

№ 5 (91) сентябрь-октябрь 2015

Издается с 2002 года. Выходит 6 раз в год

Учредитель – федеральное государственное бюджетное образовательное учреждение
высшего профессионального образования

«Государственный университет —
учебно-научно-производственный комплекс» (Госуниверситет – УНПК)

Редакционный совет

Голенков В.А., председатель
Радченко С.Ю., заместитель председателя
Борзенков М.И., секретарь

Астафичев П.А., Иванова Т.Н., Киричек А.В.,
Колчунов В.И., Константинов И.С.,
Новиков А.Н., Попова Л.В., Степанов Ю.С.

Главный редактор

Константинов И.С.

Редколлегия

Архипов О.П. (Орел, Россия)
Аверченков В.И. (Брянск, Россия)
Еременко В.Т. (Орел, Россия)
Иванников А.Д. (Москва, Россия)
Коськин А.В. (Орел, Россия)
Подмастерьев К.В. (Орел, Россия)
Поляков А.А. (Москва, Россия)
Савина О.А. (Орел, Россия)
Раков В.И. (Орел, Россия)

Сдано в набор 15.08.2015 г.

Подписано в печать 26.08.2015 г.

Формат 60x88 1/8.

Усл. печ. л. 7,5. Тираж 300 экз.

Заказ №

Отпечатано с готового оригинал-макета
на полиграфической базе
ФГБОУ ВПО «Госуниверситет - УНПК»
302030, г. Орел, ул. Московская, 65

Подписной индекс 15998

по объединенному каталогу

«Пресса России»

Материалы статей печатаются в авторской редакции.

Право использования произведений предоставлено
авторами на основании п. 2 ст. 1286 Четвертой части
ГК РФ.

Журнал входит в Перечень ведущих рецензируемых
научных журналов и изданий, определенных ВАК для
публикации трудов на соискание ученых степеней
кандидатов и докторов наук.

Рубрики номера

1. Математическое и компьютерное моделирование.....5-44
2. Информационные технологии в социально-экономических и организационно-технических системах45-71
3. Автоматизация и управление технологическими процессами и производствами.....72-85
4. Математическое и программное обеспечение вычислительной техники и автоматизированных систем.....86-109
5. Телекоммуникационные системы и компьютерные сети.....110-143
6. Информационная безопасность и защита информации.....144-151

Редакция

О.И. Константинова
А.А. Митин

Адрес учредителя журнала

302020, г. Орел, Наугорское шоссе, 29
(4862) 42-00-24; www.gu-unpk.ru;
E-mail: unpk@ostu.ru

Адрес редакции

302020, г. Орел, Наугорское шоссе, 40
(4862) 43-40-39; www.gu-unpk.ru;
E-mail: konstaoksana@yandex.ru; isit@ostu.ru

Зарег. в Федеральной службе по надзору в сфере
связи, информационных технологий
и массовых коммуникаций.
Св-во о регистрации средства массовой
информации ПИ № ФС77-47350 от 03.11.2011 г.

№ 5 (91) September-October 2015

The journal is published since 2002, leaves six times a year

The founder – State University – Education-Science-Production Complex

Editorial council

Golenkov V.A., president
Radchenko S.Y., vice-president
Borzenkov M.I., secretary

Astafichev P.A., Ivanova T.N., Kirichek A.V.,
Kolchunov V.I., Konstantinov I.S.,
Novikov A.N., Popova L.V., Stepanov Y.S.

Editor-in-chief

Konstantinov I.S.

Editorial board

Arhipov O.P. (Orel, Russia)
Averchenkov V.I. (Bryansk, Russia)
Eremenko V.T. (Orel, Russia)
Ivannikov A.D. (Moscow, Russia)
Koskin A.V. (Orel, Russia)
Podmasteriev K.V. (Orel, Russia)
Polyakov A.A. (Moscow, Russia)
Savina O.A. (Orel, Russia)
Rakov V.I. (Orel, Russia)

*It is sent to the printer's on 15.08.2015,
26.08.2015 is put to bed
Format 60x88 1/8.
Convent. printer's sheets 7,5. Circulation 300 copies
The order №
It is printed from a ready dummy layout
on polygraphic base of State University – ESPC
302030, Orel, Moskovskaya street, 65*

*Index on the catalogue
«Pressa Rossii» 15998*

Journal is included into the list of the Higher Attestation Commission for publishing the results of theses for competition the academic degrees.

In this number

1. Mathematical and computer simulation....5-44
2. Information technologies in social and economic and organizational-technical systems.....45-71
3. Automation and control of technological processes and manufactures.....72-85
4. Software of the computer facilities and the automated systems.....86-109
5. Telecommunication systems and computer networks.....110-143
6. Information and data security.....144-151

The editors

Konstantinova O.I.
Mitin A.A.

The address of the founder of journal

302020, Orel, Highway Naugorskoe, 29
(4862) 42-00-24; www.gu-unpk.ru;
E-mail: unpk@ostu.ru

The address of the editorial office

302020, Orel, Highway Naugorskoe, 40
(4862) 43-40-39; www.gu-unpk.ru;
E-mail: konstaoksana@yandex.ru; isit@ostu.ru

Journal is registered in Federal Service for Supervision in the Sphere of Telecom, Information Technologies and Mass Communications.

