy kommertsii [Model action attacker in electronic commerce systems]. *Mezhdunarodnyy nauchno-issledovatel'skiy zhurnal – International Research Journal*. 7(38). pp. 83-85.
7. Russian Federation. (2011) *Federal'nyy zakon ot 27.06.2011 № 161-FZ (red. ot 29.12.2014) "O natsional'noy platezhnoy sisteme" (s izm. i dop., vstup. v silu s 01.03.2015) (27 iyunya 2011 g.)* [Federal Law N 161-FZ of June 27, 2011 (ed. by December 29, 2014) "On the national payment system" (rev. and ext., joined. in force from March 1, 2015) (27 June 2011)]. [Online] Available from: http://www.consultant.ru/document/cons_doc_LAW_173643. (Accessed: 15th March 2015).
8. Maksimov, V. (n.d.) *№ 161-FZ "O natsional'noy platezhnoy sisteme": roli, pravila, trebovaniya Banka Rossii k zashchite informatsii, sroki ispolneniya, posledstviya* [N161-FZ "On the national payment system": roles, rules, requirements of the Bank of Russia to the protection of information, deadlines, and consequences]. [Online] Available from: http://www.andek.ru/ehkspertiza/banki/nacionalnaya-platezhnaya-sistema/. (Accessed: 15th March 2015).
9. The Bank of Russia. (2012) *Polozhenie Banka Rossii ot 09.06.2012 g. № 382-P: "Polozhenie o trebovaniyakh k obespecheniyu zashchity informatsii pri osushchestvlenii perevodov denezhnykh sredstv i o poryadke osushchestvleniya Bankom Rossii kontrolya za soblyudeniem trebovaniy k obespecheniyu zashchity informatsii pri osushchestvlenii perevodov denezhnykh sredstv"* [The Bank of Russia on June 9, 2012, the N382-P: "Regulations on requirements to ensure the protection of information in the implementation of remittances and the exercise by the Bank of Russia control over compliance with the requirements for the protection of inforation in the implementation of remittances"]. [Online] Available from: http://www.consultant.ru/document/cons_doc_LAW_131473/. (Accessed: 24th August 2015).
10. The Bank of Russia. (2012) *Ukazanie Banka Rossii ot 09.06.2012 № 2831-U "Ob otchetnosti po obespecheniyu zashchity informatsii pri osushchestvlenii perevodov denezhnykh sredstv operatorov platezhnykh sistem, operatorov uslug platezhnoy infrastruktury, operatorov po perevodu denezhnykh sredstv"* [The Bank of Russia Resolution N2831-U of June 9, 2012, "O accountability to ensure the protection of information in the implementation of remittance payment system operators, service operators of the payment infrastructure operators for the transfer of funds"]. [Online] Available from: http://www.consultant.ru/document/cons_doc_LAW_131471/ (Accessed: 23rd August 2015).
11. The Bank of Russia. (2012) *Polozhenie Banka Rossii ot 31.05.2012 g. № 379-P: "Polozhenie o bespereboynosti funktsionirovaniya platezhnykh sistem i analize riskov v platezhnykh sistemakh"* [The Bank of Russia Resolution N379-P of May 31, 2012: "Regulaions on the smooth functioning of payment systems and risk analysis in payment systems"]. [Online] Available from: http://base.consultant.ru/cons/CGI/ online.cgi?req=doc;base=LAW;n=167370. (Accessed: 23rd August 2015).
12. Erokhin, S. (2011) *Metodika otsenki riskov narusheniya informatsionnoy bezopasnosti* [Methods of assessing the risks of violation of information security]. [Online] Available from: http://esstm.blogspot.com/2011/09/microsoft.html. (Accessed: 16th April 2013).
13. Banokin, P.I. & Vichugov, V.N. (2014) Implementation of Software system for prevention of internal data leaks. *Vestnik Tomskogo gosudarstvennogo universiteta. Upravlenie, vychislitel'naya tekhnika i informatika – Tomsk State University Journal of Control and Computer Science*. 1(26). pp. 84-88.
14. Atkina, V.S. & Olad'ko, A.Yu. (2015) Model of e-commerce security assessment. *Vestnik komp'yuternykh i in-formatsionnykh tekhnologiy – Herald of Computer and Information Technologies*. 3. pp. 33-37. (In Russian).

ВЕСТНИК ТОМСКОГО ГОСУДАРСТВЕННОГО УНИВЕРСИТЕТА
2015 Управление, вычислительная техника и информатика № 4 (33)

УДК 330.342.146
DOI: 10.17223/19988605/33/7

А.А. Сидоров, П.В. Сенченко, В.Ф. Тарасенко

МОНИТОРИНГ СОЦИАЛЬНО-ЭКОНОМИЧЕСКОГО РАЗВИТИЯ МУНИЦИПАЛЬНЫХ ОБРАЗОВАНИЙ В КОНТЕКСТЕ ЖИЗНЕННОГО ЦИКЛА ПЕРЕРАБОТКИ ИНФОРМАЦИИ

Работа выполнена при частичной финансовой поддержке РФФИ (проект РФФИ 12-06-33012-МОЛ_А_ВЕД_2012 «Методология комплексной оценки социально-экономического развития территориальных образований и эффективности государственного управления»).

В статье предложены различные уровни моделей мониторинга социально-экономического развития муниципальных образований на основе формального представления жизненного цикла обработки информации. Рассмотрен инструмент определения требований, обеспечивающих эффективное функционирование системы информационно-аналитического сопровождения управления в органах власти. Предложена цветная (раскрашенная) временная сеть Петри, описывающая динамическую модель мониторинга социально-экономического развития муниципальных образований. Указаны основные области применения разработанных моделей.

Ключевые слова: мониторинг социально-экономического развития; структурно-функциональная модель; динамическая модель; сети Петри.

Развитие современного общества характеризуется повышенными требованиями к использованию актуальной информации в управлении социально-экономическими процессами. Важным фактором в таком управлении является проведение глубокого анализа деятельности органов государственной власти и местного самоуправления, без которого сложно представить не только построение полноценной модели развития территории, но и проведение оценки деятельности властных структур. В качестве инструмента, с помощью которого возможно проводить информационно-аналитическую подготовку и принятие решений по широкому кругу вопросов в области публичного управления, целесообразно рассматривать систему мониторинга.

Изначально мониторинг появился и стал развиваться в естественных науках (например, гидрологический, метеорологический, радиологический и т.п.). Впоследствии он стал применяться в технических (промышленный мониторинг, мониторинг качества продукции и т.п.) и социальных (педагогика, психология и т.п.) отраслях знаний. Необходимо отметить, что здесь, как правило, существует четкая методологическая основа мониторинга, разработаны инструменты измерения, сформирована организационная структура реализации, а также закреплен на законодательном уровне его статус. Мониторинг в сфере публичного управления можно представить как существующую на постоянной основе систему наблюдения, сбора, анализа и прогнозирования социально-экономических показателей развития территории. Его цель – предоставление органам власти актуальной информации о ситуации в различных сферах жизнедеятельности.

Управленческая сущность мониторинга раскрывается через его функцию обслуживания процессов подготовки и принятия решений, что связано с тем, что только при наличии необходимой информации, отвечающей требованиям достоверности, полноты и своевременности, эффективно можно осуществлять планирование, организацию исполнения, контроль и регулирование административных и бизнес-процессов. Несмотря на широкое распространение терминов «социально-экономический мониторинг» и «мониторинг социально-экономического развития» как на нормативном уровне, так и в научной литературе, раскрытию их сущности посвящено немного работ. Практически предопределено, что указанные термины представляются общепринятыми и однозначно детерминированными. Однако при детальном рассмотрении оказывается, что априорная ясность отсутствует. Таким образом, выявленное

противоречие является основанием для формирования модели системы мониторинга социально-экономического развития территорий, рассмотренной на примере местных сообществ.

1. Функциональная модель мониторинга социально-экономического развития муниципальных образований

Неоднозначная природа по отношению к разным сферам, равно как и сложившееся многообразие в понимании мониторинга, предопределяет на первоначальном этапе разработки проблемы исследования необходимость формирования универсальной (инвариантной к части приложения) трактовки рассматриваемой категории: мониторинг – специальным образом сформированный инструмент информационного обеспечения управленческой деятельности для контроля, оценки, анализа и прогнозирования развития объекта управления на основе непрерывного процесса, состоящего из процедур жизненного цикла переработки информации (сбор, обработка, хранение, отображение и распространение), каждая из которых, в свою очередь, реализуется через свойственные ей методические приемы [1].

Мониторинг как инструмент управления реализуется только в определенной упорядоченности и взаимосвязи. Их внутреннее содержание можно раскрыть, используя формальную модель декомпозиции социально-экономических систем: согласно [2] представляется возможным сделать следующие сопоставления:

– процесс деятельности → мониторинговые процедуры;

– предмет деятельности → исходная информация, получаемая из источников;

– средства деятельности → методический инструментарий реализации мониторинговых процедур; организационный регламент; программно-технические средства;

– субъекты деятельности → носители мониторинговых функций;

– конечный продукт → результаты мониторинга.

Мониторинговые процедуры вытекают из модели жизненного цикла обработки информации, что позволяет вести речь о функциональной системе мониторинга. В качестве графического представления декомпозиции рассматриваемого процесса используется методология моделирования IDEF0 (рис. 1), в идеологии которой диаграммы являются главным компонентом модели и состоят из блоков – отображений функций системы и дуг – отображений потоков информации или объектов. Место соединения дуг и блоков определяет тип дуги: управляющие данные входят в блок сверху; объекты или информация, подвергаемые воздействию – входы, отображаются с левой стороны блока; механизмы – исполнитель или необходимый для воздействия инструментарий – представляются дугой, входящей в блок снизу, выходы – результат воздействия на входящие потоки – находятся с правой стороны блока. Каждый из блоков, изображенных на рис. 1, в свою очередь, также можно декомпозировать.

Элементы представленной модели могут быть охарактеризованы следующим образом:

1. Преобразуемое в «результаты мониторинга» (конечный продукт) в виде расчетных величин «информационное поле» (предмет деятельности) образует массив данных, обеспечивающих представление об объекте изучения и изменений в нем. Детерминированный перечень показателей формируется с учетом целей управления, критериев их эффективности и используемых методик контроля, оценки, анализа и прогноза мониторируемого объекта (в частности, социально-экономического развития).

2. Организационный регламент, упорядочивающий ход процесса, призван распределить нормативно полномочия (права, обязанности и ответственность) между участниками мониторинговых процедур, а также закрепить механизмы взаимодействия между ними.

3. Методический инструментарий, представляющий правила преобразования массива данных, содержит описание конкретных способов сбора, обработки, хранения и отображения информации.

4. Программно-технические средства осуществляют реализацию поддерживающей функции и позволяют повысить эффективность функционирования всей системы. Более того, широкое использование современных информационных технологий делает мониторинг, в том числе и социально-экономический, качественно иным, выводя его на иную ступень в развитии, что обусловлено расширением сервисных потенциалов реализации каждого этапа. Так, существенно могут быть увеличены объ-

емы хранимой и перерабатываемой информации, облегчены расчеты показателей, а также существенно расширены презентационные возможности.

5. Носителями мониторинговых функций выступают различные организации, структурные подразделения, а также отдельные сотрудники.

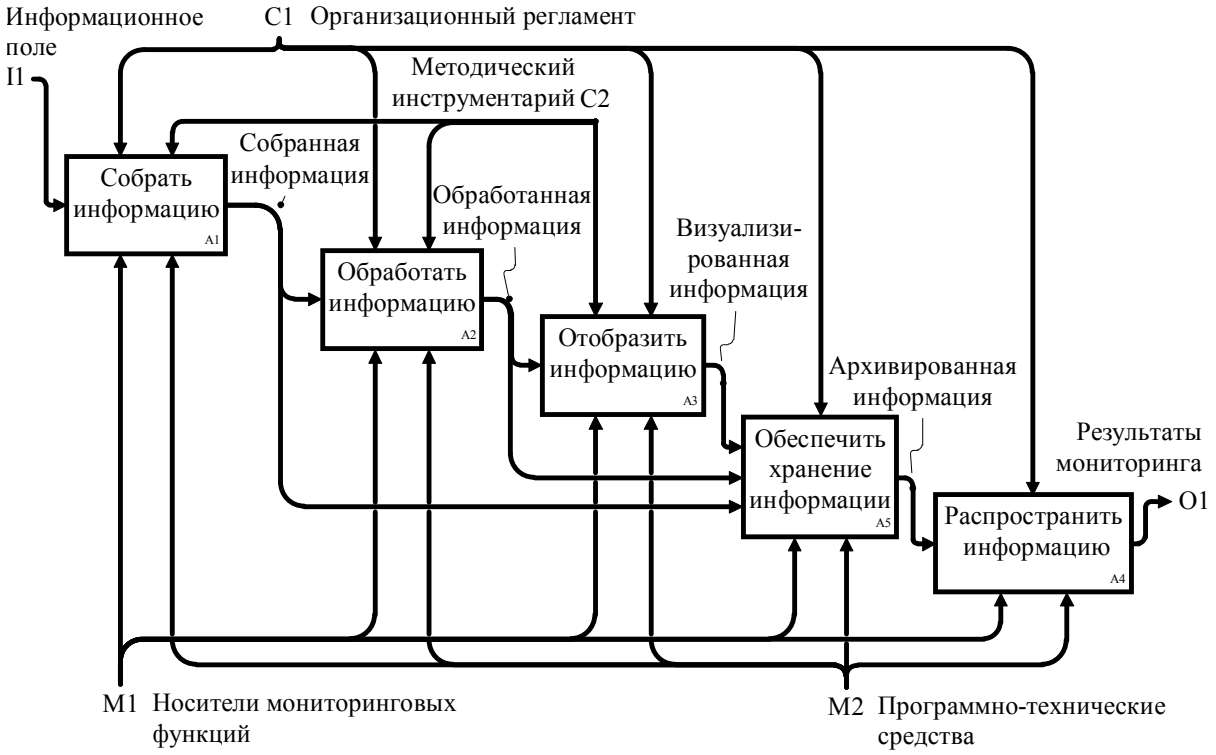


Рис. 1. Уровень А-0 «Провести мониторинг социально-экономического развития муниципальных образований»

Несмотря на общую прозрачность такого представления, вопрос внутреннего содержания обозначенных компонентов не является столь очевидным. Данный тезис является основанием для формирования перечня нормативных требований к организации системы мониторинга. Пусть $S = \{s_i\}$, $Z = \{z_j\}$ – множества мониторинговых процедур и обеспечивающих частей системы мониторинга соответственно. Поэлементно соотнося S и Z, получается произведение рассматриваемых множеств системы мониторинга. В то же самое время не для всех комбинаций их целесообразно формулировать. Например, при организации метеорологического мониторинга крайне важными являются требования к программно-техническим средствам на этапе сбора исходных данных, в то время как при организации системы мониторинга социально-экономического развития территорий эти ограничения не являются столь важными. Соответственно, представляется возможным сформулировать следующие примеры синтеза требований [3]:

– s_3z_1 – требования к возможностям программного обеспечения по визуализации результатов мониторинга;

– s_2z_2 – требования к адекватности применяемых методик оценки, анализа и прогноза социально-экономического развития;

– s_1z_3 – инструкции по технологии взаимодействия носителей мониторинговых функций (заказчиков, с одной стороны, регистраторов, интервьюеров и т.п. – с другой, носителей информации – с третьей) на этапе сбора информации;

– s_2z_4 – квалификационные и профессиональные требования к сотрудникам, осуществляющим оценку, анализ и прогноз социально-экономического развития.

2. Динамическая модель мониторинга социально-экономического развития муниципальных образований

Предложенная в разделе 1 структурно-функциональная модель описывает этапы сбора, обработки и хранения информации, не учитывая при этом динамику процесса. Использование математического аппарата сетей Петри позволяет перейти к динамическому моделированию мониторинга социально-экономического развития муниципальных образований и оптимизации деятельности соответствующих служб, что обосновывается следующими факторами [4]:

– сети Петри позволяют реализовать условия выбора, основанные на логических операциях AND и OR;

– с помощью сетей Петри возможно моделировать не только структурную часть процесса, но и отображать динамику выполнения основных его функций посредством перемещения фишек из одной позиции сети в другие позиции;

– структура процесса представляется в наглядном графическом виде с помощью графов.

Классическая структура сетей Петри определяется позициями, переходами, входной и выходной функцией и формально описывается следующим образом [5]:

$$C = (P, T, I, O), \quad (1)$$

где $P = \{p_1, p_2, \ldots, p_n\}$ – конечное множество позиций, $n \geq 0$; $T = \{t_1, t_2, \ldots, t_m\}$ – конечное множество переходов, $m \geq 0$; I: $T \rightarrow P^\infty$ – входная функция, определяющая отображение из переходов в комплекты позиций; O: $T \rightarrow P^\infty$– выходная функция, определяющая отображение из переходов в комплекты позиций.

Множество позиций и множество переходов не пересекаются: $P \cap T = \varnothing$, мощность множества P есть число n, а мощность множества T есть число m. Произвольный элемент P обозначается символом p_i, где $i = 1,\ldots, n$, а произвольный элемент T – символом t_j, где $j = 1,\ldots, m$.

Поскольку процесс мониторинга предполагает наличие различных временных характеристик, наиболее важными из которых являются время на сбор, обработка, предоставление (распространение) информации, в данном случае целесообразно применение аппарата цветных (раскрашенных) временных сетей Петри (Coloured Timed Petri net – CTPN). Использование такого расширения в сетях Петри позволяет более адекватно описывать реальные процессы по сравнению с обычными сетями Петри. Применение цветных (раскрашенных) временных сетей Петри позволяет исследовать технологию мониторинга до момента его практической реализации.

При рассмотрении цветной (раскрашенной) сети Петри необходимо учитывать многомерность функций I и O, т.е. $I = (I^1, I^2, \ldots, I^L)$, $O = (O^1, O^2, \ldots, O^L)$, где $L = |D|$, $D = \{d_1, d_2, \ldots, t_L\}$ – множество цветов (пометок). На множестве P задается функция $\mu(P)$ как совокупность целых неотрицательных чисел, определяющих количество цветных фишек (маркеров) в позициях. Следовательно, CTPN, в отличие от (1), определяется следующей шестеркой:

$$C = (P, T, I, O, \mu, \tau),$$

где τ – один из вариантов задания временных параметров элементов множества T; μ – разметка (маркировка) сети.

При этом временные параметры задаются следующим образом:

– продолжительность выполнения каждого задания мониторинга и время начала выполнения первого задания;

– продолжительность выполнения каждого задания и время окончания выполнения последнего задания;

– продолжительность выполнения каждого задания и время начала и окончания выполнения каждого из них.

Временные параметры здесь будем задавать длительностью исполнения задания, временем начала и завершения его исполнения. В общем случае моделирование процесса мониторинга предполагает изменение следующих исходных условий (в терминах сети Петри) [3, 6]:

– периодическое изменение состояния CTPN путем изменения маркировки фишек;

– изменение пользователем временных характеристик либо ресурсных параметров.

Для изменения состояния сети путем маркировки фишек необходимо использовать алгоритм маркировки (разметки) сетей Петри, при этом начальная маркировка μ_0 соответствует начальному состоянию сети. Переход сети считается разрешенным (произойдет некоторое событие), а соответствующее задание – активным, если все его условия соблюдены (имеются все необходимые фишки).

Условия срабатывания перехода могут быть заданы следующим образом: в выходных позициях перехода фишки появляются сразу после того, как будет отмечено выполнение любой из работ, входящих в соответствующий данному переходу этап деятельности, вместе с тем отметка об исполнении остальных работ данного этапа осуществляется в течение некоторого времени, отведенного для данного этапа. Это необходимо для избегания тупиковых ситуаций по времени срабатывания для следующих параллельных переходов. В случае истечения регламентного времени выполнения этапа и наличия в нем невыполненных работ (задание не исполнено в срок) пользователь сети имеет возможность самостоятельно принять решение о дальнейших действиях. Окончание этапа порождает новую маркировку μ позиций и определяет условия активизации следующих переходов.