*The certificate of registration
ПИ № ФС77-47350 from 03.11.2011.*

© State University – ESPC, 2015

СОДЕРЖАНИЕ

МАТЕМАТИЧЕСКОЕ И КОМПЬЮТЕРНОЕ МОДЕЛИРОВАНИЕ

<i>О.В. ЗАХАРОВА, Т.А. ПОТЛОВА</i> Новое представление АЛУ НФ для операций с плавающей точкой.....	5-13
<i>И.Е. ЛАНДОВСКАЯ, В.Д. ФРОЛОВСКИЙ, В.В. ЛАНДОВСКИЙ</i> Исследование деформационных свойств и автоматизация моделирования тканых материалов.....	14-23
<i>А.В. МАСЛОБОВЕВ, П.А. ЛОМОВ</i> Технология онтологической интеграции на основе автоматического формирования и сравнения контекстов понятий разнородных онтологий.....	24-31
<i>В.И. РАКОВ, Ю.В. ЧЕРНОБРОВКИНА</i> О возможностях использования дискретной формулы академика Котельникова.....	32-44

ИНФОРМАЦИОННЫЕ ТЕХНОЛОГИИ В СОЦИАЛЬНО-ЭКОНОМИЧЕСКИХ И ОРГАНИЗАЦИОННО-ТЕХНИЧЕСКИХ СИСТЕМАХ

<i>А.Ф. АНТИПИН</i> К вопросу о семантическом анализе программ автоматизированных систем управления.....	45-52
<i>В.Н. ВОЛКОВ, Р.А. ЛУНЁВ, А.А. СТЫЧУК, А.Е. ЯСТРЕБКОВ, А.С. БЫЧКОВА, А.Б. НЕЧАЕВА</i> Анализ возможностей совместного использования социальных сетей и геоинформационных сервисов при оказании электронных услуг населению.....	53-61
<i>Н.Г. ПЕНЬКОВ, С.В. ЕРЕМЕНКО, А.Г. СТАВЦЕВА</i> Моделирование учета и приема учетных бланков в системе электронного документооборота.....	62-71

АВТОМАТИЗАЦИЯ И УПРАВЛЕНИЕ ТЕХНОЛОГИЧЕСКИМИ ПРОЦЕССАМИ И ПРОИЗВОДСТВАМИ

<i>В.В. ВЛАСОВ, Н.А. ВЛАСОВА, А.В. ДЕМИДОВ, Н.В. КАНАТНИКОВ, М.В. СМОЛЯКОВ</i> К вопросу о применении беспилотных летательных аппаратов в сфере точного земледелия.....	72-77
<i>М.Н. ОРЕШИНА, А.А. ЛЯКИШЕВ</i> Алгоритм оптимизации процессов информационного обмена в модульных структурах систем сбора и обработки данных АСУ городской сети газоснабжения.....	78-85

МАТЕМАТИЧЕСКОЕ И ПРОГРАММНОЕ ОБЕСПЕЧЕНИЕ ВЫЧИСЛИТЕЛЬНОЙ ТЕХНИКИ И АВТОМАТИЗИРОВАННЫХ СИСТЕМ

<i>А.Н. ГРИГОРЬЕВА</i> Алгоритм эффективного построения множества ключей индекса на основе мультиграмм.....	86-93
<i>Н.А. ОРЕШИН, С.А. ЧЕРЕПКОВ, В.С. ШУМИЛИН, А.Т. ТАШМАТОВ</i> Однопродуктовая двухполюсная потоковая модель транспортной сети.....	94-98
<i>Е.А. ТИТЕНКО</i> MISD-метод и аппаратный семафор готовности результата.....	99-104
<i>Д.А. ТОРОПЧИН, К.В. ЛЫСЫХ, Н.И. КОРСУНОВ</i> Распознавание изображений на основе аппроксимации функций и комбинированных нейронных сетей.....	105-109

ТЕЛЕКОММУНИКАЦИОННЫЕ СИСТЕМЫ И КОМПЬЮТЕРНЫЕ СЕТИ

<i>Г.В. БАБИН, В.И. ХАЛИМОН</i> Программный комплекс для получения и анализа данных в измерительных системах с использованием источника меченых нейтронов.....	110-116
<i>Е.П. ДОЛГОВ</i> Реализация мультипротокольной коммутации по меткам на базе операционных систем с открытым исходным кодом.....	117-122
<i>А.Г. ДУБРОВИН, А.И. ОФИЦЕРОВ, Н.Н. ЕРОФЕЕВ, И.С. ОВЕЧКИН</i> Способ повышения эффективности использования выделенного ресурса в мультисервисных сетях спутниковой связи.....	123-129
<i>С.Ю. МИКОВА, В.С. ОЛАДЬКО</i> Программный комплекс для исследования качества алгоритмов обнаружения сетевых аномалий.....	130-138
<i>С.В. МИТИН</i> Адаптация метода максимального правдоподобия Витерби к декодированию хаотических последовательностей.....	139-143

ИНФОРМАЦИОННАЯ БЕЗОПАСНОСТЬ И ЗАЩИТА ИНФОРМАЦИИ

<i>А.П. ФИСУН, Р.А. ФИСУН</i> Методика оценки эффективности средств комплексного обеспечения информационной безопасности информационных систем и информационно-телекоммуникационных сетей объектов информатизации органов публичной власти.....	144-151
--	---------

CONTENT

MATHEMATICAL AND COMPUTER SIMULATION

- ZAXAROVA O.V., POTLOVA T.A.
New representation of high-speed ALU for operations with floating point.....5-13
- LANDOVSKAYA I.E., FROLOVSKIY V.D., LANDOVSKIY V.V.
Deformation properties investigation and fabric simulation automatization.....14-23
- MASLOBOEV A.V., LOMOV P.A.
Semantic integration technology based on automatic formation and comparison of heterogeneous ontology concept contexts.....24-31
- RAKOV I.V., ChERNOBROVKINA Yu.V.
About the possibilities of using the discrete formula of academician Kotelnikov.....32-44

INFORMATION TECHNOLOGIES IN SOCIAL AND ECONOMIC AND ORGANIZATIONAL-TECHNICAL SYSTEMS

- ANTIPIN A.F.
To the question about semantic analysis of programs of automated control systems.....45-52
- VOLKOV V.N., LUNYOVA R.A., STY'CHUK A.A., YaSTREBKOV A.E., BY'CHKOVA A.S., NEChAEVA A.B.
The analysis of opportunities of sharing of social networks and geoinformation services when rendering electronic services to the population.....53-61
- PEN'KOV N.G., ERYOMENKO S.V., STAVCEVA A.G.
Modeling of the account and the receiving accounting forms in the electronic document management system.....62-71

AUTOMATION AND CONTROL OF TECHNOLOGICAL PROCESSES AND MANUFACTURES

- VLASOV V.V., VLASOVA N.V., DEMIDOV A.V., KANATNIKOV N.V., SMOLYAKOV M.V.
On application of usage unmanned aerial vehicles in sphere of precision agriculture.....72-77
- OREShINA M.N., LYAKISHEV A.A.
Algorithm of optimization of the process information exchange in the modular structure of data collection and processing systems in the ACS of a city gas supply network.....78-85

SOFTWARE OF THE COMPUTER FACILITIES AND THE AUTOMATED SYSTEMS

- GRIGOR'EVA A.N.
The algorithm for effective keyset construction for a multigram-based index.....86-93
- OREShIN N.A., ChEREPKOV S.A., ShUMILIN V.S., TASHMATOV A.T.
Single-product double-pole streaming model of the transport network.....94-98
- TITENKO E.A.
Misd-method and hardware semaphore of final result.....99-104
- TOROPChIN D.A., LY'SY'X K.V., KORSUNOV N.I.
Image recognition based on approximation of functions and neural networks combined.....105-109

TELECOMMUNICATION SYSTEMS AND COMPUTER NETWORKS

- BABIN G.V., XALIMON V.I.
Software for obtaining and analysing of data in measuring systems using tagged neutron source.....110-116
- DOLGOV E.P.
Implementation of multiprotocol label switching based on open source operating systems.....117-122
- DUBROVIN A.G., OFICEROVA A.I., EROFEEV N.N., OVEChKIN I.S.
Method of increase of efficiency of usage of the selected resource on multiservice communication satellite networks.....123-129
- MIKOVA S.Yu., OLAD'KO V.S.
Software for research quality algorithms to detect network anomalies.....130-138
- MITIN S.V.
Adaptation maximum likelihood Viterbi's method for decoding of chaotic sequences.....139-143

INFORMATION AND DATA SECURITY

- FISUN A.P., FISUN R.A.
Methods of assessing the effectiveness of an integrated information security of information systems and information and telecommunication networks objects of informatization of public authorities.....144-151

УДК 681.3.063

О.В. ЗАХАРОВА, Т.А. ПОТЛОВА

НОВОЕ ПРЕДСТАВЛЕНИЕ АЛУ НФ ДЛЯ ОПЕРАЦИЙ С ПЛАВАЮЩЕЙ ТОЧКОЙ

В статье предлагается логическая структура быстродействующего арифметико-логического устройства (АЛУ) для операций с плавающей точкой, разработанная на основе технологии непосредственного формирования результата. Предлагаются инструментальные средства для моделирования АЛУ непосредственного формирования (АЛУ НФ) с плавающей точкой и поиска подходящей структуры устройства. Даются ориентировочные оценки сложности кристалла СБИС, обрабатывающего вычислительные операции (для АЛУ НФ с плавающей точкой).

Ключевые слова: производительность; арифметико-логическое устройство (АЛУ); вычисления с плавающей точкой.

Известно [1-3], что в ЭВМ используется естественная или нормальная форма представления двоичных чисел. Естественная форма двоичного числа записывается только с помощью набора значащих цифр, а для представления нормальной формы числа используется запись

$$A = \pm m_A \cdot q^{\pm P_A}, \quad (1)$$

где m_A – мантисса числа A ; q – основание системы счисления; P_A – порядок.

Анализ показал, что при представлении чисел в формате с плавающей точкой имеется значительно больший диапазон значений и высокая точность результатов вычислительных операций, а при выполнении операций над числами в естественной форме не исключена возможность того, что значение результирующего операнда окажется вне диапазона допустимых значений. Поскольку известные структуры арифметико-логических устройств непосредственного формирования (АЛУ НФ) предложены в формате с фиксированной точкой и полезность представления чисел в формате с плавающей точкой очевидна, то актуально создание структур АЛУ НФ, ориентированных на формат с плавающей точкой.

ТЕХНОЛОГИЯ НЕПОСРЕДСТВЕННОГО ФОРМИРОВАНИЯ РЕЗУЛЬТАТА

Технология непосредственного формирования результата заключается в следующем (рис. 1) [4-7]:

- формирование таблицы соответствия (автоматной таблицы);
- построение унифицированных совершенных нормальных форм для каждого разряда результирующего операнда:
- построение унифицированных дизъюнктивных совершенных нормальных форм (УДСНФ):

$$Y_\varphi = \sum_{i=1}^{2^{2^n}} (f_i \cdot \varphi_i),$$

где Y_φ – УДСНФ для соответствующего разряда результирующего операнда; f_i – значение функции на соответствующем наборе входных переменных; φ_i – характеристическая функция единицы для соответствующего набора входных переменных; n – разрядность входного операнда;

– построение унифицированных конъюнктивных совершенных нормальных форм (УКСНФ):

$$Y_{\psi} = \prod_{i=1}^{2^{2n}} (f_i \vee \psi_i),$$

где Y_{ψ} – УКСНФ для соответствующего разряда результирующего операнда; ψ_i – характеристическая функция нуля для соответствующего набора входных переменных.