Графическое представление сетей Петри является важным аспектом для обеспечения наглядности восприятия сети Петри и проведения качественного анализа моделируемых процессов. В общем понимании структура простой сети Петри состоит из двух типов узлов:

– символ **O** (круг) представляет позицию;

– символ │(планка) является переходом и определяет некоторое примитивное событие, описываемое в сети.

Соединение позиций и переходов обеспечивается путем представления ориентированных дуг. Дуга, направленная от позиции p_i к переходу t_j, определяет позицию, которая является входом перехода. Дуга, направленная от перехода t_j к позиции p_i, определяет позицию, которая является выходом из перехода. При этом условия, т.е. логические состояния системы, принимающие значения ИСТИНА или ЛОЖЬ, моделируются позициями, а события или действия, происходящие в системе, моделируются переходами.

Выполнение некоторого условия (значение ИСТИНА) представляется фишкой в позиции, которая соответствует данному условию, при этом запуск перехода удаляет разрешающие фишки, определяющие выполнение так определенных предусловий, и образуют новые фишки, которые определяют выполнение постусловий. Следует отметить, что в сетях Петри непримитивные события представляются уже не в виде планок, а в виде прямоугольников, что позволяет упростить некоторые виды сетей Петри.

Если необходимо обеспечить переход от функциональной к динамической модели, то при переходе от диаграммы, представленной в нотации IDEF0, к сети Петри функциональный блок диаграммы заменяется фрагментом сети Петри, имитирующим работу этого блока. При этом позициями сети становятся возможные потоки информации или объекты системы, подвергаемые воздействию, управляющие данные, механизмы (исполнитель или необходимый инструментарий), а также результат воздействия на входящие потоки.

Основываясь на предложенных выше определениях и порядке трансформации функциональной модели в сеть Петри, можно предложить цветную (раскрашенную) временную сеть Петри, описывающую динамическую модель ведения мониторинга (рис. 2).

Предложенные в функциональной модели (рис. 1) взаимосвязи элементов рассматриваемой сети Петри можно представить следующим образом:

– t_1 – собрать информацию: $I(t_1) = \{p_1, p_2, p_3\}$ – входная функция, где p_1 – информационное поле; p_2 – носители мониторинговых функций; p_3 – программно-технические средства; $O(t_1)= \{p_4, p_2, p_3\}$ – выходная функция, где p_4 – собранная информация.

– t_2 – обработать информацию: $I(t_2) = \{p_4, p_2, p_3\}$, $O(t_1)= \{p_5, p_2, p_3\}$, где p_5 – обработанная информация;

– t_3 – отобразить информацию: $I(t_3) = \{p_5, p_2, p_3\}$, $O(t_3) = \{p_6, p_2, p_3\}$, где p_6 – визуализированная информация;

– t_4 – обеспечить хранение информации – данный переход сработает в случае наличия фишек в одной из позиций: p_4, p_5 или p_6, а также при наличии фишек в позициях p_2 и p_3: $I(t_4) = \{p_4, p_5, p_6, p_2, p_3\}$, $O(t_4) = \{p_7, p_2, p_3\}$, где p_7 – архивированная информация;

– t_5 – распространить информацию: $I(t_3) = \{p_5, p_7, p_3\}$, $O(t_5) = \{p_8, p_2, p_3\}$, где p_8 – результат мониторинга.

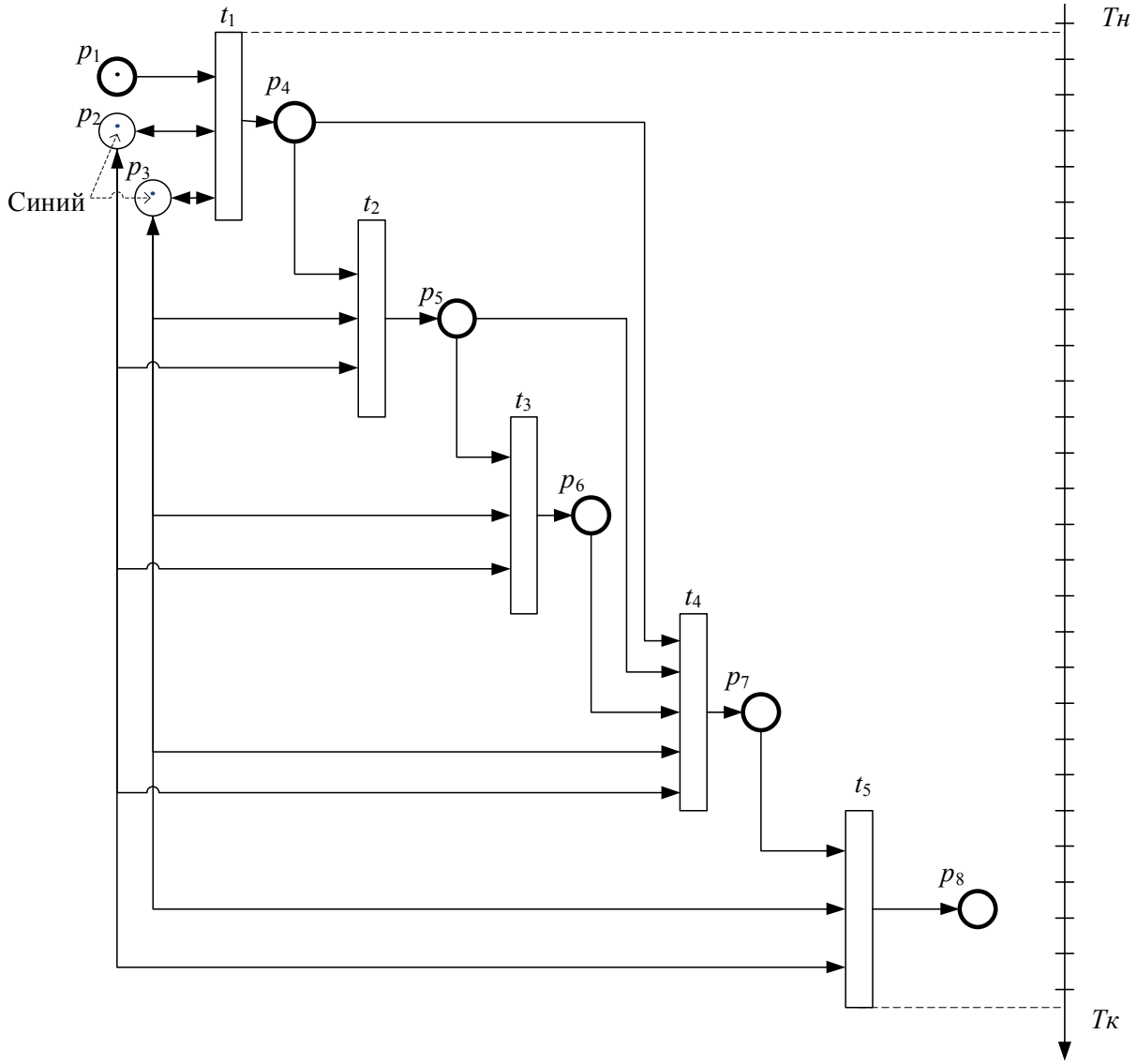


Рис. 2. Цветная (раскрашенная) временная сеть Петри, описывающая процесс мониторинга социально-экономического развития муниципальных образований

Преобразовав функциональную модель (рис. 1), получим цветную (раскрашенную) временную сеть со свободным выбором – позиция p_4 сети может иметь различную смысловую нагрузку (в зависимости от цвета метки), что позволяет использовать ее для запуска разных переходов:

– если p_4 описывает информацию, которую необходимо подвергнуть дополнительной обработке, то информация поступит на вход перехода t_2;

– если p_4 содержит информацию, для которой необходимо обеспечить сохранность (без предварительной обработки), то такая информация поступит на вход перехода t_4.

Различные характеристики могут быть также заданы для позиции p_5:

– если p_5 описывает обработанную информацию, которую необходимо визуализировать, то информация поступит на вход перехода t_3;

– если p_5 содержит информацию, для которой необходимо обеспечить сохранность (без визуализации), то такая информация поступит на вход перехода t_4.

Определенную смысловую нагрузку несет и временная составляющая сети τ_i. Например, время срабатывания перехода t_3 зависит от продолжительности выполнения мероприятий по сбору и обработ-

ке информации – переходов t_1и t_2. Соответственно, переход t_5 сработает только после выполнения задач на предыдущих переходах и появления фишек в позициях p_7, p_2 и p_3.

Аналогично можно провести динамическое моделирование всех выделенных в разделе 1 этапов мониторинга социально-экономического развития муниципального образования. Фактически такие цветные (раскрашенные) временные сети Петри будут являться декомпозициями соответствующих переходов базовой сети.

Поведение модели может быть проанализировано посредством компьютерного моделирования. Использование существующих в настоящее время программных средств, имитирующих работу цветной временной сети Петри (например, Design / CPN [7]), позволяет, манипулируя временными характеристиками сети и изменяя число и цвет фишек в позициях, моделировать различные ситуации, возникающие в процессе ведения мониторинга. Так, например, появляется возможность руководителю организации, отвечающему за ведение мониторинга, использующему данную динамическую модель, распределить человеческие и временные ресурсы для выполнения поставленной задачи в зависимости от количества собираемых данных, необходимого времени на их обработку и вида носителей мониторинговых функций.

Заключение

Предлагаемые модели являются базовыми при разработке информационных технологий мониторинга социально-экономического развития муниципальных образований и используются при построении общих моделей деятельности органов государственной власти и местного самоуправления.

Использование математического аппарата сетей Петри (вспомогательного инструментария для специалистов по мониторингу) облегчает мониторинг. Эмуляция представленной на рис. 2 сети Петри при разрешении конкретных проблемных ситуаций с включением в ее состав дополнительных позиций и переходов позволяет определять, например, состав, количество и загрузку исполнителей мониторинга. Использование результатов моделирования повышает степень обоснованности принятия решений руководителем при мониторинге.

ЛИТЕРАТУРА

1. *Сидоров А.А., Сенченко П.В.* Структурно-функциональная и динамическая модели мониторинга социально-экономического развития муниципальных образований // Доклады Томского государственного университета систем управления и радиоэлектроники. 2012. № 2 (26), ч. 1. С. 258–264.
2. *Перегудов Ф.И., Тарасенко Ф.П.* Основы системного анализа : учеб. 2-е изд., доп. Томск : Изд-во НТЛ, 1997. 396 с.
3. *Сидоров А.А., Сенченко П.В.* Мониторинг социально-экономического развития территорий в контексте информационного обеспечения системы управления по результатам // Ползуновский альманах. 2009. № 1. С. 293–299.
4. *Ехлаков Ю.П., Тарасенко В.Ф., Жуковский О.И., Сенченко П.В., Грищенко Ю.Б.* Динамические модели бизнес-процессов. Теория и практика реинжиниринга. Томск : Том. гос. ун-т систем управления и радиоэлектроники, 2014. 203 с.
5. *Питерсон Дж.* Теория сетей Петри и моделирование систем : пер. с англ. М. : Мир, 1984. 264 с.
6. *Тарасенко В.Ф.* Нелинейные математические модели и информационные системы в финансовом менеджменте / под ред. В.З. Ямпольского. Томск : Изд-во ТПУ, 1998. 191 с.
7. Design / CPN – Computer Tool for Coloured Petri Nets. URL: http://www.daimi.au.dk/designCPN/, свободный (дата обращения: 11.11.2015).

Сидоров Анатолий Анатольевич, канд. экон. наук, доцент. E-mail: saa@muma.tusur.ru
Сенченко Павел Васильевич, канд. техн. наук, доцент. E-mail: pvs@tusur.ru
Томский государственный университет систем управления и радиоэлектроники
Тарасенко Владимир Феликсович, д-р техн. наук, профессор. E-mail: vtara54@mail.ru
Томский государственный университет

Поступила в редакцию 16 июня 2015 г.

Sidorov Anatoly A., Senchenko Pavel V. (Tomsk State University of Control System and Radioelectronics, Russian Federation), *Tarasenko Vladimir F.* (Tomsk State University, Russian Federation).

Monitoring of the social and economic development of municipalities in the life-cycle of information processing.

Keywords: management; monitoring; socio-economic development; structural-functional model; dynamic model; Petri nets.

DOI: 10.17223/19988605/33/7

Monitoring in state and municipal government can be represented as a system of observation, data collection, analysis and forecasting of socio-economic indicators of development of the territory. The main objective of such monitoring is to provide the authorities of relevant information on the situation in the various spheres of life. Monitoring as a management tool is realized only in a certain order and relationship. Their internal contents can be opened using a formal model of decomposition of socio-economic systems in which there are the following elements: the process of activity, scope of activities, the final product, means the activities of stakeholders.

Monitoring procedures are derived from the model of the life cycle of information processing. In fact, here is a question of a functional monitoring system. As a graphic representation of the process of decomposition is used IDEF0-model.

Monitoring Indicators are a set of basic (object of activity) and calculated values (final product), providing an idea of the socio-economic system and the changes in it. A specific list is formed with the objectives of the control criteria of their effectiveness and the techniques used for monitoring, assessment, analysis and forecast of the monitored object (the socio-economic development).

The structural-functional model gives an idea of the organization of monitoring, data collection, processing and storage of information, not taking into account the dynamics of the process. Using the mathematical apparatus of Petri nets allows us to go to the dynamic modeling of monitoring of socio-economic development of the municipalities and the optimization of the compliance of corresponding services related to monitoring. The offered color timed Petri net describes dynamic model of monitoring of social and economic development of municipalities.

Use of currently available software that simulates colored timed Petri net (e.g., Design / CPN) allows, manipulating the time components of the network and changing the quantity and color of chips in positions, to simulate various situations that arise in the process of monitoring. For example, the head of the organization responsible for conducting the monitoring, use this dynamic model to allocate human and time resources for re-perform this task, depending on the amount of information collected, the time required to process it and the type of media monitoring functions.

The proposed article multidimensional model can be taken as a basis in the formulation of the problem and the development of information technology for monitoring of socio-economic development of municipal entities and may be used to construct more general models of public authorities and local government.

REFERENCES

1. Sidorov, A.A. & Senchenko, P.V. (2012) Structural-functional and dynamic models of monitoring of social and economic development in municipalities. *Doklady Tomskogo gosudarstvennogo universiteta sistem upravleniya i radio-elektroniki – Proceedings of Tomsk State University of Control Systems and Radioelectronics*. 2(26). Part 1. pp. 258-264. (In Russian).
2. Peregudov, F.I. & Tarasenko, F.P. (1997) *Osnovy sistemnogo analiza* [Basics of System Analysis]. 2nd ed. Tomsk: NTL.
3. Sidorov, A.A. & Senchenko, P.V. (2009) Monitoring sotsial'no-ekonomicheskogo razvitiya territoriy v kontekste informatsionnogo obespecheniya sistemy upravleniya po rezul'tatam [Monitoring the socio-economic development of territories in the context of an information security management system based on the results]. *Polzunovskiy al'manakh*. 1. pp. 293-299.
4. Ekhlakov, Yu.P., Tarasenko, V.F., Zhukovskiy, O.I., Senchenko, P.V. & Gritsenko, Yu.B. (2014) *Dinamicheskie modeli biznes-protsessov. Teoriya i praktika reinzhiniringa* [Dynamic models of business processes. Theory and practice of reengineering]. Tomsk: Tomsk State University of Control Systems and Radioelectronics.
5. Piterson, J. (1981) *Petri Net theory and the modeling of systems*. Englewood Cliffs, New Jersey: Prentice Hall, Inc.
6. Tarasenko, V.F. (1998) *Nelineynye matematicheskie modeli i informatsionnye sistemy v finansovom menedzhmente* [Non-linear mathematical models and information systems in financial management]. Tomsk: Tomsk Polytechnics Univesity.
7. *Design / CPN – Computer Tool for Coloured Petri Nets*. [Online] Available from: http://www.daimi.au.dk/designCPN/. (Accessed: 11th November 2015).

ВЕСТНИК ТОМСКОГО ГОСУДАРСТВЕННОГО УНИВЕРСИТЕТА
2015 Управление, вычислительная техника и информатика № 4 (33)

ПРОЕКТИРОВАНИЕ И ДИАГНОСТИКА ВЫЧИСЛИТЕЛЬНЫХ СИСТЕМ

УДК 510.63
DOI: 10.17223/19988605/33/8

В.В. Андреева, Т.П. Тарновская

СОКРАЩЕНИЕ РАНГА КОНЬЮНКЦИИ, ПРЕДСТАВЛЯЮЩЕЙ КОРЕНЬ ЛОГИЧЕСКОГО УРАВНЕНИЯ

Работа выполнена при финансовой поддержке Российского научного фонда (*проект № 14-19-00218*).

Предлагается метод поиска корня логического уравнения по возможности меньшего ранга. Поиск корня осуществляется с помощью *дерева разложения*. Сформулированы правила выбора переменной разложения, способствующие сокращению ранга корня логического уравнения. Сформулированы достаточные условия существования корня логического уравнения.
Ключевые слова: логическое уравнение; дерево разложения; троичные векторы.

Логические уравнения используются в различных приложениях [1]. Они широко применяются в электротехнике, комбинаторике, логике высказываний, целочисленном программировании и др. Принято представлять корни таких уравнений конъюнкциями. Требования к качеству корня в общем случае зависят от контекста задачи. Например, построение тестовых наборов для неисправности сводится к решению соответствующего логического уравнения с представлением решения в виде конъюнкции. Проверяющий тест для класса неисправностей задается совокупностью таких конъюнкций. Чем меньше ранги конъюнкций, тем больше возможностей для сокращения длины проверяющего теста [2].

В данной работе предлагается модификация алгоритма Закревского [3] поиска одного корня логического уравнения. Целью модификации является построение конъюнкции по возможности меньшего ранга. В общем случае поиск корня сводится к сокращенному обходу дерева. С целью сокращенного обхода на каждом шаге выбирается переменная разложения, способствующая ускорению поиска корня уравнения. В работе [3] предлагаются правила выбора переменной разложения, причем предпочтение отдается однородным переменным, которые способствуют быстрому поиску, но не сокращению ранга. В предлагаемой модификации выбирается такая переменная в очередной вершине дерева разложения, которая способствует сокращению ранга конъюнкции. Выбор переменной осуществляется среди конъюнкций минимального ранга, сопоставляемых вершине, причем выбирается не обязательно однородная переменная.

1. Постановка задачи

Рассмотрим ДНФ $D = K_1 \vee, \ldots, \vee K_s$, зависящую от множества переменных $X = \{x_1, \ldots, x_n\}$. Представим D в виде троичной матрицы M (множеством троичных векторов), сопоставив каждой конъюнкции троичный вектор. Компоненты троичного вектора принимают значения из множества {0, 1, x}. Здесь символ x означает неопределенную компоненту, которая может принимать значение либо 0, либо 1. Значения {0, 1} переменной троичного вектора будем называть *определенными*, значение x будем называть *неопределенным*. Троичный вектор, в котором определенные компоненты принимают только значение 1, будем называть *положительно однородным вектором* и *отрицательно однородным*, если определённые компоненты принимают только значение 0. Векторы, содержащие среди определенных компонент как 0, так и 1, будем называть *неоднородными*.

Однородным столбцом, соответствующим переменной x_i, в матрице троичных векторов M будем называть столбец, состоящий либо из значений {0, x}, либо {1, x}. Матрицу M будем называть совместимой, если в ней все столбцы однородны. *Максимально определенным* столбцом будем называть столбец, состоящий из максимального числа определенных компонент. В качестве примера рассмотрим ДНФ $D = x_1 \vee x_4 \vee x_2x_3x_4x_5 \vee x_1x_4x_5x_6 \vee \bar{x}_1x_3x_4x_5x_6$, матрица M в этом случае будет иметь следующий вид. Здесь столбцы, соответствующие переменным x_1, x_2, x_3, являются *однородными*, а столбец, соответствующий переменной x_4, является *максимально определенным*. Очевидно, что некоторая конъюнкция r_i представляет корень уравнения $D = 0$, если она ортогональна каждой конъюнкции, представляемой строкой матрицы M. Ранг конъюнкции обозначим $rang_i$.

При анализе матрицы M будем придерживаться следующих правил.

Правило 1. Если в троичной матрице M присутствует троичный вектор, в котором определена всего одна компонента, то соответствующая ей переменная в конъюнкции r_i содержится с противоположным знаком инверсии. В этой ситуации из матрицы M могут быть удалены все троичные векторы, ортогональные троичному вектору, представляющему конъюнкцию r_i. У оставшихся троичных векторов соответствующая компонента может быть заменена значением x. Например, в рассматриваемом примере (табл. 1) два троичных вектора веса 1, тогда в конъюнкции r_i могут присутствовать две переменные с противоположным знаком инверсии, т.е. $\bar{x}_2, x_4$.

Таблица 1

$M =$

	x_1	x_2	x_3	x_4	x_5	x_6
K_1	x	1	x	x	x	x
K_2	x	x	x	0	x	x
K_3	x	1	0	1	1	x
K_4	0	x	x	1	0	0
K_5	0	x	0	1	0	1

После упрощения матрица M будет иметь вид, представленный в табл. 2.

Таблица 2

$M =$

	x_1	x_2	x_3	x_4	x_5	x_6
K_4	0	x	X	x	0	0
K_5	0	x	0	x	0	1

Если в результате упрощения в оставшейся матрице будет получен вектор, состоящий только из значений x, то рассматриваемое уравнение *не имеет решения*.

Правило 2. Если в матрице присутствует однородный полностью определенный монотонный столбец, то соответствующая переменная включается в решение конъюнкцию r_i с противоположным знаком инверсии. Троичный вектор, представляющий конъюнкцию, будет ортогонален всем векторам матрицы M, следовательно, является корнем.

Например, в матрице M (табл. 2), полученной после упрощения, переменной x_1 соответствует однородный столбец, следовательно, эту переменную можно включить в r_i, тогда $r_i = x_1\bar{x}_2x_4$.

Эти правила будут использоваться при отыскании корня уравнения $D = 0$, где D – произвольная ДНФ.

2. Поиск корня логического уравнения $D = 0$

Рассмотрим общую процедуру построения дерева разложения относительно некоторой переменной x_i троичной матрицы M [4]. Из вершины дерева, отмеченной переменной x_i, исходят три дуги, помеченные значениями {0,1,x}. Каждой дуге сопоставляется соответствующее подмножество троичных векторов, полученное путем разбиения исходного множества M относительно переменной x_i на три

подмножества, обозначим их следующим образом: $M^1(x_i = 1)$, $M^0(x_i = 0)$, $M^x(x_i = \mathrm{x})$ по значениям {0,1,x}(рис. 1).

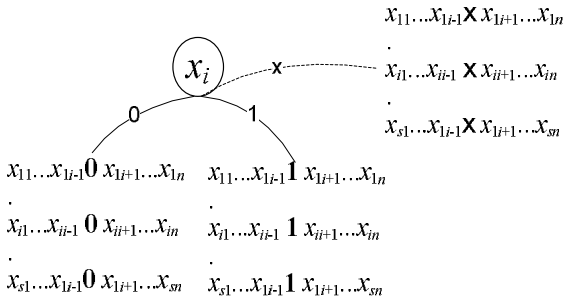


Рис. 1. Дерево разложения по переменной x_i

Среди множеств $M^1(x_i = 1)$, $M^0(x_i = 0)$ выбирается множество большей мощности, и переменная x_i включается в конъюнкцию r_i со знаком инверсии, обеспечивающим ортогональность всем элементам выбранного множества. Во втором из рассматриваемых множеств переменной x_i присваивается значение x, и полученное в результате множество троичных векторов объединяется с $M^x(x_i = \mathrm{x})$. Новое множество M сопоставляется со следующим шагом разложения. При этом дерево корректируется, так что из вершины, помеченной x_i, исходят две дуги, одна из которых помечена определенным значением, а другая значением x. Если на очередном шаге разложения выясняется, что конъюнкция r_i не может быть достроена до корня уравнения, т.е. в множестве, сопоставляемом дуге, отмеченной переменной x, присутствует вектор из одних неопределенных компонент, то возвращаемся в ближайшую вершину дерева разложения, для которой оба множества $M^1(x_i = 1)$, $M^0(x_i = 0)$ не пусты.

2.1. Достаточные условия существования корня логического уравнения

Ведем следующие обозначения:

M^+ – подмножество положительно однородных векторов;

M^- – подмножество отрицательно однородных векторов;

M^* – подмножество неоднородных векторов.

Утверждение 1. Если множество $M = M^*$, то корень уравнения существует.

Доказательство. Пусть множество M состоит только из неоднородных троичных векторов. Значения определенных компонент в неоднородном троичном векторе принимают 0 и 1, следовательно, найдется такой троичный вектор, представляющий корень уравнения, который будет ортогонален всем векторам из M либо только по компонентам 1, либо только по компонентам 0. Утверждение доказано.

Рассмотрим множество M (табл. 3), вектор 1x11xx, ортогональный всем троичным векторам по определенному значению 0, а троичный вектор x0x0xx ортогонален всем троичным векторам по значению 1.

Таблица 3

$M =$

	x_1	x_2	x_3	x_4	x_5	x_6
K_1	**0**	1	x	x	**0**	x
K_2	1	1	x	**0**	x	x
K_3	x	1	**0**	1	1	x
K_4	**0**	x	x	1	**0**	**0**

Утверждение 2. Если множество $M = M^* \cup M^-$ либо $M = M^* \cup M^+$, то корень уравнения существует.

Доказательство. Рассмотрим случай, когда $M = M^* \cup M^-$, т.е. наряду с неоднородными векторами присутствуют однородно отрицательные векторы, следовательно, в каждом векторе присутствуют

определенные компоненты со значением 0, тогда найдется такой троичный вектор, представляющий корень уравнения, который будет ортогонален всем векторам из M по компонентам со значением 0. В случае $M = M^* \cup M^+$ найдется такой троичный вектор, представляющий корень уравнения, который будет ортогонален всем векторам из M по компонентам со значением 1. Утверждение доказано.

Например, во множестве M (табл. 4) вектор 1x1xx ортогонален всем троичным векторам по определенным компонентам со значением 0.

Таблица 4

$M =$

	x_1	x_2	x_3	x_4	x_5	x_6
K_1	**0**	x	x	x	**0**	x
K_2	**0**	**0**	x	**0**	x	x
K_3	x	1	**0**	1	1	x
K_4	**0**	x	x	1	**0**	**0**

В множестве троичных векторов M выполним упрощения по правилам 1, 2.

Пусть $M^{**} = M^+ \cup M^-$.

Утверждение 3. Если множество M представляется в виде $M = M^* \cup M^+ \cup M^-$ и M^{**} совместимо, то корень уравнения существует.

Доказательство. Пусть подмножество M^{**} совместимо, тогда найдем пересечение между этими троичными векторами. Результатом пересечения является *неоднородный* вектор V. Получим вектор V^* путем инвертирования всех значений вектора V, таким образом, V^* ортогонален всем векторам из M^{**}. Так как M^* состоит из неоднородных векторов, то найдется вектор V^{**}, совместимый с вектором V^*, ортогональный всем элементам из M^*. Утверждение доказано**.**

Например, в множестве M (табл. 5) выделим множество M^{**}, которое состоит из двух троичных векторов, 11xxxx и xxx000. Результатом их пересечения будет вектор V = 11x000, получим вектор V^* = 00x111. По утверждению 3 найдется вектор, совместимый с V^* и ортогональный всем векторам из M^*. В нашем случае V^{**} = x0xx1x.

Таблица 5

$M=$

	x_1	x_2	x_3	x_4	x_5	x_6
K_1	1	**1**	x	x	X	x
K_2	x	x	x	0	**0**	0
K_3	x	**1**	0	1	1	x
K_4	0	x	x	1	**0**	0
K_5	0	x	0	1	**0**	1

Например, полученный вектор V^{**} поглощает вектор V^*, тогда результирующий вектор $R = V^{**}$, он содержит меньшее количество компонент и ортогонален всем элементам матрицы M (табл. 5).

2.2. Выбор переменой разложения

Предлагаются следующие правила выбора переменной разложения.

Рассмотрим множество векторов, представляемое матрицей M. Исключим из матрицы все поглощаемые векторы, далее упорядочим векторы по возрастанию их рангов. Выделим группу векторов, содержащую троичные векторы *минимального ранга*.

В этой группе выделяем множество монотонных векторов $M^+(M^-)$ и среди последних векторов выполняем поиск максимально определенных столбцов, среди выбранных столбцов предпочтение отдается тому, который имеет максимальное число определенных компонент в множестве M. Сопоставляемая этому столбцу переменная выбирается в качестве переменной разложения.

В случае отсутствия однородных векторов в группе предпочтение отдается столбцу, который имеет максимальное число определенных компонент в множестве M. Если существует несколько неоднородных столбцов с максимальным числом определенных компонент, то для каждого столбца подсчи-

тываем число единичных и нулевых компонент. Предпочтение отдаем столбцу с максимальным числом либо значений 0, либо значений 1. Иначе выбираем любой столбец с максимальным числом определенных компонент. Сопоставляемая ему переменная выбирается в качестве переменной разложения.

Рассмотрим пример (табл. 6). Здесь K_1, K_2, K_3, K_4 образуют подмножество троичных векторов минимального ранга, в котором присутствуют однородно отрицательные троичные векторы K_1, K_2.

Таблица 6

$M=$

	x_1	x_2	x_3	x_4	x_5	x_6
K_1	**x**	**x**	**x**	**x**	**0**	**0**
K_2	**x**	**x**	**x**	**0**	**x**	**0**
K_3	**1**	**x**	**0**	**x**	**x**	**x**
K_4	**x**	**x**	**x**	**1**	**0**	**x**
K_5	1	x	1	0	x	x
K_6	1	1	x	0	x	x
K_7	x	1	0	x	0	x
K_8	1	1	x	x	0	x
K_9	x	1	1	x	x	0

По утверждению 2 корень уравнения существует. В качестве переменной разложения выбираем максимально определенную компоненту из множества однородно отрицательных векторов, а именно компоненту, отмеченную переменной x_6. Из вершины, помеченной x_6, исходят две дуги: одна дуга, помеченная значением 0, вторая дуга, помеченная значением x. Для множества, сопоставляемого дуге со значением x, разложение продолжается, как показано на (рис. 2). Корень уравнения найден, если на очередном шаге разложения множество, сопоставляемое дуге со значением x, оказывается пустым. Построенная конъюнкция представляет корень уравнения. Для рассматриваемого примера $r = \bar{x}_1 x_5 x_6$.

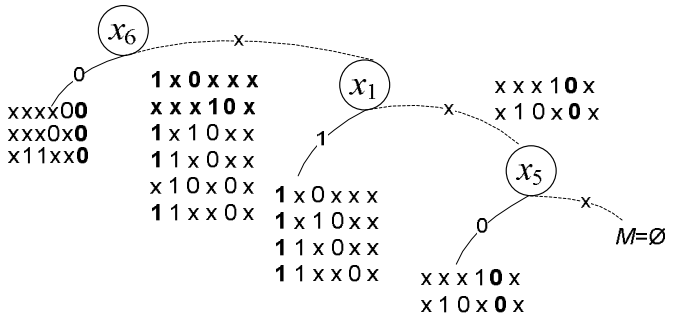


Рис. 2. Дерево разложения по переменной x_i

Рассмотрим множество M, представленное в табл. 7 и 8, для которого корня не существует. Дерево разложения представлено на рис 3.

Таблица 7

	x_1	x_2	x_3	x_4	x_5	x_6
K_1	1	x	x	x	x	0
K_2	1	x	1	x	1	x
K_3	1	x	x	x	0	1
K_4	1	0	x	x	1	1
K_5	1	1	0	x	x	1
K_6	0	0	x	x	1	x
K_7	0	1	x	x	x	1

Таблица 8

	x_1	x_2	x_3	x_4	x_5	x_6
K_8	0	x	x	x	0	1
K_9	0	x	x	0	1	0
K_{10}	0	1	x	1	0	0
K_{11}	0	1	x	1	1	0
K_{12}	0	x	1	x	0	0
K_{13}	0	0	0	x	0	0

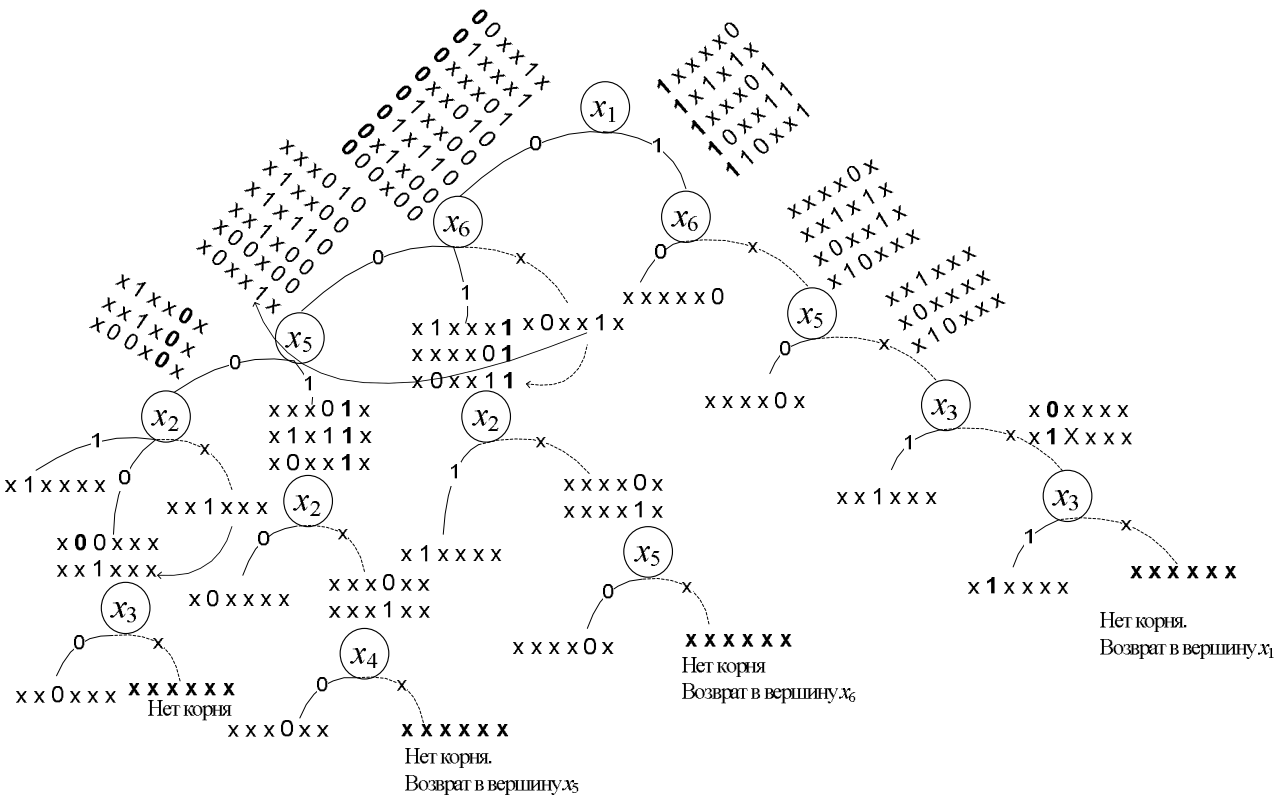


Рис. 3. Дерево разложения множества *M* в случае отсутствия корня

3. Экспериментальные результаты

Предложенная модификация программно реализована и испытана на случайно сгенерированных примерах, полученных с помощью программы [5]. Генерировались примеры разной размерности и разной разреженности по неопределённым компонентам. Результаты сравнивались с алгоритмом Закревского. Как показали экспериментальные результаты (см. табл. 9, 10), предложенная модификация сокращает ранг конъюнкции.

Таблица 9

№	*i*	*t*	*%dc*	*rZ*	*rDR*
1	100	500	15	39	37
2	100	500	25	22	19
3	100	1000	50	37	34
4	200	1000	60	40	37
5	200	1300	70	57	51
6	300	1500	75	35	30
7	300	1500	85	92	63

Таблица 10

№	*i*	*t*	*%dc*	*rZ*	*rDR*
8	350	2000	25	150	148
9	350	3000	30	120	116
10	350	3000	35	220	213
11	400	3000	45	223	219
12	400	3000	65	131	126
13	500	3000	70	157	150
14	500	3000	85	194	186

Здесь ***i*** – длина вектора, ***t*** – число векторов матрицы *M*, ***%dc*** – процент неопределенных компонент, *rZ* – ранг корня, полученного с помощью алгоритма Закревского, ***rDR*** – ранг корня, полученного с помощью предложенного алгоритма.

Заключение

В работе предложена модификация алгоритма Закревского поиска одного корня логического уравнения. Целью модификации является построение корня по возможности меньшего ранга. Корень

уравнения строится с помощью дерева разложения. Предложены эвристики выбора переменной разложения, ориентированные на сокращение ранга конъюнкции, представляющей корень уравнения. Эффективность предложенной модификации подтверждается экспериментально. Сформулированы достаточные условия существования корня уравнения $D = 0$.

ЛИТЕРАТУРА

1. *Crama Y.*, *Hammer P.L.* Boolean functions: theory, algorithms, and applications. Cambridge ; New York : Cambridge University Press, 2011. P. 440.
2. *Andreeva V.* Test minimization technique for multiple stuck-at faults of combinational circuits // Proc. 8th East-West Design & Test International Symposium. Saint Petersburg, Russia. 2010. P. 168–170.
3. *Закревский А.Д.* Логические уравнения. М. : Едиториал УРСС, 2003. 96 с.
4. *Sorudeykin K.*, *Andreeva V.* Decomposition Tree – based Compaction Procedure with Iteration Steps for Interconversional Layouts of Tasks // Proc. 12th IEEE East-West Design & Test Symposium, 2014. Kiev, Ukraine, 2014. P. 173–178.
5. *Měchura T.* Random Circuits Generators // Czech Technical University in Prague. 2008. URL: http://ddd.fit.cvut.cz/prj/Circ_Gen (дата обращения: 15.11.2015).

***Андреева Валентина Валерьевна*,** канд. техн. наук. E-mail: valenrina.andreeva@mail.tsu.ru
Тарновская Татьяна Павловна. E-mail: tarnovskayat@mail.ru
Томский государственный университет

Поступила в редакцию 7 октября 2015 г.

Andreeva Valentina V., Tarnovskaya Tatyana P. (Tomsk State University, Russian Federation).

Reducing the rank of the conjunction, representing the root of the logical equation.

Keywords: boolean equation; decomposition tree; ternary vectors.

DOI: 10.17223/19988605/33/8

In this paper the modification of Zakrevskij algorithm of finding a root of logical equation has been suggested. A root of logic equation $D = 0$ is representing by product r_i. The purpose of this modification is to find a product that has possibly smaller rank. As a rule finding the root of logic equation is reduced to tree traversal. Here, we deal with $D = 0$, where D is SoP (Sum of Products), that is represented as a set of ternary vectors. The ternary vector is a partially specified bit vector, i.e. elements (variables x_i) of it can contain values {1,0, x}. The values '1' and '0' are *determined values*. The value x is *don't care symbol*. Note this set as *M*.

The root of the equation is formed with the decomposition tree, by variables x_i. In general, from the vertex of the tree marked by variable x_i, run three branches labeled with the values {0,1,x}. Each branch is associated with a corresponding subset of ternary vectors obtained by dividing a set *M* with respect to variable x_i into three subsets $M^1(x_i = 1)$, $M^0(x_i = 0)$, $M^x(x_i = \mathrm{x})$. The variable x_i from the set with maximum power is included into to product r_i with the opposite sign of inversion. The decomposition tree is changed so that from the vertex labeled x_i, run two branches, one of which is labeled with a determined value, but another is labeled with value x. If the next step of decomposition fails that is the product r_i cannot be derived as the root of the equation, then we return to the closest vertex of decomposition tree for which both sets $M^1(x_i = 1)$, $M^0(x_i = 0)$ is not empty. The root of the equation is found when the set corresponding to branch labeled x is empty.