– построение структурных единиц АЛУ НФ по унифицированным совершенным нормальным формам (УДСНФ или УКСНФ).

ВЫПОЛНЕНИЕ ОПЕРАЦИЙ НАД ЧИСЛАМИ С ПЛАВАЮЩЕЙ ТОЧКОЙ В АЛУ НФ

Входными данными АЛУ НФ для операций с плавающей точкой являются:

- операнд $A = \pm m_A \cdot q^{\pm P_A}$;
- операнд $B = \pm m_B \cdot q^{\pm P_B}$;
- код операции KOn .

Будем полагать, что код операции (KOn) указывает на операции, которые должны быть исполнены, а также определяет одновременно вычисляемые операции. Для этого положим, что число разрядов в коде операции будет равняться количеству реализуемых элементарных операций.

На выходе АЛУ НФ с плавающей точкой формируется результат $C = \pm m_C \cdot q^{\pm P_C}$.

Учитывая (1), операция умножения с плавающей точкой реализуется на основе известного соотношения [1-3]:

$$A \cdot B = (m_A \cdot q^{P_A}) \cdot (m_B \cdot q^{P_B}) = (m_A \cdot m_B) \cdot q^{(P_A + P_B)}. \quad (2)$$

Таким образом, для реализации операции умножения над числами с плавающей точкой в АЛУ НФ необходимы:

- ядро АЛУ НФ для перемножения мантисс ($m_C = m_A \cdot m_B$);
- ядро АЛУ НФ для сложения порядков ($P_C = P_A + P_B$).

Деление с плавающей точкой реализуется на основе следующего соотношения [1-3]:

$$\frac{A}{B} = \frac{(m_A \cdot q^{P_A})}{(m_B \cdot q^{P_B})} = \frac{m_A}{m_B} \cdot \frac{q^{P_A}}{q^{P_B}} = \frac{m_A}{m_B} \cdot q^{(P_A - P_B)}. \quad (3)$$