In this paper variable selection heuristics have been proposed. The experimental results confirm advantages of this approach. The sufficient conditions for the existence of the root of logic equation $D = 0$ have been formulated.

REFERENCES

1. Crama, Y. & Hammer, P.L. (2011) *Boolean functions: theory, algorithms, and applications*. Cambridge, New York: Cambridge University Press. pp. 440.
2. Andreeva, V. (2010) Test minimization technique for multiple stuck-at faults of combinational circuits. *Proc*. 8*th East-West Design&Test International Symposium*. St. Petersburg. pp. 168-170.
3. Zakrevskiy, A.D. (2003) *Logical equations*. Moscow: Editorial URSS.
4. Sorudeykin, K. & Andreeva, V. (2014) Decomposition Tree – based Compaction Procedure with Iteration Steps for Interconversional Layouts of Tasks. *Proc.* 12*th IEEE East-West Design & Test Symposium.* Kiev. pp. 173-178.
5. Měchura, T. (2008) *Random Circuits Generators*. [Online] Available from: http://ddd.fit.cvut.cz/prj/Circ_Gen

ВЕСТНИК ТОМСКОГО ГОСУДАРСТВЕННОГО УНИВЕРСИТЕТА
2015 Управление, вычислительная техника и информатика № 4 (33)

УДК 681.518.5:004.052.32
DOI: 10.17223/19988605/33/9

Д.В. Ефанов

АНАЛИЗ СПОСОБОВ ПОСТРОЕНИЯ КОДОВ С СУММИРОВАНИЕМ С УЛУЧШЕННЫМИ ХАРАКТЕРИСТИКАМИ ОБНАРУЖЕНИЯ СИММЕТРИЧНЫХ ОШИБОК В ИНФОРМАЦИОННЫХ ВЕКТОРАХ

Предложено несколько способов построения модифицированных кодов с суммированием, обладающих улучшенными характеристиками обнаружения ошибок в информационных векторах в сравнении с классическими кодами Бергера. Показано, что новые коды с суммированием обнаруживают примерно вдвое больше ошибок, чем классические коды Бергера. Установлено условие, при котором модифицированные коды Бергера имеют такое же количество контрольных разрядов, как и коды Бергера. Освещены схемотехнические особенности синтеза генераторов контрольных разрядов модифицированных кодов с суммированием.
Ключевые слова: техническая диагностика; система функционального контроля; код Бергера; модифицированный код с суммированием; информационный вектор; необнаруживаемая ошибка; симметричная ошибка; свойства обнаружения ошибок.

Принципы помехоустойчивого кодирования используются при передаче информации и при синтезе надежных систем управления [1–7]. Часто для этих целей применяют коды с суммированием [8–10]. Простота их построения (данные коды являются систематическими, т.е. значения разрядов контрольного вектора вычисляются по заранее установленным правилам суммирования разрядов информационного вектора) определяет их частое приложение в задачах технической диагностики при аппаратном и программном контроле технического состояния [11–15].

На рис. 1 приведена структура системы функционального контроля, организованная по кодам с суммированием. Блок $F(x)$ является контролируемым логическим устройством и имеет m выходов $f_1 f_2 \ldots f_m$, соответствующих выходному рабочему вектору $<f_1 f_2 \ldots f_m>$ (информационному вектору длины m). Для организации контроля технического состояния в систему вносится избыточность: блок $F(x)$ снабжается специализированным контрольным оборудованием в составе блока контрольной логики $G(x)$ и самопроверяемого тестера (*TSC*). Блок контрольной логики $G(x)$ имеет k выходов и вычисляет значения системы контрольных функций $<g_1\ g_2 \ldots\ g_k>$ (формирует контрольный вектор длины k), а самопроверяемый тестер проверяет факт соответствия значений разрядов информационного и контрольного векторов, которое устанавливается на этапе проектирования системы функционального контроля. При наличии неисправностей в любом из блоков системы функционального контроля на ее контрольных выходах формируется непарафазный сигнал <00> или <11> [9].

От выбранного на этапе проектирования системы функционального контроля кода с суммированием зависят такие важные характеристики самой системы, как аппаратурная избыточность и свойства обнаружения ошибок в блоке $F(x)$. Они, в свою очередь, напрямую влияют на энергопотребление системы и быстродействие [16].

Известно большое количество правил построения кодов, которые можно свести в диаграмму, приведенную на рис. 2. Между разрядами информационного вектора может быть установлено неравноправие, что делается приписыванием им различных последовательностей весовых коэффициентов $[w_1, w_2, \ldots, w_m]$, либо же все информационные разряды могут быть равноправными [17, 18]. При построении кодов анализируется вес информационного вектора W, для него может определяться наименьший неотрицательный вычет по заранее установленному модулю M, а могут также проводиться некоторые модификации [19–24].

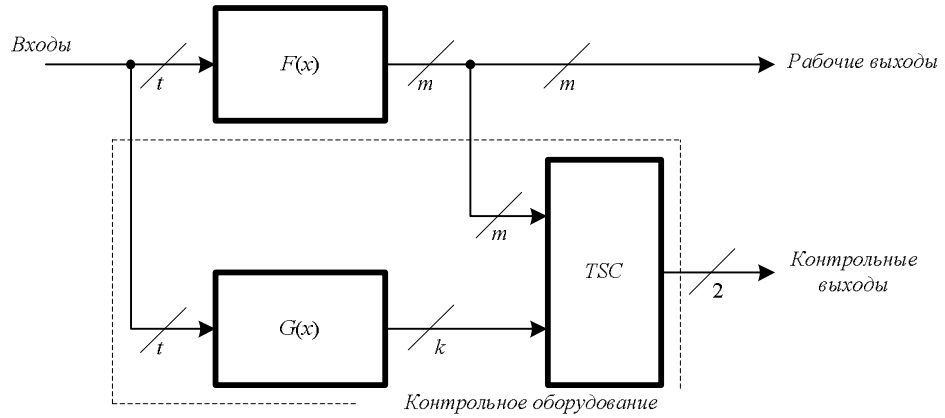


Рис. 1. Структура системы функционального контроля

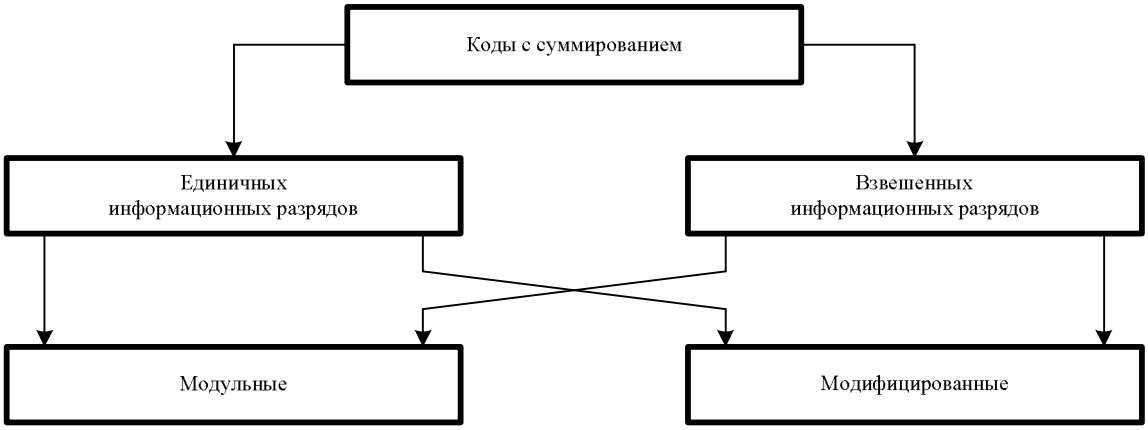


Рис. 2. Классификация кодов с суммированием

Данная работа посвящена изложению нескольких способов построения модифицированных кодов с суммированием, позволяющих улучшить возможности обнаружения ошибок классическими кодами с суммированием. При этом сохраняются все основные особенности классических кодов с суммированием.

1. Задача построения кода с суммированием с улучшенными характеристиками обнаружения ошибок в информационных векторах

Классический код с суммированием, или код Бергера [17], содержит в разрядах контрольного вектора двоичный эквивалент количества единичных разрядов информационного вектора (веса r информационного вектора). Исходя из этого, количество разрядов контрольного вектора кода Бергера $k = \lceil \log_2 (m+1) \rceil$, где m – длина информационного вектора, а запись $\lceil ... \rceil$ обозначает целое сверху от вычисляемого значения.

Обозначим код Бергера как $S(m,k)$-код.

Каждому информационному вектору с весом r соответствуют контрольные векторы с одинаковыми значениями разрядов. Поскольку всего существует C_m^r информационных векторов с весом r, их распределение между контрольными векторами с различными значениями разрядов крайне неравномерно (с увеличением значения m оно стремится к нормальному распределению). Например, в табл. 1 приведены все информационные векторы $S(5,3)$-кода, распределенные между контрольными векторами.

От того, как распределены информационные векторы между контрольными векторами, зависят свойства обнаружения ошибок кодом. Любая ошибка в информационном векторе, переводящая его в информационный вектор с тем же контрольным вектором, кодом обнаружена не будет. $S(m,k)$-кодами

не обнаруживается 100% разнонаправленных ошибок четной кратностью в информационных векторах, которые происходят при одинаковом количестве искажений 0 и 1 (так называемые симметричные ошибки [25]). В [26] предложена формула расчета количества необнаруживаемых ошибок в информационных векторах кодов Бергера:

$$N_m = \sum_{d=2}^{m,(m-1)} \left(\sum_{r=\frac{d}{2}}^{m-\frac{d}{2}} C_m^r C_m^{\frac{d}{2}} C_{m-r}^{\frac{d}{2}} \right). \quad (1)$$

Таблица 1

Распределение информационных векторов S(5,3)-кода между контрольными векторами

Вес							
0	1	2	3	4	5	6	7
Контрольный вектор							
000	**001**	**010**	**011**	**100**	**101**	**110**	**111**
00000	00001	00011	00111	01111	11111		
	00010	00101	01011	10111			
	00100	00110	01101	11011			
	01000	01001	01110	11101			
	10000	01010	10011	11110			
		01100	10101				
		10001	10110				
		10010	11001				
		10100	11010				
		11000	11100				

Данная формула позволяет подсчитать общее количество необнаруживаемых ошибок для $S(m,k)$-кода по каждой четной кратности d. В [27] эта формула приведена к виду

$$N_m = C_{2m}^m - 2^m. \quad (2)$$

Используя формулу (2) для $S(5,3)$-кода, получаем

$$N_5 = C_{2\cdot 5}^5 - 2^5 = C_{10}^5 - 2^5 = \frac{10!}{5!\cdot 5!} - 32 = 220.$$

Основной особенностью $S(m,k)$-кодов, определяющей их широкое применение при построении надежных систем автоматики, является стопроцентное обнаружение монотонных ошибок в информационных векторах (при таких ошибках искажаются либо только 0, либо только 1). С использованием данного свойства в [8, 28, 29] разработаны методы синтеза систем функционального контроля со стопроцентным обнаружением одиночных неисправностей в контролируемых схемах.

Следует, однако, отметить, что $S(m,k)$-коды неэффективно используют свои контрольные разряды: контрольные векторы с близким к 0 и m значением веса соответствуют малому количеству информационных векторов, а значит, дают малое количество необнаруживаемых ошибок; некоторые контрольные векторы, в зависимости от значения m, могут не формироваться, а значит, и не использоваться (табл. 1). В [30] приводится способ построения кода Бергера с улучшенными характеристиками обнаружения ошибок в информационных векторах, но он приводит к появлению некоторого количества монотонных необнаруживаемых ошибок, а значит, накладывает ограничения на применение данных кодов при построении надежных систем автоматики [31, 32].

Информационные векторы, меняя правила построения кода, можно перераспределить между контрольными векторами так, чтобы сохранялись все основные свойства классических кодов с суммированием. Например, некоторым информационным векторам из групп, соответствующих контрольным векторам со значением веса, близким к $m/2$, можно поставить в соответствие неиспользуемые контрольные векторы (см. табл. 2). Количество необнаруживаемых ошибок уменьшится, а свойства стопроцентного обнаружения монотонных ошибок сохранятся. Например, $S(5,3)$-кодом не обнаруживается 240 ошибок в информационных векторах, а кодом, которому соответствует табл. 2, – 180 ошибок, что в 1,33 раза меньше. При этом все необнаруживаемые ошибки будут являться симметричными.

Таблица 2

Перераспределение информационных векторов S(5,3)-кода между контрольными векторами

Вес							
0	1	2	3	4	5	6	7
Контрольный вектор							
000	**001**	**010**	**011**	**100**	**101**	**110**	**111**
00000	00001	00011	00111	01111	11111	01100	10101
	00010	00101	01011	10111		10001	10110
	00100	00110	01101	11011		10010	11001
	01000	01001	01110	11101		10100	11010
	10000	01010	10011	11110		11000	11100

2. Способы модификации кодов с суммированием

Формализованные правила построения кодов с суммированием с улучшенными характеристиками обнаружения симметричных ошибок в информационных векторах даны в нескольких работах, например [27, 33, 34]. Каждый из описанных способов основан на «сдвиге» части информационных векторов в группы, соответствующие контрольным векторам, большего веса, чем истинный вес информационных векторов.

В [27] описывается процедура подсчета модифицированного веса W информационного вектора следующим образом (автор данной работы называет данный «сдвиговый» код комбинированным кодом Бергера).

Алгоритм 1. Получение контрольного вектора комбинированного кода Бергера.

1. Подсчитывается вес m–1 информационного разряда (информационного вектора за исключением одного информационного разряда – вектора $<f_1 f_2 \ldots f_{m-1}>$).

2. Вычисляется функция паритета по всем разрядам информационного вектора: $\alpha = f_1 \oplus f_2 \oplus \ldots \oplus f_m$.

3. Старшему разряду контрольного вектора $<g_1\, g_2 \ldots g_k>$ соответствует значение функции паритета α, а оставшимся младшим разрядам – значение веса m–1 информационного разряда.

Обозначим код, получаемый по алгоритму 1, как $KS(m,k)$-код. Для информационных векторов длины $m = 5$ применение алгоритма 1 дает распределение информационных векторов между контрольными векторами, показанное в табл. 3.

Общее количество необнаруживаемых ошибок в $KS(m,k)$-кодах в два раза больше, чем в соответствующих $S(m–1,k^*)$-кодах ($k^*=k$ или k–1), и его можно определить по формуле [27]:

$$N_m = 2C_{2m-2}^{m-1} - 2^m. \qquad (3)$$

Для приведенного в табл. 3 $KS(6,3)$-кода формула (3) дает такой результат:

$$N_5 = 2C_{2\cdot5-2}^{5-1} - 2^5 = 2C_8^4 - 2^5 = 2\cdot\frac{8!}{4!\cdot4!} - 32 = 108.$$

Полученный результат для $S^*(m,k)$-кода в 2,037 раза меньше, чем для $S(m,k)$-кода.

Таблица 3

Распределение информационных векторов комбинированного KS(6,3)-кода между контрольными векторами

Вес															
0	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15
Контрольный вектор															
0000	**0001**	**0010**	**0011**	**0100**	**0101**	**0110**	**0111**	**1000**	**1001**	**1010**	**1011**	**1100**	**1101**	**1110**	**1111**
00000	10001	00011	10111	01111				10000	00001	10011	00111	11111			
	10010	00101	11011						00010	10101	01011				
	10100	00110	11101						00100	10110	01101				
	11000	01001	11110						01000	11001	01110				
		01010								11010					
		01100								11100					

Информационные векторы классического кода Бергера были перераспределены между контрольными векторами, что дало новый код. Фактически половина информационных векторов с функцией паритета $\alpha = f_1 \oplus f_2 \oplus ... \oplus f_m$, равной 1, поменяли свои места – они оказались сдвинутыми на 8 контрольных групп «вправо», а оставшаяся половина информационных векторов сохранила свои позиции (см. табл. 3). При этом добавился один контрольный разряд, а характер распределения был сохранен: информационные векторы неравномерно распределены между контрольными векторами, контрольные разряды используются неэффективно, часть из контрольных векторов никогда не формируется. Для подсчета коэффициента α можно выбирать различные сочетания информационных разрядов вместе с тем разрядом, для которого не применяется правило 1 алгоритма 1. Это даст возможность получения нового множества кодов с суммированием с различными распределениями необнаруживаемых ошибок по кратностям и видам. При этом основные свойства кодов Бергера могут и не сохраниться.

Правила построения, выбранного в качестве основы системы функционального контроля кода с суммированием, определяют и способ построения тестера. Тестер в структуре, приведенной на рис. 1, состоит из генератора (кодера или шифратора) контрольных разрядов, выходы которого соединяются с входами компаратора. Последний осуществляет сравнение одноименных контрольных функций вычисленными блоками генератора и $G(x)$. Компаратор имеет стандартную структуру и строится на основе каскадного соединения k–1 модуля сравнения парафазных сигналов [35]. Структура же генератора определяется правилами построения кода с суммированием.

Для синтеза генератора $KS(m,k)$-кода требуется использование счетчика единиц m–1 разряда и формирование функции паритета. На рис. 3 изображены схемы генераторов $S(5,3)$-кода и $KS(6,3)$-кода. Младший разряд счетчика единиц m–1 разряда является функцией паритета [9, 10], а значит, достаточно использовать один сумматор по модулю 2 для формирования старшего контрольного разряда. Общая структура генераторов $KS(m,k)$-кодов дана на рис. 4.

Анализ показывает, что $KS(m,k)$-код сохраняет свойства обнаружения ошибок, если коэффициент α равен единственному неконтролируемому информационному разряду (см. рис. 5). Это следует из свойства функции сложения по модулю 2 – она принимает равное количество единичных и нулевых значений на всех 2^m входных наборах.

Свойства кода никак не меняются, а сложность технической реализации генератора уменьшается. Это позволяет предложить другой алгоритм построения модифицированного кода с суммированием.

Алгоритм 2. Получение контрольного вектора модифицированного кода Бергера.

1. Подсчитывается вес m–1 информационного разряда (информационного вектора за исключением одного информационного разряда – вектора $<f_1 f_2 \dots f_{m-1}>$).

2. Старшему разряду контрольного вектора $<g_1 g_2 \dots g_k>$ соответствует значение неконтролируемого информационного разряда, а оставшимся младшим разрядам – значение веса m–1 информационного разряда.

Код, полученный по алгоритму 2, обозначим как $MS(m,k)$-код.

В табл. 4 дается распределение информационных векторов между контрольными векторами $MS(5,4)$-кода, при этом $\alpha = f_5$. Информационные векторы $KS(m,k)$-кода распределены иначе, чем у $MS(m,k)$-кода, однако общий характер сохранен. Генератор $MS(5,4)$-кода изображен на рис. 6, его можно сравнить с генераторами, приведенными на рис. 3.

Нетрудно заметить, что тот же результат дают и другие алгоритмы «сдвига» информационных векторов, например описанные в [21]. Они изоморфны тем правилам, которые описаны в [27]. Например, процедура построения $MS(m,k)$-кода может быть описана следующим образом.

Алгоритм 3. Построение взвешенного кода с суммированием.

1. Информационному вектору ставится в соответствие вектор весовых коэффициентов [1, 1, ..., 1, 0].

2. Подсчитывается вес W информационного вектора.

3. Число W представляется в двоичном виде и записывается в младшие разряды контрольного вектора.

4. В старший разряд контрольного вектора записывается значение разряда, для которого вес $w_m = 0$.

Этот же результат дает взвешивание одного информационного разряда. Коды с суммированием с одним взвешенным информационным разрядом детально описаны в [36, 37].

а *б*

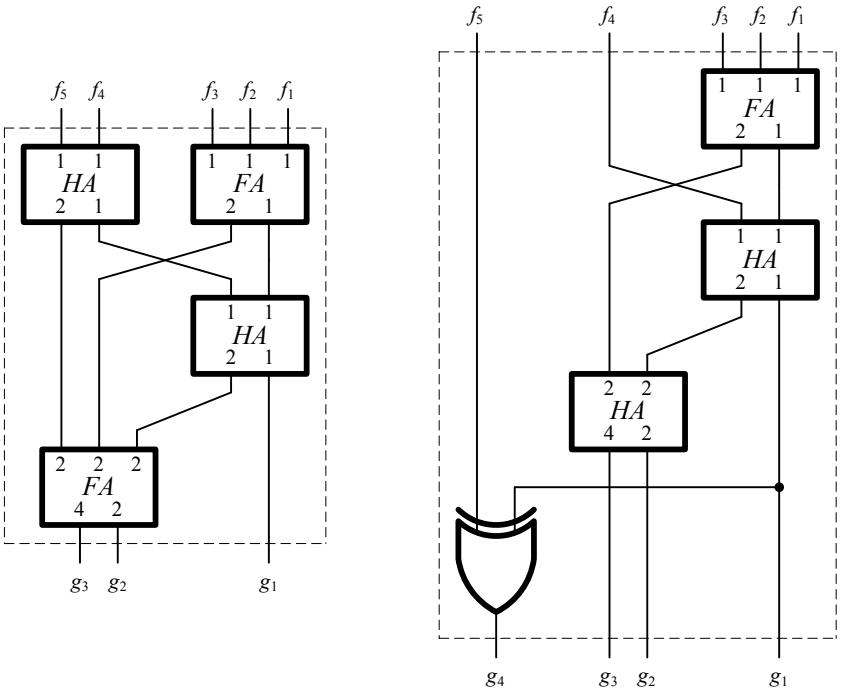


Рис. 3. Генераторы кодов с суммированием: *а* – *S*(5,3)-кода; *б* – *KS*(6,3)-кода

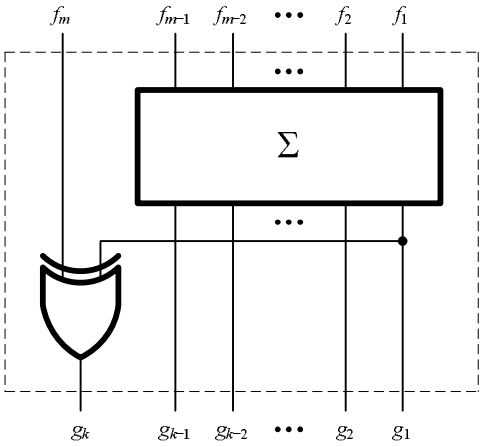


Рис. 4. Структура генераторов *KS*(*m*,*k*)-кодов

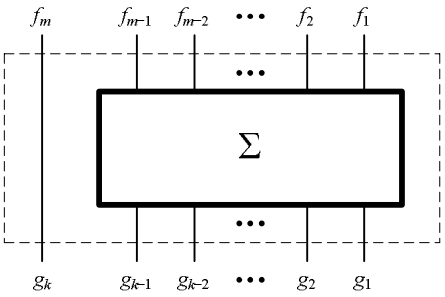


Рис. 5. Структура генераторов *MS*(*m*,*k*)-кодов

Таблица 4

Распределение информационных векторов MS(5,4)-кода между контрольными векторами

Вес															
0	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15
Контрольный вектор															
0000	**0001**	**0010**	**0011**	**0100**	**0101**	**0110**	**0111**	**1000**	**1001**	**1010**	**1011**	**1100**	**1101**	**1110**	**1111**
00000	00001	00011	00111	01111				10000	10001	10011	10111	11111			
	00010	00101	01011						10010	10101	11011				
	00100	00110	01101						10100	10110	11101				
	01000	01001	01110						11000	11001	11110				
		01010								11010					
		01100								11100					

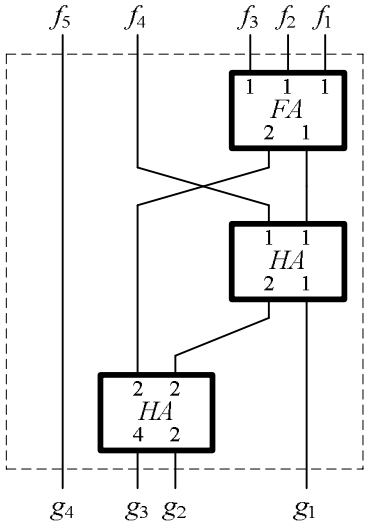


Рис. 6. Генератор *MS*(5,4)-кода

Анализируя алгоритм модификации кодов Бергера, приведенный в [30], можно предложить аналогичный приведенным выше способ построения кода с суммированием.

Алгоритм 4. Построение модифицированного кода Бергера.

1. Вес информационного вектора подсчитывается по модулю кода Бергера: $M = 2^{\lceil \log_2 (m+1) \rceil}$.

2. Вычисляется значение поправочного коэффициента α, равное значению одного произвольного информационного разряда.

3. Определяется модифицированный вес информационного вектора: $W = r+\alpha M$.

4. Число W представляется в двоичном виде и записывается в разряды контрольного вектора.

Общее количество необнаруживаемых ошибок для кодов, построение которых ведется по алгоритмам 2–4, также может быть определено по формуле (3).

Следует заметить, что приведенные способы построения модифицированных кодов Бергера улучшают свойства обнаружения ошибок в информационных векторах кодом, однако в большинстве случаев приводят к увеличению количества контрольных разрядов на единицу по сравнению с кодом Бергера.

Утверждение 1. Модифицированный код Бергера будет всегда обнаруживать 100% любых видов ошибок в информационных векторах, кроме некоторой доли симметричных ошибок, которая, однако, будет меньше, чем у кодов Бергера, и при этом будет иметь такое же количество контрольных разрядов при выполнении следующего условия:

$$m = 2^t \quad \left(t \in \{1,2,...\}\right). \tag{4}$$

Формулировка утверждения 1 следует непосредственно из алгоритмов 1–4. Только в одном случае, когда сохраняется условие (4), не потребуется при модификации кода добавления еще одного контрольного разряда. В самом деле, в табличной форме задания кода с суммированием (табл. 4) заполняются только m+1 контрольные группы, при этом общее их количество равно

$$2^{\lceil \log_2 (m+1) \rceil} = 2^{\lceil \log_2 (2^t+1) \rceil} = 2^{t+1}, \tag{5}$$

где $t = \log_2 m$.

Подставим в (5) значение t:

$$2^{t+1} = 2^{\log_2 m+1} = 2 \cdot 2^{\log_2 m} = 2m. \tag{6}$$

Из выражения (6), таким образом, следует, что количество контрольных групп при выполнении условия (4) будет равным $2m$. Изначально заполнены m+1 группы. В контрольной группе, соответствующей максимальному весу информационного вектора, в группе с номером m будет присутствовать только один информационный вектор. При модификации по любому из предложенных алгоритмов сдвигается вправо по таблице задания кода (см. табл. 4) ровно половина информационных векторов,

включая и данный вектор, заполняя тем самым оставшуюся m–1 группу. Добавления нового контрольного разряда не требуется.

Этим объясняется утверждение 1.

Определим, во сколько раз уменьшается количество симметричных необнаруживаемых ошибок у модифицированных кодов Бергера с увеличением длины информационного вектора:

$$\vartheta = \frac{C_{2m}^{m} - 2^{m}}{2C_{2m-2}^{m-1} - 2^{m}}. \tag{7}$$

Перейдем к пределу:

$$\lim_{m\to\infty} \vartheta = \lim_{m\to\infty} \frac{C_{2m}^{m} - 2^{m}}{2C_{2m-2}^{m-1} - 2^{m}} = \lim_{m\to\infty} \frac{\dfrac{(2m)!}{m!\cdot m!} - 2^{m}}{2\cdot\dfrac{(2m-2)!}{(m-1)!\cdot(m-1)!} - 2^{m}} =$$

$$= \lim_{m\to\infty} \frac{\dfrac{\sqrt{2\pi}e^{-2m}(2m)^{2m+\frac{1}{2}}}{\sqrt{2\pi}e^{-m}m^{m+\frac{1}{2}}\cdot\sqrt{2\pi}e^{-m}m^{m+\frac{1}{2}}} - 2^{m}}{2\cdot\dfrac{\sqrt{2\pi}e^{-2m-2}(2m-2)^{2m-2+\frac{1}{2}}}{\sqrt{2\pi}e^{-m-1}(m-1)^{m-1+\frac{1}{2}}\cdot\sqrt{2\pi}e^{-m-1}(m-1)^{m-1+\frac{1}{2}}} - 2^{m}} = \lim_{m\to\infty} \frac{\dfrac{(2m)^{2m+\frac{1}{2}}}{\left(m^{m+\frac{1}{2}}\right)^{2}} - 2^{m}}{2\cdot\dfrac{(2m-2)^{2m-2+\frac{1}{2}}}{\left((m-1)^{m-1+\frac{1}{2}}\right)^{2}} - 2^{m}} =$$

$$= \lim_{m\to\infty} \frac{\dfrac{(2m)^{2m+\frac{1}{2}}}{2^{m}\left(m^{m+\frac{1}{2}}\right)^{2}} - 1}{2\cdot\dfrac{(2m-2)^{2m-2+\frac{1}{2}}}{2^{m}\left((m-1)^{m-1+\frac{1}{2}}\right)^{2}} - 1} = \lim_{m\to\infty} \frac{\dfrac{(2m)^{2m+\frac{1}{2}}}{m^{m+\frac{1}{2}+m+\frac{1}{2}}}}{2\cdot\dfrac{(2m-2)^{2m-2+\frac{1}{2}}}{(m-1)^{m-1+\frac{1}{2}+m-1+\frac{1}{2}}}} = \lim_{m\to\infty} \frac{2^{2m+\frac{1}{2}}m^{2m+\frac{1}{2}-2m-1}}{2\cdot 2^{2m-2+\frac{1}{2}}(m-1)^{2m-2+\frac{1}{2}-2m+2-1}} =$$

$$= \lim_{m\to\infty} \frac{2^{2m+\frac{1}{2}}m^{-\frac{1}{2}}}{2\cdot 2^{2m-2+\frac{1}{2}}(m-1)^{-\frac{1}{2}}} = \lim_{m\to\infty} 2^{2m+\frac{1}{2}-1-2m+2-\frac{1}{2}} \frac{\dfrac{1}{\sqrt{m}}}{\dfrac{1}{\sqrt{m}-1}} = 2\lim_{m\to\infty}\sqrt{\frac{m-1}{m}} = 2\lim_{m\to\infty}\sqrt{1-\frac{1}{m}} = 2. \tag{8}$$

Таким образом, справедливо следующее положение.

Утверждение 2. Отношение количества необнаруживаемых симметричных ошибок в классических и модифицированных кодах Бергера с увеличением длины информационного вектора уменьшается и в пределе при $m \to \infty$ стремится к 2.

Таким образом, модифицированные коды Бергера примерно в два раза эффективнее обнаруживают ошибки, чем классические коды Бергера.

Анализируя схемотехнические особенности контрольного оборудования систем функционального контроля, заметим, что генераторы модифицированных кодов Бергера являются более простыми, чем генераторы классических $S(m,k)$-кодов. Но в системе функционального контроля во всех случаях, кроме случая (4), блок контрольной логики $G(x)$ будет реализовывать еще одну контрольную функцию, а компаратору, соответственно, потребуется на один модуль сравнения парафазных сигналов больше, чем компаратору классического кода Бергера.

В табл. 6 приводятся результаты экспериментальных исследований модифицированных кодов Бергера в сравнении с классическими кодами Бергера в системах функционального контроля на ряде контрольных комбинационных схем из набора LGSynth`89 [38]. В эксперименте оценивалась сложность технической реализации систем функционального контроля, организованных по кодам с суммированием (табл. 5). Показателем сложности технической реализации является площадь, занимаемая устройством на кристалле. Для анализа данного показателя была выбрана библиотека функциональных элементов stdcell2_2.genlib. Контрольные комбинационные схемы в LGSynth`89 представлены в виде фай-

лов списочной формы задания структуры логического устройства – *net*-листа. С использованием специально разработанного программного обеспечения были получены *net*-листы всех блоков системы функционального контроля, а затем с применением известного интерпретатора SIS [39] получены значения площадей в условных единицах. В последней графе табл. 5 представлено значение величины ζ – отношение площадей систем функционального контроля, организованных по классическим и модифицированным кодам Бергера. Анализ данного коэффициента показывает, что использование модифицированных кодов Бергера эффективно по показателю сложности технической реализации в том случае, если выполняется полученное выше условие (4). Этот результат был ожидаем, однако не следует забывать, что характеристики обнаружения ошибок в системах функционального контроля, организованных по модифицированным кодам Бергера, улучшены, а увеличение сложности есть «цена» за это свойство.

Таблица 5

Сложность технической реализации систем функционального контроля

Контрольная комбинационная схема	**Число выходов**	**Площадь системы функционального контроля**		
		***S*(*m*,*k*)-код**	***MS*(*m*,*k*)-код**	ζ
cm42a	10	8928	10712	0,833
f51m	8	13112	10152	1,292
pm1	13	16144	20880	0,773
x2	7	4096	4856	0,843
z4ml	4	4192	2976	1,409

Общий недостаток всех приведенных алгоритмов заключается в том, что не дает возможности построения кода, эффективно использующего свои контрольные разряды. Повышение эффективности возможно за счет применения модулей представления веса, меньших, чем модуль кода Бергера. Например, в табл. 6 дается распределение информационных векторов между контрольными векторами *MS*(*m*,*k*)-кода, у которого для подсчета веса информационного вектора используется модуль $M = 2^{\lceil \log_2 (m+1) \rceil - 1}$. Применение модуля дает уменьшение количества контрольных разрядов. Но при этом появляется некоторое количество монотонных ошибок, все они имеют кратность $d = M$ (векторы, допускающие монотонные ошибки, в табл. 6 выделены серым фоном). Это свойство можно использовать при построении надежных устройств автоматики и вычислительной техники [32].

Таблица 6

Распределение информационных векторов MS(5,3)-кода между контрольными векторами

Вес							
0	1	2	3	4	5	6	7
Контрольный вектор							
000	**001**	**010**	**011**	**100**	**101**	**110**	**111**
00000	00001	00011	00111	10111	10000	10001	10011
01111	00010	00101	01011	11011	11111	10010	10101
	00100	00110	01101	11101		10100	10110
	01000	01001	01110	11110		11000	11001
		01010					11010
		01100					11100

Заключение

В статье приводятся способы модификации кода Бергера, позволяющие построить модифицированный код с суммированием, обладающий основным свойством кода Бергера – способностью идентифицировать все монотонные ошибки. Модифицированные коды Бергера обнаруживают, кроме того, некоторое количество симметричных ошибок (коды Бергера не обнаруживают стопроцентных симметричных ошибок [25, 32, 40]). Это свойство можно использовать для уменьшения избыточности систем функционального контроля при применении известных алгоритмов модификации контролируемых схем [8, 28, 29].

Предложенные алгоритмы построения модифицированных кодов с суммированием, в отличие от описанных в [27], дают более простые структуры генераторов кодов с суммированием и в некоторых случаях более простые структуры систем функционального контроля. Установлено, что максимальный эффект при обнаружении ошибок в информационных векторах кодов при минимальных аппаратурных затратах достигается для кодов с длинами информационных векторов $m = 2^t \ \left(t \in \{1,2,...\}\right)$.

Стоит отметить, что описанные коды с суммированием не являются оптимальными в смысле использования контрольных разрядов, а значит, правила построения кодов могут быть «усовершенствованы», например, могут быть использованы свойства реальных схем, на выходах которых формируются не все 2^m информационных векторов, и код может формироваться для некоторого множества информационных векторов.

ЛИТЕРАТУРА

1. *Сапожников В.В., Сапожников Вл.В., Христов Х.А., Гавзов Д.В.* Методы построения безопасных микроэлектронных систем железнодорожной автоматики. М. : Транспорт, 1995. 272 с.
2. *Микропроцессорные* системы централизации / Вл.В. Сапожников, В.А. Кононов, С.А. Куренков, А.А. Лыков, О.А. Наседкин, А.Б. Никитин, А.А. Прокофьев, М.С. Трясов ; под ред. Вл.В. Сапожникова. М. : Учебно-методический центр по образованию на железнодорожном транспорте, 2008. 398 с.
3. *Ubar R., Raik J., Vierhaus H.-T.* Design and Test Technology for Dependable Systems-on-Chip (Premier Reference Source) // Information Science Reference. Hershey ; New York : IGI Global, 2011. 578 p.
4. *Рабочее* диагностирование безопасных информационно-управляющих систем / А.В. Дрозд, В.С. Харченко, С.Г. Антощук, Ю.В. Дрозд, М.А. Дрозд, Ю.Ю. Сулима ; под ред. А.В. Дрозда и В.С. Харченко. Харьков : Национальный аэрокосмический университет им. Н.Е. Жуковского «ХАИ», 2012. 614 с.
5. *Кравцов Ю.А., Архипов Е.В., Бакин М.Е.* Перспективные способы кодирования рельсовых цепей тональной частоты // Автоматика на транспорте. 2015. Т. 1, № 2. С. 119–126.
6. *Ходаковский В.А., Ходаковский Т.В.* Мера сходства узкополосных сигналов // Автоматика на транспорте. 2015. Т. 1, № 2. С. 180–194.
7. *Шаманов В.И.* Управление процессом модернизации комплексов систем железнодорожной автоматики и телемеханики // Автоматика на транспорте. 2015. Т. 1, № 3. С. 237–250.
8. *Согомонян Е.С., Слабаков Е.В.* Самопроверяемые устройства и отказоустойчивые системы. М. : Радио и связь, 1989. 208 с.
9. *Сапожников В.В., Сапожников Вл.В.* Самопроверяемые дискретные устройства. СПб. : Энергоатомиздат, 1992. 224 с.
10. *Piestrak S.J.* Design of Self-Testing Checkers for Unidirectional Error Detecting Codes. Wrocław : Oficyna Wydawnicza Politechniki Wrocławskiej, 1995. 111 p.
11. *Fujiwara E.* Code Design for Dependable Systems: Theory and Practical Applications. John Wiley & Sons, 2006. 720 p.
12. *Matrosova A.Yu., Levin I., Ostanin S.A.* Self-Checking Synchronous FSM Network Design with Low Overhead // VLSI Design. 2000. Vol. 11, Issue 1. P. 47–58.
13. *Lala P.K.* Self-Checking and Fault-Tolerant Digital Design. San Francisco : Morgan Kaufmann Publishers, 2001. 216 p.
14. *Matrosova A., Levin I., Ostanin S.* Survivable Self-Checking Sequential Circuits // Proc. of 2001 IEEE International Symposium on Defect and Fault Tolerance in VLSI Systems (DFT 2001). Oct. 24–26. San Francisco, CA, 2001. P. 395–402.
15. *Abdulhadi A.H., Maamar A.H.* Self Checking Register File Using Berger Code // 6th WSEAS International Conference on Circuits, systems, control & signal processing. 2007. Cairo, Egypt. December 29–31. P. 62–68.
16. *Cheremisinov D., Cheremisinova L.* Low-Power Design of Combinational CMOS Networks // Proceedings of 11th IEEE East-West Design & Test Symposium (EWDTS`2013). Rostov-on-Don, Russia, September 27–30, 2013. P. 208–211.
17. *Berger J.M.* A Note on Error Detecting Codes for Asymmetric Channels // Information and Control. 1961. Vol. 4, Issue 1. P. 68–73.
18. *Das D., Touba N.A.* Weight-Based Codes and Their Application to Concurrent Error Detection of Multilevel Circuits // Proc. 17th IEEE Test Symposium. USA, California, 1999. P. 370–376.
19. *Das D., Touba N.A., Seuring M., Gossel M.* Low Cost Concurrent Error Detection Based on Modulo Weight-Based Codes // Proceedings of IEEE 6th International On-Line Testing Workshop (IOLTW), Spain, Palma de Mallorca, July 3–5, 2000. P. 171–176.
20. *Mehov V., Saposhnikov V., Sapozhnikov Vl., Urganskov D.* Concurrent Error Detection Based on New Code with Modulo Weighted Transitions between Information Bits // Proceedings of 7th IEEE East-West Design & Test Workshop (EWDTW`2007). Erevan, Armenia, September 25–30, 2007. P. 21–26.
21. *Мехов В.Б., Сапожников В.В., Сапожников Вл.В.* Контроль комбинационных схем на основе модифицированных кодов с суммированием // Автоматика и телемеханика. 2008. № 8. С. 153–165.
22. *Блюдов А.А., Сапожников В.В., Сапожников Вл.В.* Модифицированный код с суммированием для организации контроля комбинационных схем // Автоматика и телемеханика. 2012. № 1. С. 169–77.
23. *Блюдов А.А., Ефанов Д.В., Сапожников В.В., Сапожников Вл.В.* Коды с суммированием для организации контроля комбинационных схем // Автоматика и телемеханика. 2013. № 6. С. 153–164.
24. *Блюдов А.А., Ефанов Д.В., Сапожников В.В., Сапожников Вл.В.* О кодах с суммированием единичных разрядов в системах функционального контроля // Автоматика и телемеханика. 2014. № 8. С. 131–145.