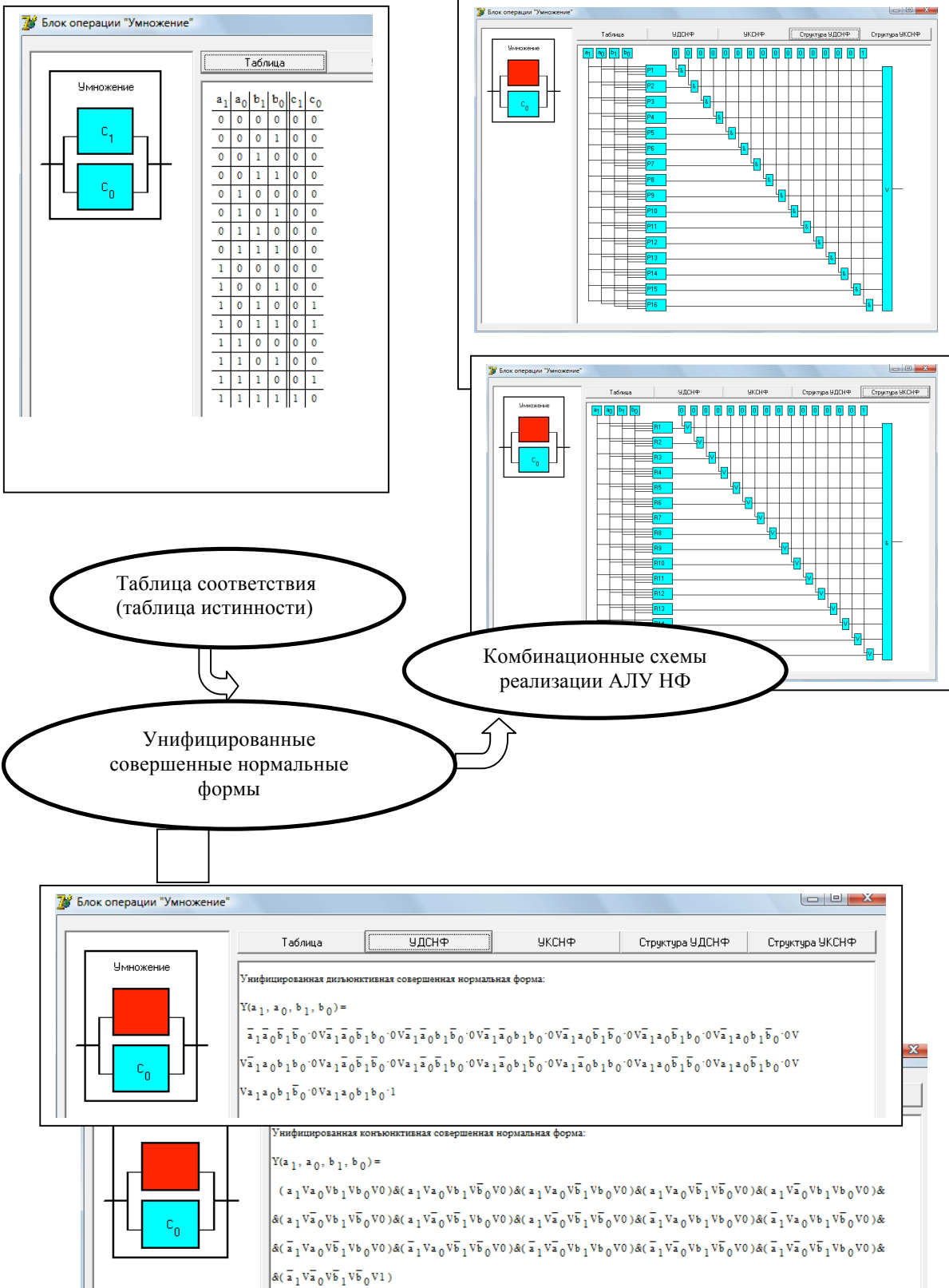
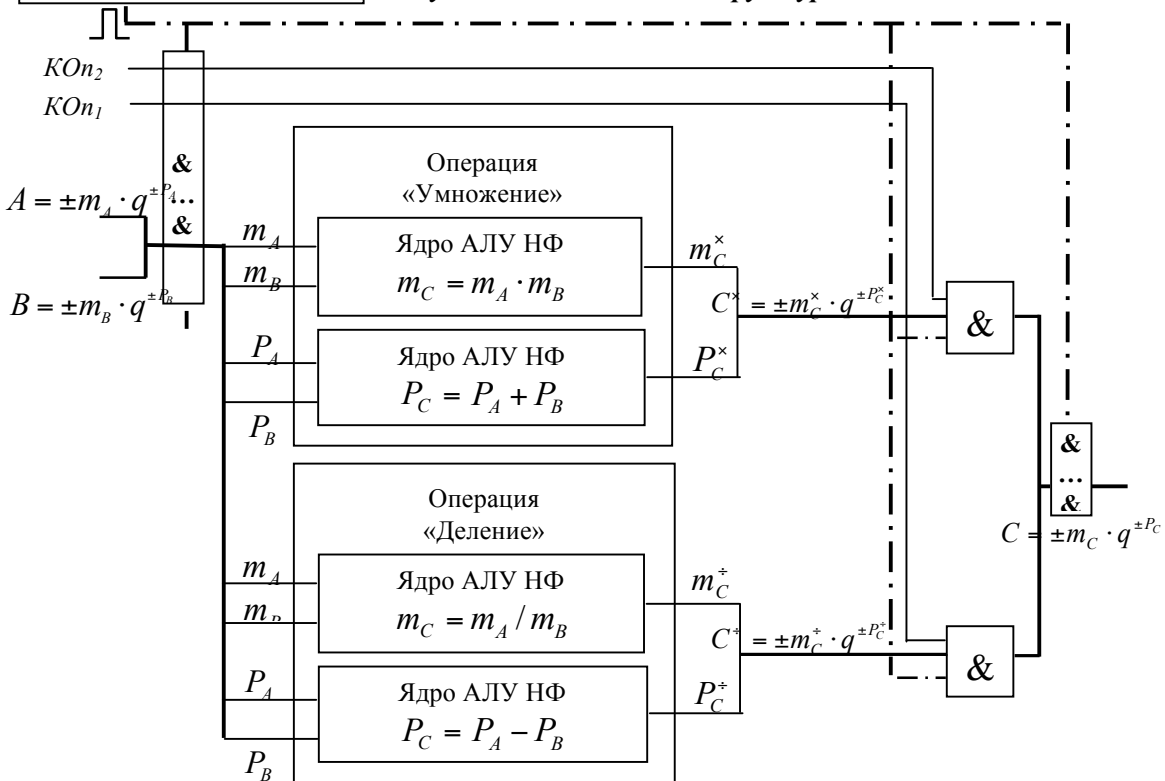
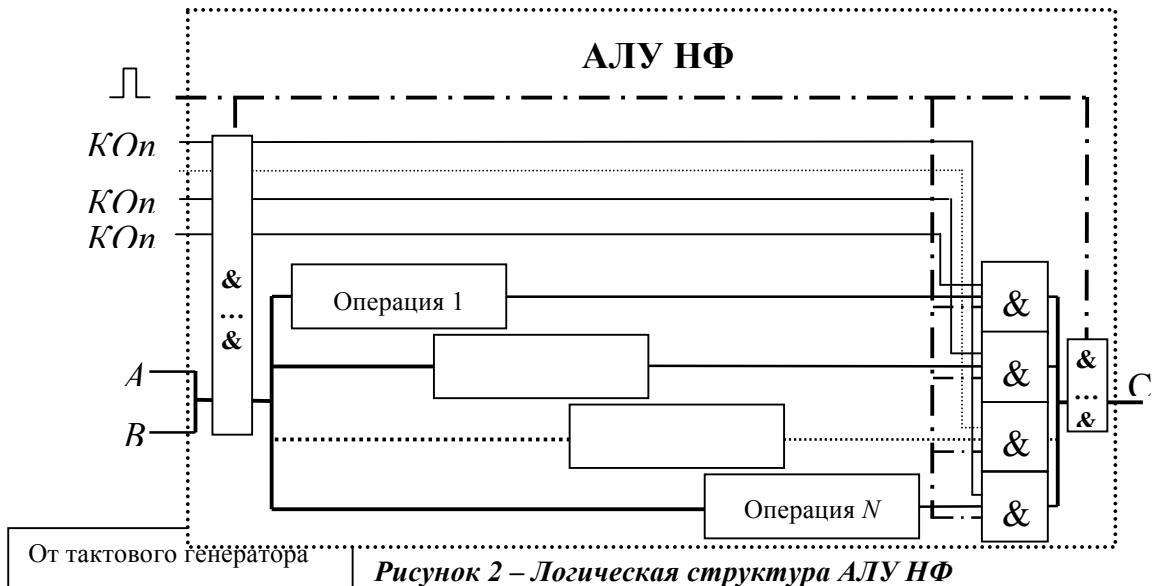


Рисунок 1 – Принцип реализации технологии непосредственного формирования результата

Таким образом, для реализации операции деления над операндами с плавающей точкой в АЛУ НФ необходимы:

- ядро АЛУ НФ для деления мантисс ($m_c = \frac{m_A}{m_B}$);
- ядро АЛУ НФ для вычитания порядков ($P_c = P_A - P_B$).