25. *Сапожников В.В., Сапожников Вл.В., Ефанов Д.В.* Классификация ошибок в информационных векторах систематических кодов // Известия вузов. Приборостроение. 2015. Т. 58, № 5. С. 333–343.
26. *Ефанов Д.В., Сапожников В.В., Сапожников Вл.В.* О свойствах кода с суммированием в схемах функционального контроля // Автоматика и телемеханика. 2010. № 6. С. 155–162.
27. *Черкасова Т.Х.* Обнаружение ошибок в системах автоматики и вычислительной техники с помощью кодов Бергера и его модификаций // Проблемы безопасности и надежности микропроцессорных комплексов : сб. тр. науч.-практ. конф. / под ред. Вал.В. Сапожникова. СПб. : ПГУПС, 2015.
28. *Busaba F.Y., Lala P.K.* Self-Checking Combinational Circuit Design for Single and Unidirectional Multibit Errors // Journal of Electronic Testing: Theory and Applications. 1994. Issue 5. P. 19–28.
29. *Morosow A., Sapozhnikov V.V., Sapozhnikov Vl.V., Goessel M.* Self-Checking Combinational Circuits with Unidirectionally Independent Outputs // VLSI Design. 1998. Vol. 5, Issue 4. P. 333–345.
30. *Блюдов А.А., Ефанов Д.В., Сапожников В.В., Сапожников Вл.В.* Построение модифицированного кода Бергера с минимальным числом необнаруживаемых ошибок информационных разрядов // Электронное моделирование. 2012. Т. 34, № 6. С. 17–29.
31. *Сапожников В.В., Сапожников Вл.В., Ефанов Д.В.* Об использовании свойств кодов с суммированием по обнаружению монотонных ошибок в системах функционального контроля комбинационных схем // Вестник Томского государственного университета. Управление, вычислительная техника и информатика. 2014. № 3. С. 76–88.
32. *Сапожников В.В., Сапожников Вл.В., Ефанов Д.В.* Применение кодов с суммированием при синтезе систем железнодорожной автоматики и телемеханики на программируемых логических интегральных схемах // Автоматика на транспорте. 2015. Т. 1, № 1. С. 84–107.
33. *Sapozhnikov V., Sapozhnikov Vl., Efanov D., Nikitin D.* Combinational Circuits Checking on the Base of Sum Codes with One Weighted Data Bit // Proceedings of 12th IEEE East-West Design & Test Symposium (EWDTS`2014). Kyev, Ukraine, September 26–29, 2014. P. 126–136.
34. *Сапожников В.В., Сапожников Вл.В., Ефанов Д.В.* Построение кодов с суммированием с наименьшим количеством необнаруживаемых симметричных ошибок в информационных векторах // Радиоэлектроника и информатика. 2014. № 4. С. 46–55.
35. *Мельников А.Г., Сапожников В.В., Сапожников Вл.В.* Синтез самопроверяющихся тестеров для кодов с суммированием // Проблемы передачи информации. 1986. Т. 22, № 2. С. 85–97.
36. *Сапожников В.В., Сапожников Вл.В., Ефанов Д.В., Никитин Д.А.* Метод построения кода Бергера с повышенной эффективностью обнаружения ошибок в информационных разрядах // Электронное моделирование. 2013. Т. 35, № 4. С. 21–34.
37. *Сапожников В.В., Сапожников Вл.В., Ефанов Д.В., Никитин Д.А.* Исследование свойств кодов с суммированием с одним взвешенным информационным разрядом в системах функционального контроля // Электронное моделирование. 2015. Т. 37, № 1. С. 25–48.
38. *Collection* of Digital Design Benchmarks. URL: http://ddd.fit.cvut.cz/prj/Benchmarks
39. *Sentovich E.M., Singh K.J., Lavagno L., Moon C., Murgai R., Saldanha A., Savoj H., Stephan P.R., Brayton R.K., Sangiovanni-Vincentelli A.* SIS: A System for Sequential Circuit Synthesis // Electronics Research Laboratory, Department of Electrical Engineering and Computer Science. University of California, Berkeley, 4 May 1992. 45 p.
40. *Сапожников В.В., Сапожников Вл.В., Ефанов Д.В.* Обнаружение опасных ошибок на рабочих выходах комбинационных логических схем // Автоматика на транспорте. 2015. Т. 1, № 2. С. 195–211.

Ефанов Дмитрий Викторович, канд. техн. наук, доцент. E-mail: TrES-4b@yandex.ru
Петербургский государственный университет путей сообщения Императора Александра I (г. Санкт-Петербург)

Поступила в редакцию 7 сентября 2015 г.

Efanov Dmitry V. (Petersburg State Transport University, St. Petersburg, Russian Federation).

Analysis of formation methods of sum codes with improved characteristics of detection of symmetrical errors in data vectors

Keywords: technical diagnostics; concurrent error detection (CED); Berger code; modified sum code; data bits; undetectable error; symmetric error; detection ability properties.

DOI: 10.17223/19988605/32/9

The principles of noise-resistant coding are widely used for transmitting the information and synthesis of reliable management systems. Quite often, the sum codes are used for the purposes specified. Sum codes are systematic codes that allow to detect some error rate in data vectors. Sum codes have simple rules of formation and relatively low redundancy that determines the priority in its selection, in comparison with error correction codes.

Thus, using of the sum codes in systems of technical diagnostics results in a small hardware redundancy for providing the error detection characteristics.

Classic sum codes or Berger codes do not detect 100% of multidirectional errors with even multiplicity containing $\{0 \to 1, 1 \to 0\}$ distortion groups (100% of so-called symmetric errors). However, Berger codes detect 100% of monotonous distortions in data vectors that is used to transfer the information and to build the discrete devices with a 100% detection of single faults.

Analysis of the characteristics of Berger codes shows, that they often use the check bits inefficiently. For example, all possible check vectors are formed for Berger code only in special cases when data vector length is $m = 2^t - 1 \ \left(t \in \{1, 2, ...\}\right)$. In all other cases, a number of check vectors does not used at all. There is a problem of providing the modified Berger code which has all the characteristics of the classic Berger code. The article presents the algorithms for formation of modified Berger codes which detect almost twice as many symmetrical errors in data vectors, as classic Berger codes. The article also covers the conditions under which there is no need to

increase the number of check bits in the code. It also highlights the schematics particularities of synthesis of generators (encoders) of new codes. The experiment with making use of a set of test cases LGSynth`89 confirmed theoretical results.

Note that the modified Berger code is promising in solving the problems of technical diagnostics.

REFERENCES

1. Sapozhnikov, V.V., Sapozhnikov, Vl.V., Khristov, Kh.A. & Gavzov, D.V. (1995) *Metody postroeniya bezopasnykh mikroelektronnykh sistem zheleznodorozhnoy avtomatiki* [Methods for Constructing Safety Microelectronic Systems of Railway Automation]. Moscow: Transport.
2. Sapozhnikov, Vl.V. et al. (2008) *Mikroprotsessornye sistemy tsentralizatsii* [Microprocessor Interlocking System]. Moscow: Educational Center for Education in railway transport.
3. Ubar, R., Raik, J. & Vierhaus, H.-T. (2011) *Design and Test Technology for Dependable Systems-on-Chip (Premier Reference Source)*. Hershey; New York: IGI Global.
4. Drozd, A.V. & Kharchenko, V.S. (eds) (2012) *Rabochee diagnostirovanie bezopasnykh informatsionno-upravlyayushchikh sistem* [Objects and Methods of On-Line Testing for Safe Instrumentation and Control Systems]. Kharkov National Aerospace University.
5. Kravtsov, Yu.A, Arkhipov, E.V. & Bakin, M.E. (2015) Perspektivnye sposoby kodirovaniya rel'sovykh tsepey tonal'noy chastoty [Advanced Coding Schemes of Voice-Frequency Track Circuits]. *Avtomatika na transporte*. 1(2). pp. 119-126.
6. Khodakovskiy, V.A. & Khodakovskiy, T.V. (2015) Mera skhodstva uzkopolosnykh signalov [On Similarity Measure of Narrow Band Signals]. *Avtomatika na transporte*. 1(2). pp. 180-194.
7. Shamanov, V.I. (2015) Upravlenie protsessom modernizatsii kompleksov sistem zheleznodorozhnoy avtomatiki i telemekhaniki [Control of Railway Automation and Remote Control System Complexes Modernization Process]. *Avtomatika na transporte*. 1(3). pp. 237-250.
8. Sogomonyan, E.S. & Slabakov, E.V. (1989) *Samoproveryaemye ustroystva i otkazoustoychivye sistemy* [Self-Checking Devices and Fault-Tolerant Systems]. Moscow: Radio i svyaz'.
9. Sapozhnikov, V.V. & Sapozhnikov, Vl.V. (1992) *Samoproveryaemye diskretnye ustroystva* [Self-Checking Digital Devices]. St. Petersburg: Energoatomizdat.
10. Piestrak, S.J. (1995) *Design of Self-Testing Checkers for Unidirectional Error Detecting Co*des. Wrocław: Oficyna Wydawnicza Politechniki Wrocławskiej.
11. Fujiwara, E. (2006) *Code Design for Dependable Systems: Theory and Practical Applications*. John Wiley & Sons.
12. Matrosova, A.Yu., Levin, I. & Ostanin, S.A. (2000) Self-Checking Synchronous FSM Network Design with Low Overhead. *VLSI Design*. 11(1). pp. 47-58.
13. Lala, P.K. (2001) *Self-Checking and Fault-Tolerant Digital Design*. San Francisco: Morgan Kaufmann Publishers.
14. Matrosova, A., Levin, I. & Ostanin, S. (2001) Survivable Self-Checking Sequential Circuits. *Proc. of 2001 IEEE International Symposium on Defect and Fault Tolerance in VLSI Systems (DFT 2001)*. 24th – 26th October 2001. San Francisco, CA. pp. 395-402.
15. Abdulhadi, A.H. & Maamar, A.H. (2007) Self Checking Register File Using Berger Code. *6th WSEAS International Conference on Circuits, systems, control & signal processing*. 29th – 31st December 2007. Cairo, Egypt. pp. 62-68.
16. Cheremisinov, D. & Cheremisinova, L. (2013) Low-Power Design of Combinational CMOS Networks. *Proceedings of 11th IEEE East-West Design & Test Symposium (EWDTS`2013)*. 27th –30th September 2013. Rostov-on-Don, Russia. pp. 208-211.
17. Berger, J.M. (1961) A Note on Error Detecting Codes for Asymmetric Channels. *Information and Control*. 4(1). pp. 68-73. DOI: 10.1016/S0019-9958(61)80037-5
18. Das, D. & Touba, N.A. (1999) Weight-Based Codes and Their Application to Concurrent Error Detection of Multilevel Circuits. *Proc. 17th IEEE Test Symposium*. USA, California. pp. 370-376.
19. Das, D., Touba, N.A., Seuring, M. & Gossel, M. (2000) Low Cost Concurrent Error Detection Based on Modulo Weight-Based Codes. *Proceedings of IEEE 6th International On-Line Testing Workshop (IOLTW)*. 3rd to 5th July 2000. Spain, Palma de Mallorca. pp. 171-176.
20. Mehov, V., Saposhnikov, V., Sapozhnikov, Vl. & Urganskov, D. (2007) Concurrent Error Detection Based on New Code with Modulo Weighted Transitions between Information Bits. *Proceedings of 7th IEEE East-West Design & Test Workshop (EWDTW`2007)*. 25th to 30th September 2007. Erevan, Armenia. pp. 21-26.
21. Mekhov, V.B., Sapozhnikov, V.V. & Sapozhnikov, Vl.V. (2008) Checking of Combinational Circuits Basing on Modification Sum Codes. *Automation and Remote Control*. 8. pp. 153-165. DOI: 10.1134/S0005117908080134
22. Blyudov, A.A., Sapozhnikov, V.V. & Sapozhnikov, Vl.V. (2012) A Modified Summation Code for Organizing Control of Combinatorial Circuits. *Automation and Remote Control*. 73(1). pp. 153-160. DOI: 10.1134/S0005117912010122
23. Blyudov, A.A., Efanov, D.V., Sapozhnikov, V.V. & Sapozhnikov, Vl.V. (2013) Summation Codes for Organization of Control of Combinational Circuits. *Automation and Remote Control*. 74(6). pp. 1020-1028. DOI: 10.1134/S0005117913060118.
24. Blyudov, A.A., Efanov, D.V., Sapozhnikov, V.V. & Sapozhnikov, Vl.V. (2014) On codes with summation of data bits in concurrent error detection systems. *Automation and Remote Control*. 75(8). pp. 1460-1470. DOI: 10.1134/S0005117914080098
25. Sapozhnikov, V.V., Sapozhnikov, Vl.V. & Efanov, D.V. (2015) Errors classification in information vectors of systematic codes. *Izvestiya vuzov. Priborostroenie – Journal of Instrument engineering*. 58(5). pp. 333-343. DOI: 10.17586/0021-3454-2015-58-5-333-343. (In Russian).
26. Efanov, D.V., Sapozhnikov, V.V. & Sapozhnikov, Vl.V. (2010) On Summation Code Properties in Functional Control Circuits. *Automation and Remote Control*. 71 (6). pp. 155–162. DOI: 10.1134/S0005117910060123
27. Cherkasova, T.Kh. (2015) [Error Detection in Automation and Computer Equipment Using Codes Berger and Its Modifications]. *Problemy bezopasnosti i nadezhnosti mikroprotsessornykh kompleksov* [Problems of safety and reliability of microprocessor systems]. Proc. of Scientific-Practical Conference. St. Petersburg: PGUPS.

28. Busaba, F.Y. & Lala, P.K. (1994) Self-Checking Combinational Circuit Design for Single and Unidirectional Multibit Errors. *Jounal of Electronic Testing: Theory and Applications*. 5. pp. 19-28. DOI: 10.1007/BF00971960
29. Morosow, A., Sapozhnikov, V.V., Sapozhnikov, Vl.V. & Goessel, M. (1998) Self-Checking Combinational Circuits with Unidirectionally Independent Outputs. *VLSI Design*. 5(4). pp. 333-345.
30. Blyudov, A.A., Efanov, D.V., Sapozhnikov, V.V. & Sapozhnikov, Vl.V. (2012) Formation of the Berger Modified Code with Minimum Number of Undetectable Errors of Data Bits. *Elektronnoe modelirovanie – Electronic Modeling*. 34(6). pp. 17-30. (In Russian).
31. Sapozhnikov, V.V., Sapozhnikov, Vl.V. & Efanov, D.V. (2014) On the Use of the Properties of Sum Code for Unidirectional Error Detection in Concurrent Error Detection (CED) Systems of Combinational Circuits. *Vestnik Tomskogo gosudarstvennogo universiteta. Upravlenie, vychislitel'naya tekhnika i informatika – Tomsk State University Journal of Control and Computer Science*. 3. pp. 76-88. (In Russian).
32. Sapozhnikov, V.V., Sapozhnikov, Vl.V. & Efanov, D.V. (2015) Primenenie kodov s summirovaniem pri sinteze sistem zheleznodorozhnoy avtomatiki i telemekhaniki na programmiruemykh logicheskikh integral'nykh skhemakh [Application of Codes with Summation for the Synthesis of Railway Automation and Remote Control Systems Based on Field-Programmable Gate Arrays]. *Avtomatika na transporte*. 1(1). pp. 84-107.
33. Sapozhnikov, V., Sapozhnikov, Vl., Efanov, D. & Nikitin, D. (2014) Combinational Circuits Checking on the Base of Sum Codes with One Weighted Data Bit. *Proceedings of 12th IEEE East-West Design & Test Symposium (EWDTS`2014)*. 26th to 29th September 2014. Kyev, Ukraine. pp. 126-136.
34. Sapozhnikov, V.V., Sapozhnikov, Vl.V. & Efanov, D.V. (2014) Postroenie kodov s summirovaniem s naimen'shim kolichestvom neobnaruzhivaemykh simmetrichnykh oshibok v informatsionnykh vektorakh [Formation of Codes with Summation with the Smallest Number of Undetectable Errors of Data Bits]. *Radioelektronika i informatika – RadioElectronics & Informatics Journal*. 4. pp. 46-55.
35. Mel'nikov, A.G., Sapozhnikov, V.V., Sapozhnikov, Vl.V. (1986) Sintez samoproveryayushchikhsya testerov dlya kodov s summirovaniem [Synthesis of Self-Checking Checkers for Berger Codes]. *Problemy peredachi informatsii – Problem of Transmission*. 22(2). pp. 85-97.
36. Sapozhnikov, V.V., Sapozhnikov, Vl.V., Efanov, D.V. & Nikitin, D.A. (2013) Method of Constructing the Berger Code with High Error Detecting Efficiency in Information Bits. *Electronic Modeling*. 35(4). pp. 21-34. (In Russian).
37. Sapozhnikov, V.V., Sapozhnikov, Vl.V., Efanov, D.V., Nikitin, D.A. (2015 Research of Properties of Codes With Summation With One Weighted Data Bit in Concurrent Error Detection Systems. *Electronic Modeling*. 37(1). pp. 25-48. (In Russian).
38. *Collection of Digital Design Benchmarks*. [Online] Available from: http://ddd.fit.cvut.cz/prj/Benchmarks.
39. Sentovich, E.M., Singh, K.J., Lavagno, L., Moon, C., Murgai, R., Saldanha, A., Savoj, H., Stephan, P.R., Brayton, R.K. & Sangiovanni-Vincentelli, A. (1992) *SIS: A System for Sequential Circuit Synthesis*. Electronics Research Laboratory, Department of Electrical Engineering and Computer Science. University of California, Berkeley.
40. Sapozhnikov, V.V., Sapozhnikov, Vl.V. & Efanov, D.V. (2015) Obnaruzhenie opasnykh oshibok na rabochikh vykhodakh kombinatsionnykh logicheskikh skhem [Dangerous Errors Detection at the Operational Outputs of Combinational Logic Circuits]. *Avtomatika na transporte*. 1(2). pp. 195-211.

ВЕСТНИК ТОМСКОГО ГОСУДАРСТВЕННОГО УНИВЕРСИТЕТА

2015 Управление, вычислительная техника и информатика № 4 (33)

УДК 681.5.09
DOI: 10.17223/19988605/33/10

A.Yu. Matrosova, S.A. Ostanin, E.A. Nikolaeva, I.E. Kirienko

FULLY DELAY AND MULTIPLE STUCK-AT FAULT TESTABLE SEQUENTIAL CIRCUIT DESIGN

Supported by Russian Science Foundation (project № 14-19-00218).

In this paper we show that it is possible to derive a sequential circuit either from a transition table or from a State Transition Graph (STG), so that the sequential circuit has the short test detecting all multiple stuck-at faults at gate poles of the circuit, and delay of each circuit path is detectable

Keywords: finite state machine (FSM); state transition graph (STG); path delay faults (PDFs); multiple stuck-at faults; monotonous functions.

Under development of nanometer technologies it is not enough to test single stuck-at faults at gate poles of logical circuits. We need to test multiple stuck-at faults at gate poles along with delay faults of circuits. A path delay fault (PDF) model is often used at delay testing. In accordance with the conditions of fault manifestation, PDFs are divided into robust testable faults and non-robust testable faults [1, 2]. The PDF is robust testable on condition that the fault manifestation on a test pair does not depend on delays of other circuit paths. The PDF is non-robust testable if the fault manifestation is possible only when all other paths of a circuit are fault-free. Note that the robust testable PDF may be located and measured.

It is known that even for a simple circuit it is impossible to enumerate its multiple stuck-at faults, and only 20% of PDFs are robust testable. That is why it is very important to provide testability of logical circuits both for multiple stuck-at faults and delays during the circuit design.

Application of these synthesis methods to a combinational part of a sequential circuit may generate rather long paths in the obtained circuit. We suggest [4] a new method of the fully delay testable combinational part design that is oriented to cutting down the lengths of circuit paths in contrast to the lengths in the circuits obtained by the method suggested in [3]. In this method [4] the circuit behavior is represented with a composition of ROBDDs and monotonous products extracted from a state transition graph (STG). The experimental results have shown that the circuit paths in the resulted combinational parts [4] are cut from two to twenty times. Unfortunately, for these circuits we cannot find rather a short test detecting all multiple stuck-at faults at their gate poles.

For a high performance circuit it is desirable to detect multiple stuck-at faults together with PDFs. In paper [5] it is shown that if we apply the multilevel logic minimization method, based on algebraic division, to the system of irredundant sum of products (SoPs) describing the circuit behavior, we get a circuit testable for multiple stuck-at faults. We mean that a test detecting single faults of the literals of the system of irredundant SoPs detects all multiple stuck-at faults at gate poles of the resulted circuit.

In this paper we suggest the design method that allows deriving the sequential circuit that is both fully delay testable and multiple stuck-at fault testable.

The circuit behavior is represented either with a transition table or with the STG in which symbols of input and output alphabets have already encoded. For the STG we need to use additional inputs to provide above mentioned testability properties.

In Section 1 the approach to a sequential circuit design is suggested. In Section 2 the testability properties of the resulted circuit are investigated, and some means of cutting a test length are noted.

1. Sequential circuit design

Let the discrete device behavior be represented either with the transition table or the STG. We suggest the method of deriving the sequential circuit that implements the behavior and has the above mentioned testability properties. An example of the STG description is given in table 1.

Note that for encoding states we may use any unordered code. A code is unordered if any two code words are not comparable. Two code words α, β are not comparable if $\alpha \not\leq \beta$ and $\beta \not\leq \alpha$. For example, the code words 01010, 101111 are not comparable. We suggest encoding symbols of a state alphabet with unordered code words, in partly, code words of (m, n) code. Here n is a length of a code word and m is a number of its 1value components. To cut the length of code words we may choose value m being approximately equal to $\lceil 1/2 \rceil$. An example of the STG description after encoding states is given in table 2. Here we use (1, n) code).

Table 1

An example of the STG description

$x_1\, x_2\, x_3$	S	S	$y_1\, y_2\, y_3\, y_4\, y_5$
0 - -	1	1	0 0 0 1 0
- 0 -	1	1	0 0 0 1 0
1 1 -	1	2	1 0 0 1 0
- - 0	2	2	0 0 1 1 0
- - 1	2	3	1 0 1 1 0
1 0 -	3	3	0 1 0 0 0
0 - -	3	4	1 1 0 0 0
- 1 -	3	4	1 1 0 0 0
- - 0	4	4	0 1 0 0 1
- - 1	4	1	1 1 0 0 1

Table 2

The STG description after encoding states

$x_1\, x_2\, x_3$	$z_1\, z_2\, z_3\, z_4$	z_1', z_2', z_3', z_4'	$y_1\, y_2\, y_3\, y_4\, y_5$
0 - -	1 0 0 0	1 0 0 0	0 0 0 1 0
- 0 -	1 0 0 0	1 0 0 0	0 0 0 1 0
1 1 -	1 0 0 0	0 1 0 0	1 0 0 1 0
- - 0	0 1 0 0	0 1 0 0	0 0 1 1 0
- - 1	0 1 0 0	0 0 1 0	1 0 1 1 0
1 0 -	0 0 1 0	0 0 1 0	0 1 0 0 0
0 - -	0 0 1 0	0 0 0 1	1 1 0 0 0
- 1 -	0 0 1 0	0 0 0 1	1 1 0 0 0
- - 0	0 0 0 1	0 0 0 1	0 1 0 0 1
- - 1	0 0 0 1	1 0 0 0	1 1 0 0 1

We consider each line of table 2 as pair (u_i, h_i) of vectors. The first ternary vector (cube) u_i of the pair is represented with two first columns of Table 2, and has the length $n + p$ (in the considered example the length is 3 + 4 = 7). Here n is the number of circuit's inputs and p is the number of state variables (loops). The second Boolean vector h_i represented with two last columns of table 2 is a characteristic of the first vector. In this characteristic 1 values point to the functions taking 1 value on the first vector (cube), but 0 values point to the functions taking 0 value on the first vector (cube). Note that table 2 represents the system of incompletely specified Boolean functions that are divided into transition functions and output functions. Take into consideration that in Table 2 each incompletely specified function f_i of the system is represented with the cubes of on-set $M^1_{f_i}$ and the cubes of off-set $M^0_{f_i}$.

Give some definitions.

Call (w_j, h_j) as a system cube of an incompletely specified Boolean functions system, where characteristic h_j means that its 1 value component points to the function that take 1 value on the cube w_j; here we don't pay attention on 0 values of this characteristic.

System cube (w_j, h_j) is essential if $w_j \cap M^1_{f_i} \neq \emptyset$ and $w_j \cap M^0_{f_i} \neq \emptyset$ for each function f_i taking 1 value in characteristic h_j.

A system cube (w_j, h_j) covers a system cube (u_i, h_i) if $h_j = h_i$, and w_j covers u_i.

We differ cubes represented by ternary vectors without characteristics from system cubes supplied with characteristics.

Note that two cubes are orthogonal each other if there is at least one component that has 1 value in one cube and 0 value in another cube. For example, the cubes 010--01, 010--11 are orthogonal by the sixth component.

Two cubes are opposite two orthogonal each other if there are two components: in one of these components the cubes have 10 values, but in another component they have 01 values. For example, the cubes 0--101, 1-0001 are opposite two orthogonal by the first and the forth components.

In a set of unordered code words any two words are opposite two orthogonal.

Note that the cubes in table 2 corresponding to the different states are opposite two orthogonal; the cubes corresponding to the same state and having the different characteristics are surely orthogonal, but may be not opposite two orthogonal. To provide monotonous implementation of the system of incompletely specified Boolean functions represented by table 2, we need to support opposite two orthogonality of cubes correspond-

ing to the same state and having different characteristics. It demands adding inputs for the sequential circuit. It is desirable to cut the number of additional inputs as much as possible.

1.1. Injecting additional inputs

We suggest the algorithm of finding the minimal number of additional input variables to provide monotonous implementation of the system of incompletely specified Boolean functions, represented by the STG in which the states are encoded with unordered code words (here (m, n) code words).

For that we have to derive a direct graph G from a section of table 2. We include all system cubes corresponding to the same state in the section. The graph vertexes correspond to the cubes of the section of table 2, represented by the first column. Two vertexes are connected with edge if the corresponding cubes u_i, u_j are orthogonal but not opposite two orthogonal. The edge is directed from the vertex having 0-value in the cube u_i component to the vertex having 1 value in the same component of the cube u_j. If there is a cube, that is opposite two orthogonal any cube of table 2, then its corresponding vertex is absent in the graph G.

Two edges are compatible if the end of one edge is not the beginning of another one. A set of edges is compatible if any pair of its edges is compatible. A set of edges is maximal compatible if it is not contained in any compatible set of edges.

Algorithm of finding the minimal number of additional inputs.

1. Find for the graph G all compatible sets deriving sets of capacities 1, 2,…, q.

2. Form maximal compatible sets: S_1, S_2,…, S_t. Note that all sets of the capacity q are among maximal compatible ones.

3. Derive the Boolean matrix. Its lines correspond to the maximal compatible sets and its columns correspond to edges of the graph. Element ij of this matrix is equal to 1 if the maximal set corresponding to the line i contains the edge corresponding to the column j otherwise the element is equal to 0.

4. Find the shortest coverage of the columns by the lines.

The capacity of this *coverage is the minimal number of additional input variables for cubes of the section.*

Note that the intersections of some sets of the coverage may be not empty.

We execute injecting the input variables to the cubes of the section considered in the following way.

5. Let us consider set S_j from the shortest coverage S_1,…, S_t taking into account only the edges that are absent in the sets S_1,…, S_{j-1}. Inject the additional variable x_j and correlate 1 value of this variable to the cubes of vertexes, the edges of which belonging to S_j run from the vertexes and 0 values to the cubes of vertexes the edges of which belonging to S_j come to the vertexes. The rest cubes of the section get don't care value of the variable x_i.

6. Execute step 5 till all sets from the shortest coverage will be examined.

7. Fulfill steps 1-6 for each sections of table 2. Therefore all cubes of the same sections with different characteristics will be opposite two orthogonal.

The shortest coverage of maximal length among all sections is the minimal number of additional inputs of the sequential circuit that we derive.

Illustrate the above-mentioned algorithm for table 2. Here we have very simple graphs.

For the cubes corresponding to the state 1000 we have the graph G_{1000} (Fig. 1b) and the vectors (fig. 1, *a*). The edges (1, 3) and (2, 3) form one compatible set: it is needed one additional input variable x_4 (fig. 1, *c*). For the cubes corresponding to the state 0100 we have the graph G_{0100} (fig. 2, *b*) and the vectors (fig. 2, *a*). The edge (1, 2) forms one compatible set: it is needed one additional input variable x_4 (fig. 2, *c*).

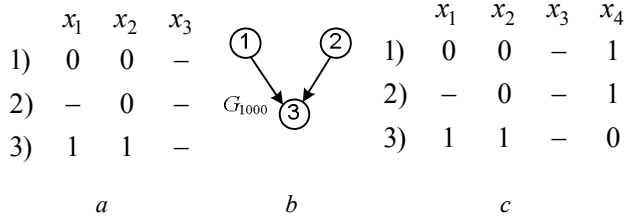


Fig. 1. Example for state 1000 in table 2

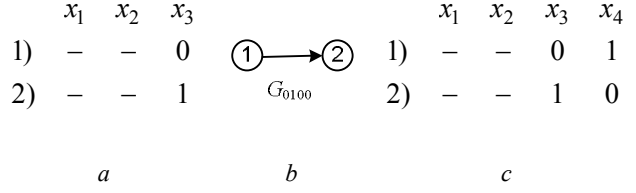


Fig. 2. Example for state 0100 in table 2

For the cubes corresponding to the state 0010 we have the graph G_{0010} (fig. 3, *b*) and the vectors (fig. 3, *a*). The edges form two compatible sets {(2, 1)} and {(1, 3)}: it is needed two additional input variables x_4 and x_5 (fig. 3, *c*). For the cubes corresponding to the state 0001 we have the graph G_{0001} (fig. 4, *b*) and the vectors (fig. 4, *a*). The edge (1, 2) forms one compatible set: it is needed one additional input variable x_4 (fig. 4, *c*).

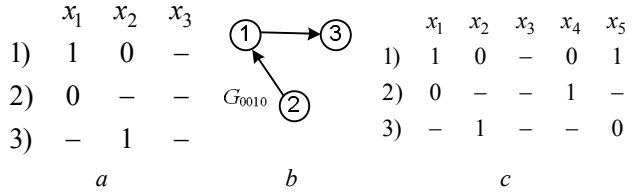


Fig. 3. Example for state 0010 in table 2

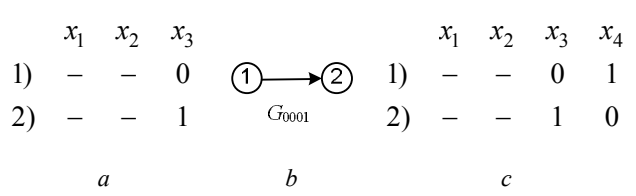


Fig. 4. Example for state 0001in table 2

The system of incompletely specified Boolean functions with additional input variables is represented in table 3. Change 0-values in the first and second columns of Table 3 for don't care executing merged system cubes and get table 4. It is known that table 4 is the implementation of the system of the incompletely specified Boolean functions represented with table 3.

Table 3

The STG description with additional input variables

$x_1\,x_2\,x_3\,x_4\,x_5$	$z_1\,z_2\,z_3\,z_4$	$z_1'\,z_2'\,z_3'\,z_4'$	$y_1\,y_2\,y_3\,y_4\,y_5$
0 - - 1 -	1 0 0 0	1 0 0 0	0 0 0 1 0
- 0 - 1 -	1 0 0 0	1 0 0 0	0 0 0 1 0
1 1 - 0 -	1 0 0 0	0 1 0 0	1 0 0 1 0
- - 0 1 -	0 1 0 0	0 1 0 0	0 0 1 1 0
- - 1 0 -	0 1 0 0	0 0 1 0	1 0 1 1 0
1 0 - 0 1	0 0 1 0	0 0 1 0	0 1 0 0 0
0 - - 1 -	0 0 1 0	0 0 0 1	1 1 0 0 0
- 1 - - 0	0 0 1 0	0 0 0 1	1 1 0 0 0
- - 0 1 -	0 0 0 1	0 0 0 1	0 1 0 0 1
- - 1 0 -	0 0 0 1	1 0 0 0	1 1 0 0 1

Table 4

The STG description with additional variables

$x_1\,x_2\,x_3\,x_4\,x_5$	$z_1\,z_2\,z_3\,z_4$	$z_1'\,z_2'\,z_3'\,z_4'$	$y_1\,y_2\,y_3\,y_4\,y_5$
- - - 1 -	1 - - -	1 0 0 0	0 0 0 1 0
1 1 - - -	1 - - -	0 1 0 0	1 0 0 1 0
- - - 1 -	- 1 - -	0 1 0 0	0 0 1 1 0
- - 1 - -	- 1 - -	0 0 1 0	1 0 1 1 0
1 - - - 1	- - 1 -	0 0 1 0	0 1 0 0 0
- - - 1 -	- - 1 -	0 0 0 1	1 1 0 0 0
- 1 - - -	- - 1 -	0 0 0 1	1 1 0 0 0
- - - 1 -	- - - 1	0 0 0 1	0 1 0 0 1
- - 1 - -	- - - 1	1 0 0 0	1 1 0 0 1

Thus we get the implementation of system $F(x_1,\ldots, x_{n+p})$ ($n(p)$ – the number of input (state) variables) of incompletely specified Boolean functions by a set of the essential cubes of this system. These cubes are represented by the lines of table 4. The completely specified Boolean functions of Table 4 are monotonous.

Table 5

The experimental results

Circuits	***n***	***m***	***s***	***p***	S_1	Σ_1	s_2	δ_1	δ_2	Σ_2
BBTAS	2	2	6	24	3	120	4	2	4	90
BEECOUNT	3	4	7	51	3	152	5	3	3	106
DK27	1	2	7	14	3	56	5	1	2	42
DK512	1	3	15	30	4	150	6	1	2	90
LION	2	1	4	15	2	40	4	2	3	28
LION9	2	1	9	25	4	150	5	2	3	89
MODULO12	1	1	12	24	4	120	6	1	2	72
S8	4	1	5	20	3	140	4	0	4	60
TRAIN11	2	1	11	25	4	150	6	2	3	83

The table 5 represents the experimental results oriented to finding the minimal number of additional inputs for benchmark circuits and simplification of circuit implementation at the expense of using monotonous functions. Here n is the number of primary inputs, m – primary outputs, s – number of states, p – number of products, s_1 and s_2 – number of bits for encoding internal states by the shortest Boolean vectors (s_1) and (m, n)-code words(s_2), δ_1 and δ_2 – number of additional inputs estimated by suggested algorithm (δ_1) and by adding

(m, n)-code words (depending on input variables) to provide opposite two orthogonality for the proper input cubes of the same section (δ_2), Σ_1 – sum of cube (product) ranks of implementation of STG description (system of completely specified Boolean function) after encoding states by the shortest Boolean vectors and without additional inputs, Σ_2 – sum of cube ranks of implementation of STG description suggested above in this paper.

We may use simplification method for this system of monotonous functions proposed by us in [6]. A simplification procedure is based on extension of the essential cubes of system $F(x_1,\ldots, x_{n+p})$ with maximal characteristics. Note that as a result of implementing system $F(x_1,\ldots, x_{n+p})$ of incompletely specified Boolean functions we get system of monotonous functions F^*.

1.2. Deriving circuit C

After getting the monotonous system F^* we suggest using the multilevel logic minimization method based on algebraic division followed covering obtained sum of products (SoPs) with gates [7]. In this case, we must previously transform each SoP of F^* into algebraic SoP in which a mergence of products is executed. Thus we get system F^* of monotonous SoPs. Applying the method [7] to this system we get circuit C that keeps system F^*. This means that when moving from the certain i-th circuit output to its inputs substituting the proper gate functions instead of internal circuit variables and eliminating brackets we obtain the corresponding monotonous SoP f_i^* from system F^*. Note that during substituting functions and eliminating brackets any simplifications are forbidden.

We may also use the special two-level synthesis method, suggested by us before [8]: it also keeps the system F^*.

Two-level synthesis methods are based on choosing a set of product factors and SoPs factors. A product factor of system F^* is derived from the certain product of system F^* by the elimination of some literals from the product. After getting a set of product factors, each system F^* product is covered by the certain factors from this set.

They say that the certain factors cover product K if multiplication of these factors gives rise to all literals in product K. If at least one literal appears two or more times under multiplication then such covering is called redundant otherwise the covering is called irredundant.

A SoP factor of system F^* is derived from the certain SoP of system F^* by elimination of some products from the system SoP. After getting a set of SoPs factors, each SoP of system F^* is covered by the certain SoP factors from this set.

They say that the certain factors cover the SoP f^* of system F^* if a disjunction of these factors gives rise to all products that are present in the SoP f^*. If the product appears two or more times under disjunction then such covering is called redundant otherwise the covering is called irredundant.

Note that a set of product factors and a set of SoP factors are chosen in such a way that each product and each SoP of system $\boldsymbol{F^*}$ has to be covered. As a result of covering, we get combinational circuit C.

We restrict our consideration of two-level synthesis methods in which only irredundant covering products and SoPs of system $\boldsymbol{F^*}$ are acceptable. Call such methods as two-level synthesis methods based on irredundant covering. These methods differ from each other by choosing factors and ways of covering,

Thus using either the multilevel logic minimization method or the two level synthesis method based on irredundant covering we get circuit C keeping system $\boldsymbol{F^*}$ of monotonous functions.

2. Testability properties of circuit *C*

2.1. Testability for multiple stuck-at faults at gate poles of circuit C

Consider single faults of literals of monotonous SoP f_i^* from F^*. Changing the certain literal in the certain product of a SoP for the constant 1 (0) call $b(a)$-fault. If several literals are fault, call it multiple fault of SoP. It is known [9] that all multiple faults of irredundant SoP consisting of prime implicants are detected with a test for single faults of literals of this SoP. This test is union of test patterns for each a-fault and test patterns

for each *b*-fault. Any f_i^* from *F** is irredundant SoP as it consists of monotonous products, and there are no product mergences.

It is also known that each *a*(*b*)-fault in irredundant SoP is detectable [10]. Test pattern α of the monotonous SoP for *a*-fault of product *K* is obtained by changing all don't care components of ternary vector corresponding to *K* for 0 values. Test pattern β for *b*-fault of product *K* and its literal x_i is obtained by changing the component corresponding to the fault literal x_i for 0 value in vector α (in test pattern for *a*-fault).

For example, we have monotonous SoP for state variable z_4' : $x_4z_3 \vee x_2z_3 \vee x_4z_4$.

Consider product $K = x_4z_3$ and literal x_4, then test pattern α for *a*-fault is Boolean vector 00010010.