Общая логическая структура АЛУ НФ представлена на рисунке 2, а логическая структура АЛУ НФ с плавающей точкой для операций «умножение» и «деление» представлена на рисунке 3.



ПРОГРАММНЫЙ ИНСТРУМЕНТАРИЙ МОДЕЛИРОВАНИЯ АЛУ НФ С ПЛАВАЮЩЕЙ ТОЧКОЙ

Разработанные программные инструментальные средства моделирования АЛУ НФ с плавающей точкой представлены на рисунке 4 [8].

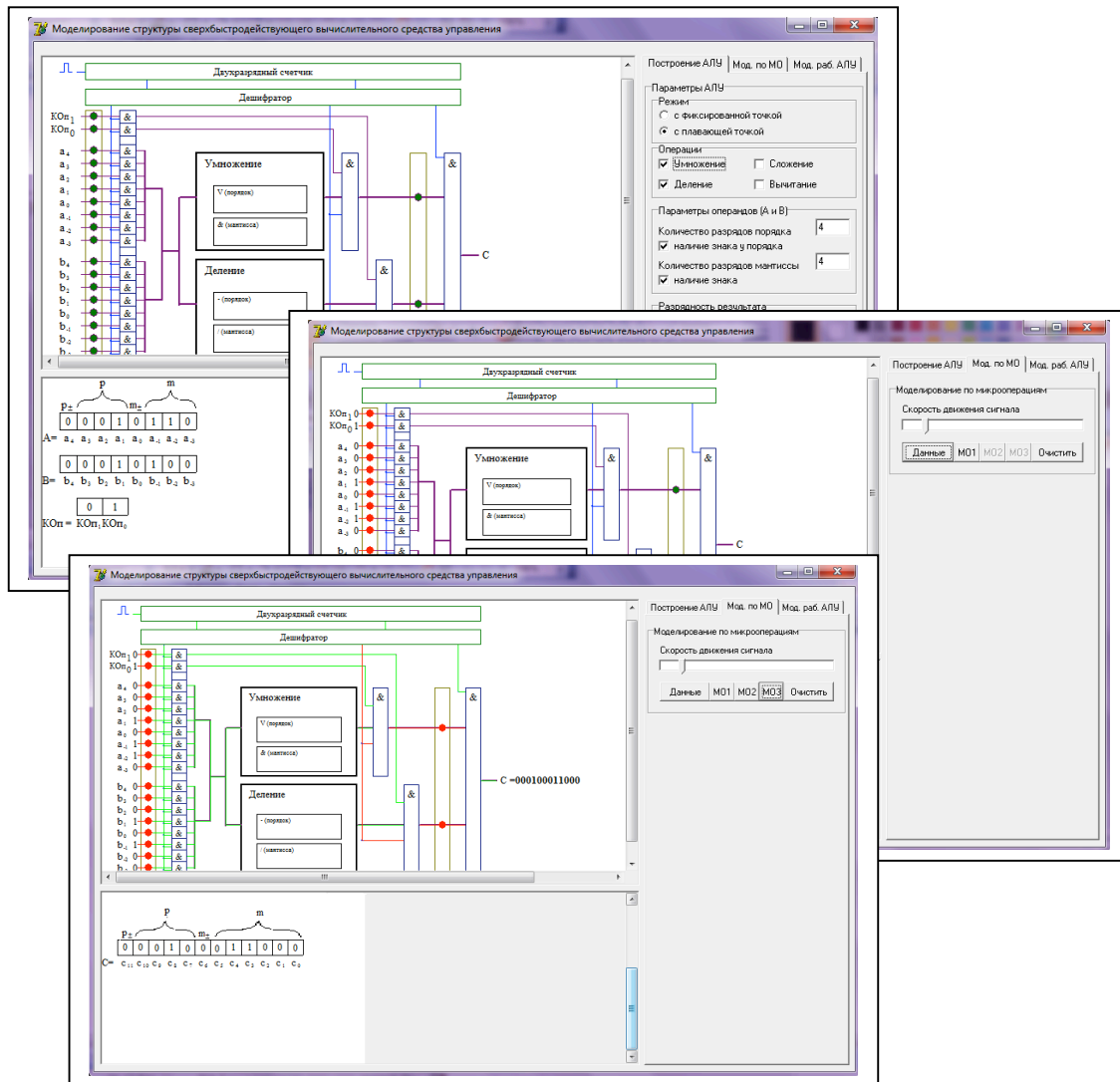


Рисунок 4 – Программный инструментальный моделирования АЛУ НФ с плавающей точкой

ВЫВОДЫ

Несмотря на то, что в представлении операндов выделяется мантисса и порядок, что приводит к возникновению фактически двух групп автоматных таблиц (в сравнении с одной группой для АЛУ НФ с фиксированной точкой), сложность кристалла СБИС по условному числу используемых транзисторов оказывается меньше. В частности, для восьмиразрядной сетки чисел, как показано в [4] (табл. 1), сложность СБИС для АЛУ НФ с фиксированной точкой составляет ориентировочно 24 117 248 транзисторов, а для АЛУ НФ с плавающей точкой – 835 824 транзистора (табл. 2), что вкладывается в границы возможного изготовления кристаллов СБИС даже по 800-нанометровой технологии микропроцессора Intel i486DX (1989 г.), обеспечивающего 1,2 млн транзисторов в кристалле, а если использовать 45-нанометровую технологию как для микропроцессора AMD Phenom II X4 (модели 920 и 940 Black Edition на ядре Deneb) (2009 г.), обеспечивающего 758 млн