Test pattern β for *b*-fault is Boolean vector 00000010.

Note that any multiple stuck-at fault at gate poles of the sub-circuit implementing monotonous SoP f_i^* is equivalent to the multiple fault of this SoP literals [5]. Consequently test for single faults of monotonous SoP f_i^* detects any multiple stuck-at fault at gate poles of the corresponding subcircuit from *C*. The length of this test is not more than sum of ranks of monotonous SoP f_i^* products and the number of this SoP products. Thus in order to derive test detecting any multiple stuck-at faults at gate poles of circuit *C* we may merge tests for single faults of literals of all monotonous SoPs of system *F**. Try to cut a length of a test detecting any multiple stuck-at fault at gate poles of circuit *C*.

For that we form SoP *f* including into it all products corresponding to the ternary vectors of Table 4 represented by the first and the second columns. Consider certain product *K* from SoP *f* and certain literal x_i. Find test pattern for *a*(*b*)-fault in above mentioned way. Let product *K* coincides with product *K** of SoP f_i^*, and literal x_i coincides with literal x_i^* from *K**.

Theorem 1. A test pattern for *a*(*b*)-fault of literal x_i from *K* is the test pattern for *a*(*b*)-fault for literal x_i^* of product *K**.

Proof. It follows from the condition: a set of the products of SoP f_i^* is subset of the products of SoP *f*. The theorem is proved.

Corollary 1. Test *T* for single *a*(*b*)-faults of SoP *f* is the test for the same faults of system *F** and, consequently, the test for multiple stuck-at faults at gate poles of circuit *C*.

Corollary 2. The length of test *T* is not more than the sum of ranks of SoP *f* products and the number of these products.

For the system of Table 4 the length of test *T* is not more than 29.

Note that the *T* length may be essentially cut as test patterns for *b*-faults of different products of SoP *f* may coincide [11].

2.2. Full path delay testability of circuit C

We say that circuit is fully delay testable if the delay of each path may be located and measured.

Consider monotonous SoP f_i^* of system *F** corresponding to the one output sub-circuit of *C*. The sub-circuit path α beginning on circuit input x_i and ending on its output originates literal x_{i_α} in the corresponding equivalent normal form (ENF) of the sub-circuit. If we exclude from ENF the sequences of numbers of gates representing the sub-circuit paths we get monotonous SoP f_i^* of system *F**. Consider one product *K** from SoP *f** and literal x_i corresponding to literal x_{i_α} of the ENF. In [12] it is suggested to consider delay fault of path α as a temporary fault of the corresponding literal x_{i_α} of the ENF, and the conditions of detecting robust PDF are formulated with using properties of ENF products. When finding test pair for falling (rising) transition of the path, it is suggested [12] to use test pattern for the corresponding $b_p(a_p)$-fault. This test pattern is vector v_2 of the test pair.

Note that a_p fault means that each appearance of literal x_{i_a} in ENF products is changed for constant 0 and b_p fault means that each appearance of literal x_{i_a} in ENF products is changed for constant 1.

Take into consideration that literal x_i^* corresponding to literal x_{i_a} in the ENF may appear in several products of monotonous SoP f_i^*.

As SoP f_i^* is monotonous and its products have no repeated variables, then for any literal of any product there exists the test pattern both for changing the literal by constant 0 and constant 1.

Consider product ***K**** from f_i^* and its literal x_i^*. Derive vector γ from the ternary vector representing product ***K**** by changing all don't care components for 0.

Theorem 2. Vector γ is the test pattern for the a_p-fault.

Proof. Vector γ turns all products of SoP f_i^* into 0 except product *K** that is it turns fault free SoP into 1. When a_p-fault appears, fault SoP takes 0 value on vector γ. The theorem is proved.

Derive vector δ from vector γ changing its 1 value component x_i^* for 0 value component.

Theorem 3. Vector δ is the test pattern for the b_p-fault.

Proof. Vector δ turns all products of fault free SoP f_i^* into 0. As literal x_i^* disappears in product *K** of fault SoP f_i^*, then the fault SoP takes 1 value on vector δ. The theorem is proved.

The test pattern for the a_p-fault is vector v_2 of the test pair for rising transition of path α. The test pattern for the b_p-fault is vector v_2 of the test pair for the falling transition of path α.

Let u be minimal cube covering vectors γ, δ.

Theorem 4. (δ γ) is the test pair for rising transition of the robust PDF.

Proof. According to Theorem 2, vector γ is vector v_2 for a_p fault. Vector v_1 must have opposite value [12] on variable x_i. This condition is executed. Cube u must be orthogonal [12] all products of SoP f_i^* except ones that contain literal x_i^* corresponding to path α. In this case, u is orthogonal all products of SoP f_i^* except product *K**. Note that the products of SoP f_i^* have no repeated variables. This means that all conditions for robust testable manifestation for rising transition of path α is executed. The theorem is proved.

Theorem 5. (γ δ) is the test pair for falling transition of the robust PDF.

The proof is similar to that of Theorem 4.

Corollary. The sequence (δ γ δ)((γ δ γ)) detects the robust testable PDFs for rising and falling transition of path α.

Thus we Have shown that each PDF of any sub-circuit of circuit *C* is robust testable. Note that test pairs consist of test patterns for $a_p(b_p)$-faults of SoP f_i^*.

Take into consideration that The fully delay and multiple stuck-at fault testability is provided, firstly, by using a description of the sequential circuit behavior with the system of monotonous functions. Secondly, we apply either the multilevel logic minimization method based on algebraic division or the two level synthesis method based on irredundant covering. Note that the conventional two-level synthesis method based on choosing product and SoP factors does not guarantee the fully delay testability property because of masking $a_p(b_p)$-faults.

Using a system of monotonous functions we increase the number of loops and the number of inputs of the sequential circuit but decrease ranks of SoP products.

Conclusion

We have shown that it is possible to derive the sequential circuit from a description of the FSM behavior either by the transition table or by the STG, so that the resulted circuit has short test detecting all multiple stuck-at faults at gate poles of the circuit, and a delay of each circuit path is detectable, i.e. robust testable. For that we need to use unordered code words for symbols of the state alphabet for the STG and symbols of input

and state alphabets for the transition table. It is necessary to add additional inputs when using the STG. As a result we get system of monotonous functions. Then applying either the multilevel logic minimization method based on algebraic division or the two-level synthesis method based on irredundant covering to the system, we derive combinational part *C,* fully testable for multiple stuck-at faults and path delay faults. The method of adding the minimal number of input variables for the STG description is suggested. This approach to the FSM design is connected with increasing of the numbers of inputs and loops of the resulted sequential circuit and decreasing its structural implementation.

REFERENCES

1. *Lin C.J., Reddy S.M.* On Delay fault testing in logic circuits // IEEE Trans. on Computer-Aided Design. 1987. V. 6, No. 5. P. 694–701.
2. *Bushnell M.L., Agrawal V.D.* Essentials of electronictesting for digital, memory and mixed-signal VLSI Circuits. Kluwer Academic Publishers. Boston, Mass, USA, 2000. P. 671.
3. *Matrosova A., Nikolaeva E., Kudin D., Singh V.* PDF testability of the circuits derived by special covering ROBDDs with gates // Proc. of the IEEE East-West Design&Test Symposium. Rostov-on-Don, 2013. P.1–5.
4. *Matrosova A.Yu., Mitrofanov E.V., Singh V.* Delay Testable Sequential Circuit Designs // Proc. of the IEEE East-West Design & Test Symposium (EWDTS 2013). Rostov-on-Don, 2013. P. 293–296.
5. *Matrosova A., Andreeva V., Ostanin S.* Easy testable combinational circuit design // Proc. Of the 6-th International workshop on Boolean problems. Germany. Freiberg, 2004. P. 237–244.
6. *Matrosova A.Yu., Andreeva V.V.* Minimizing the Boolean function system oriented toward self-checking finite-state machine design // J. of Optoelectronics, Instrumentation and Data Processing. 2008. V. 44, Issue 5. P. 459–467.
7. *Murgai R., Brayton R., Sangiovanni-Vincentelli A.* Logic Synthesis for Field Programmable Gate Arrays // Kluwer Academic Publisher. 1995. P. 425.
8. *Matrosova A., Kudin D., Nikolaeva E.* Combinational Circuits without False Paths // Proc. of the IEEE East-West Design&Test Symposium. Kiev. Ukraine, 2014. P. 1–6.
9. *Kohavi I., Kohavi Z.* Detection of multiple faults in combinational logic networks // IEEE Trans. Comput. 1975. VC-20, No. 6. P. 566–568.
10. *Matrosova A.Yu.* A fault-detection method for a synchronous device // Automation and Remote Control. 1978. V. 38, No. 12. P. 1849–1857.
11. *Andreeva V.V., Matrosova A.Yu.* Postroenie minimizirovannogo proveryaushego testa, obnaruzhivayushego neispravnost bezizbitochnoy DNA // Tomsk State University Journal. Supplement. 2006. No. 18. P. 34–39. (In russian)
12. *Matrosova A., Lipsky V., Melnikov A., Singh V.* Path delay faults and ENF // Proc. of the IEEE East-West Design&Test Symposium. St. Perersburg. Russia. 2010. P. 164–1677.

Matrosova Anjela Yurievna, doctor of science, professor. E-mail: mau11@yandex.ru
Ostanin Sergey Alexandrovich, PhD, associate professor. E-mail: sergeiostanin@yandex.ru
Nikolaeva Ekaterina Alexandrovna, PhD. E-mail: nikolaeve-ea@yandex.ru
Kirienko Irina Evgenievna. E-mail: irina.kirienko@sibmail.com
Tomsk State University, Russian Federation.

Поступила в редакцию 25 августа 2015 г.

Матросова А.Ю., Останин С.А., Николаева Е.А., Кириенко И.Е. (Томский государственный университет, Российская Федерация).

Проектирование полностью тестопригодных последовательностных схем для неисправностей задержек путей и кратных константных неисправностей.

Ключевые слова: конечный автомат; граф переходов; неисправности задержек путей; кратные константные неисправности; монотонные функции.

DOI 10.17223/19988605/33/10

При развитии нанотехнологий в электронике недостаточно тестировать одиночные константные неисправности на полюсах элементов логических схем. Необходимо тестировать кратные константные неисправности на полюсах элементов наряду с неисправностями задержек схемы. Одной из часто используемых моделей для тестирования времени задержки является модель неисправностей задержек путей схемы (НЗП). В соответствии со способом проявления неисправностей НЗП разделены на робастно тестируемые неисправности и неробастно тестируемые неисправности. НЗП робастно тестируемы, если проявление неисправности на тестовой паре не зависит от задержек на других путях схемы. НЗП неробастно тестируемы, если проявление неисправности возможно, только когда все другие пути схемы исправны.

Известно, что даже для простых схем сложно перечислить все кратные константные неисправности, и только 20% НЗП являются робастно тестируемыми. Именно поэтому очень важно обеспечить тестопригодность логических схем и для кратных константных неисправностей, и для неисправностей задержек путей на этапе проектирования схемы.

В данной работе мы предлагаем метод синтеза, который позволяет получать последовательностную схему, которая является полностью тестопригодной и для НЗП, и для кратных константных неисправностей.

Было показано, что для заданного описания поведения автомата возможно получить последовательностную схему так, что результирующая схема имеет достаточно короткий тест, обнаруживающий все кратные константные неисправности на полюсах элементов схемы, и НЗП для каждого пути схемы обнаружимы, т.е. являются робастно тестируемыми.

С помощью специального кодирования (неупорядоченными кодовыми словами) внутренних состояний автомата и введением дополнительных входов построим систему монотонных функций, являющуюся одной из реализаций, описывающих поведение автомата. Показано, что применение метода минимизации многоуровневых логических схем, основанного на алгебраическом делении, или метода двухуровневого синтеза, основанного на безызбыточном покрытии системы, позволяет синтезировать комбинационную часть схемы, которая является тестопригодной для заданного класса неисправностей. В работе также предложен метод поиска минимального числа дополнительных входных переменных.

REFERENCES

1. Lin, C.J., Reddy, S.M. (1987) On Delay fault testing in logic circuits. *IEEE Trans. on Computer-Aided Design*. 6(5). pp. 694-701. DOI: 10.1109/TCAD.1987.1270315
2. Bushnell, M.L. & Agrawal, V.D. (2000) *Essentials of electronictesting for digital, memory and mixed-signal VLSI Circuits*. Boston: Kluwer Academic Publishers.
3. Matrosova, A., Nikolaeva, E., Kudin, D., Singh, V. (2013) PDF testability of the circuits derived by special covering ROBDDs with gates. *Proc. of the IEEE East-West Design&Test Symposium*. Rostov-on-Don. pp.1-5.
4. Matrosova, A.Yu., Mitrofanov, E.V. & Singh, V. (2013) Delay Testable Sequential Circuit Designs. *EWD&T'2013. Proc. of the IEEE East-West Design&Test Symposium.* pp. 293-296. DOI: 10.1109/EWDTS.2013.6673138
5. Matrosova, A., Andreeva, V. & Ostanin, S. (2004) Easy testable combinational circuit design. *IWBP'2004. Germany, Freiberg, sept. 2004. Proc. Of the 6-th International workshop on Boolean problems*. pp. 237-244.
6. Matrosova, A.Yu. & Andreeva, V.V. (2008) Minimizing the Boolean function system oriented toward self-checking finite-state machine design. *Journal of Optoelectronics, Instrumentation and Data Processing.* 44(5). pp. 459-467. DOI: 10.3103/S8756699008050117
7. Murgai, R., Brayton, R. & Sangiovanni-Vincentelli, A. (1995) *Logic Synthesis for Field Programmable Gate Arrays*. Boston: Kluwer Academic Publisher.
8. Matrosova, A., Kudin, D. & Nikolaeva, E. (2014) Combinational Circuits without False Paths. *EWD&T'2014. Proc. of the IEEE East-West Design&Test Symposium.* September 2014. Kiev, Ukraine. pp. 1-6.
9. Kohavi, I. & Kohavi, Z. (1975) Detection of multiple faults in combinational logic networks. *IEEE Trans. Comput*. 20(6). pp. 566-568. DOI: 10.1109/TC.1972.5009008
10. Matrosova, A.Yu. (1978) A fault-detection method for a synchronous device. *Automation and Remote Control*. 38(12). pp.1849-1857.
11. Andreeva V.V. & Matrosova A.Yu. (2006) Postroenie minimizirovannogo proveryaushego testa, obnaruzhivayushego neispravnost bezizbitochnoy DNA. *Vestnik Tomskogo gosudarstvennogo universiteta – Tomsk State University Journal*. Supplement. 18. pp. 34-39. (In Russian).
12. Matrosova, A., Lipsky, V., Melnikov, A. & Singh, V. (2010) Path delay faults and ENF. *Proc. of the IEEE East-West Design&Test Symposium*. St. Perersburg. pp. 164-1677.

СВЕДЕНИЯ ОБ АВТОРАХ

АНДРЕЕВА Валентина Валерьевна – кандидат технических наук, доцент кафедры программирования факультета прикладной математики и кибернетики Томского государственного университета. E-mail: valentina.andreeva@mail.tsu.ru

АШУРОВ Михаил Фаритович – аспирант кафедры прикладной информатики факультета информатики Томского государственного университета. E-mail: ashurov.mf@gmail.com

СОЛОВЬЕВ Александр Александрович – аспирант факультета прикладной математики и кибернетики Томского государственного университета. E-mail: sizal19@mail.ru

ГОРЦЕВ Александр Михайлович – профессор, доктор технических наук, декан факультета прикладной математики и кибернетики Томского государственного университета. E-mail: gam@mail.fpmk.tsu.ru

ЕРЁМИНА Александра Рафаэловна – кандидат физико-математических наук, доцент кафедры системного программирования и компьютерной безопасности факультета математики и информатики Гродненского государственного университета им. Янки Купалы (г. Гродно, Республика Беларусь). E-mail: a.eremina@grsu.by

ЕФАНОВ Дмитрий Викторович – кандидат технических наук, доцент кафедры «Автоматика и телемеханика на железных дорогах» Петербургского государственного университета путей сообщения Императора Александра I (г. Санкт-Петербург). E-mail: TrES-4b@yandex.ru

ИГНАТЬЕВ Николай Александрович – доктор физико-математических наук, профессор кафедры программных и сетевых технологий механико-математического факультета Национального университета Узбекистана (г. Ташкент, Республика Узбекистан). E-mail: ignatev@rambler.ru

МАТРОСОВА Анжела Юрьевна – профессор, доктор технических наук, заведующая кафедрой программирования факультета прикладной математики и кибернетики Томского государственного университета. E-mail: mau11@yandex.ru

КИРИЕНКО Ирина Евгеньевна – аспирантка кафедры программирования факультета прикладной математики и кибернетики Томского государственного университета. E-mail: irina.kirienko@sibmail.com

МАЛИНКОВСКИЙ Юрий Владимирович – профессор, доктор физико-математических наук, заведующий кафедрой экономической кибернетики и теории вероятностей Гродненского государственного университета им. Янки Купалы (г. Гродно, Республика Беларусь). E-mail: Malinkovsky@gsu.by

МЕДВЕДЕВ Геннадий Алексеевич – профессор, доктор физико-математических наук, профессор факультета прикладной математики и информатики Белорусского государственного университета (г. Минск, Республика Беларусь). E-mail: MedvedevGA@cosmostv.by

ОЛАДЬКО Владлена Сергеевна – кандидат технических наук, доцент кафедры информационной безопасности Волгоградского государственного университета. E-mail: oladko.vs@yandex.ru

ОСТАНИН Сергей Александрович – доцент, кандидат технических наук, доцент кафедры программирования факультета прикладной математики и кибернетики Томского государственного университета. E-mail: sergeiostanin@yandex.ru

НИКОЛАЕВА Екатерина Александровна – кандидат технических наук, доцент кафедры программирования факультета прикладной математики и кибернетики Томского государственного университета. E-mail: nikolaeveea@yandex.ru

ПОДДУБНЫЙ Василий Васильевич – профессор, доктор технических наук, профессор кафедры прикладной информатики факультета информатики Томского государственного университета. E-mail: vvpoddubny@gmail.com

СЕНЧЕНКО Павел Васильевич – кандидат технических наук, доцент кафедры автоматизации обработки информации, декан факультета систем управления Томского государственного университета систем управления и радиоэлектроники. E-mail: pvs@tusur.ru

СИДОРОВ Анатолий Анатольевич – кандидат экономических наук, доцент кафедры автоматизации обработки информации Томского государственного университета систем управления и радиоэлектроники. E-mail: saa@muma.tusur.ru

ТАРАСЕНКО Владимир Феликсович – доктор технических наук, профессор кафедры теоретической кибернетики факультета прикладной математики и кибернетики Томского государственного университета. E-mail: vtara54@mail.ru

ТАРНОВСКАЯ Татьяна Павловна – магистрант кафедры программирования факультета прикладной математики и кибернетики Томского государственного университета. E-mail: tarnovskayat@mail.ru

Научный журнал

ВЕСТНИК ТОМСКОГО ГОСУДАРСТВЕННОГО УНИВЕРСИТЕТА

УПРАВЛЕНИЕ, ВЫЧИСЛИТЕЛЬНАЯ ТЕХНИКА И ИНФОРМАТИКА

TOMSK STATE UNIVERSITY JOURNAL OF CONTROL AND COMPUTER SCIENCE

2015. № 4 (33)

Редактор Н.А. Афанасьева
Оригинал-макет А.И. Лелоюр
Редакторы-переводчики: Г.М. Кошкин; В.Н. Горенинцева
Дизайн обложки Л.Д. Кривцова

Подписано к печати 15.12.2015 г. Формат 60x84 1/8.
Гарнитура Times.
Печ. л. 11,5; усл. печ. л. 10,7.
Тираж 250 экз. Заказы № 1470; 1470/1.

Журнал отпечатан на полиграфическом оборудовании
Издательского Дома Томского государственного университета
634050, г. Томск, Ленина, 36
Тел. 8+(382-2)–53-15-28
Сайт: http://publish.tsu.ru; E-mail: rio.tsu@mail.ru