транзисторов в кристалле, то можно реализовать АЛУ НФ для шестнадцатиразрядной сетки чисел [9].

Таблица 1 – Ориентировочные оценки сложности кристалла СБИС, обрабатывающего вычислительные операции (для АЛУ НФ с фиксированной точкой) [4]

Разрядность первого операнда	Разрядность второго операнда	Количество строк таблицы истинности для вычислительной операции: число дизъюнкций в ДСНФ	Число конъюнкций в одной дизъюнкции ДСНФ (не более)	Среднее число инверсий в одной дизъюнкции ДСНФ	Ориентировочная оценка количества транзисторов в кристалле СБИС (не более)	
					для реализации логической функции одного разряда	для реализации логических функций, формирующих все разряды результата вычислительной операции
4	4	256	7	4	$256(7+4)=2\ 816$	22 528 для 8-ми разрядов
6	6	4096	11	6	$4096(11+6)=69\ 632$	835 584 для 12-ти разрядов
7	7	16386	13	7	$16386(13+7)=327\ 720$	4 588 080 для 14-ти разрядов
8	8	65536	15	8	$65536(15+8)=1\ 507\ 328$	24 117 248 для 16-ти разрядов
9	9	262144	17	9	$262144(17+9)=6\ 815\ 744$	122 683 392 для 18-ти разрядов
10	10	1 048 576	19	10	$1048576(19+10)=30\ 408\ 704$	608 174 080 для 20-ти разрядов

Таблица 2 – Ориентировочные оценки сложности кристалла СБИС, обрабатывающего вычислительные операции (для АЛУ НФ с плавающей точкой)

Разрядность операнда			Количество строк таблицы истинности для вычислительной операции: число дизъюнкций в ДСНФ		Число конъюнкций в одной дизъюнкции ДСНФ (не более)		Среднее число инверсий в одной дизъюнкции ДСНФ		Ориентировочная оценка количества транзисторов в кристалле СБИС (не более)				ИТОГО
									для реализации логической функции одного разряда		для реализации логических функций, формирующих все разряды результата вычислительной операции		
порядок	мантисса	всего	порядок	мантисса	порядок	мантисса	порядок	мантисса	порядок	мантисса	порядок	мантисса	
2	6	8	16	4096	3	11	2	6	$16 \times (3+2) = 80$	69 632	240 для 3-х разрядов	835 584 для 12-ти разрядов	835 824
4	4	8	256	256	7	7	4	4	$256 \times (7+4) = 2 816$	2 816	14 080 для 5-ти разрядов	22 528 для 8-ми разрядов	36 608
4	6	10	256	4096	7	11	4	6	2 816	69 632	14 080 для 5-ти разрядов	835 584 для 12-ти разрядов	849 664
4	8	12	256	65536	7	15	4	8	2 816	1 507 328	14 080 для 5-ти разрядов	24 117 248 для 16-ти разрядов	24 131 328
6	10	16	4096	1 048 576	11	19	6	10	$4096 \times (11+6) = 69 632$	30 408 704	487 424 для 7-ми разрядов	608 174 080 для 20-ти разрядов	608 661 504

Исследование выполнено при поддержке Фонда содействия развитию малых форм предприятий в научно-технической сфере в рамках программы «УМНИК».

СПИСОК ЛИТЕРАТУРЫ

1. Поспелов Д.А. Арифметические основы вычислительных машин дискретного действия: учебное пособие для ВУЗов. – М.: «Высшая школа», 1970. – 308 с.
2. Шауман А.М. Основы машинной арифметики. – Л.: Изд-во Ленингр. ун-та, 1979. – 312 с.
3. Карцев М.А. Арифметика цифровых машин. – М.: Наука, 1969. – 576 с.

4. Захарова О.В., Раков В.И. Реализация быстрых вычислений посредством комбинационных структур // Вестник компьютерных и информационных технологий, 2007. – № 8. – С. 51-55.
5. Захарова О.В., Раков В.И. Моделирование комбинационных структур для реализации вычислительных операций // Известия ОрелГТУ. Информационные системы и технологии. – Орел: ОрелГТУ, 2009. – № 1/51(562). – С. 91-98.
6. Захарова О.В. Программный инструментарий для моделирования вычислительных средств управления // Информационные системы и технологии, 2012. – № 6(74). – С. 90-99.
7. Захарова О.В., Раков В.И. Программный инструментарий для моделирования вычислительных средств управления // Свидетельство о государственной регистрации программ для ЭВМ № 2012612287, 2012.
8. Захарова О.В., Раков В.И. Моделирование структуры сверхбыстродействующего вычислительного средства управления // Свидетельство о государственной регистрации программ для ЭВМ № 2013617875, 2013.
9. Захарова О.В., Раков В.И. Моделирование цифровых регуляторов: монография.

Захарова Ольга Владимировна

ФГБОУ ВПО «Госунiversитет – УНПК», г. Орел
Кандидат технических наук, доцент кафедры «Информационные системы»
Тел.: 8 906 569-20-20
E-mail: cvaig@mail.ru

Потлова Татьяна Анатольевна

ФГБОУ ВПО «Госунiversитет – УНПК», г. Орел
Студент
Тел.: 8 906 569-20-20
E-mail: +79092273044@yandex.ru

*O.V. ZAXAROVA (Candidate of Engineering Sciences,
Associate Professor of the Department «Information Systems»*

*T.A. POTLOVA (Student)
State University – ESPC, Orel*

NEW REPRESENTATION OF HIGH-SPEED ALU FOR OPERATIONS WITH FLOATING POINT

The article proposes a logical structure for high-speed arithmetic and logical unit (ALU) for floating point operations, developed on the basis of the technology of directly forming the result. Offers tools for modeling high-speed ALU floating-point and search for a suitable structure of the device. Given estimates of the complexity of the crystal VLSI, implementing the computational procedure (for high speed ALU floating-point).

Keywords: *productivity; arithmetic and logic unit (ALU); floating point.*

BIBLIOGRAPHY (TRANSLITERATED)

1. Pospelov D.A. Arifmeticheskie osnovy' vy'chislitel'ny'x mashin diskretnogo dejstviya: uchebnoe posobie dlya VUZov. – M.: «Vy'sshaya shkola», 1970. – 308 s.
2. Shauman A.M. Osnovy' mashinnoj arifmetiki. – L.: Izd-vo Leningr. un-ta, 1979. – 312 s.
3. Karcev M.A. Arifmetika cifrov'x mashin. – M.: Nauka, 1969. – 576 s.
4. Zaxarova O.V., Rakov V.I. Realizaciya by'stry'x vy'chislenij posredstvom kombinacionny'x struktur // Vestnik komp'yuterny'x i informacionny'x tehnologij, 2007. – № 8. – S. 51-55.
5. Zaxarova O.V., Rakov V.I. Modelirovanie kombinacionny'x struktur dlya realizacii vy'chislitel'ny'x operacij // Izvestiya OrelGTU. Informacionny'e sistemy' i tehnologii. – Orel: OrelGTU, 2009. – № 1/51(562). – S. 91-98.
6. Zaxarova O.V. Programmny'j instrumentarij dlya modelirovaniya vy'chislitel'ny'x sredstv upravleniya // Informacionny'e sistemy' i tehnologii, 2012. – № 6(74). – S. 90-99.

7. Zaxarova O.V., Rakov V.I. Programmny'j instrumentarij dlya modelirovaniya vy'chislitel'ny'x sredstv upravleniya // Svidetel'stvo o gosudarstvennoj registracii programm dlya E'VM № 2012612287, 2012.
8. Zaxarova O.V., Rakov V.I. Modelirovanie struktury' sverxby'strojstvuyushhego vy'chislitel'nogo sredstva upravleniya // Svidetel'stvo o gosudarstvennoj registracii programm dlya E'VM № 2013617875, 2013.
9. Zaxarova O.V., Rakov V.I. Modelirovanie cifrov'y'x regulyatorov: monografiya.

УДК 004.925.83

И.Е. ЛАНДОВСКАЯ, В.Д. ФРОЛОВСКИЙ, В.В. ЛАНДОВСКИЙ

ИССЛЕДОВАНИЕ ДЕФОРМАЦИОННЫХ СВОЙСТВ И АВТОМАТИЗАЦИЯ МОДЕЛИРОВАНИЯ ТКАНЫХ МАТЕРИАЛОВ

В работе представлен разработанный комплекс программных средств, включающий все этапы компьютерного моделирования сборки изделий из тканых материалов. Рассмотрены методы экспериментального исследования деформационных свойств ткани. Описаны варианты использования в программном продукте экспериментально полученных данных. Приводится визуальное сравнение реального изделия и его моделируемого аналога.

Ключевые слова: компьютерное моделирование ткани; исследование деформационных свойств тканых материалов; программное обеспечение для моделирования сборки изделий из ткани; физические свойства ткани; метод частиц; схема с перешагиванием.

ВВЕДЕНИЕ

На сегодняшний день наиболее прогрессивной технологией проектирования изделий многих отраслей промышленности, в том числе и швейной, является трехмерное автоматизированное конструирование. 3D-технология на базе современной компьютерной техники активно входит в практику проектирования, позволяя существенно сократить сроки разработки проектов. Одним из приоритетных направлений совершенствования современных методов трехмерного проектирования является разработка программных модулей, позволяющих еще на этапе конструирования наглядно оценить посадку проектируемого изделия и его эстетические показатели на трехмерном манекене. В качестве примера можно привести моделирование новых фасонов одежды, ортопедических конструкций, эргономичной мебели, кресел автомобилей и пр. Если моделирование будет достаточно точным, то можно обойтись без создания пробных вариантов и заранее исправить все несоответствия предлагаемого проектного решения конкретной поставленной задаче.

Большинство существующих систем автоматизированного проектирования для легкой промышленности (САПР) обладают достаточно широкими возможностями, но не позволяют задавать конкретные деформационные свойства материалов, от которых и зависит итоговый вид изделия. К тому же значительная доля этих САПР использует геометрические подходы к деформационному моделированию поверхности ткани, которые не повторяют точное поведение ткани, а лишь имитируют его, образуя случайные складки, что существенно отражается на качестве получаемых результатов. Поэтому авторы данной статьи предлагают совместить работу в уже существующих графических средах, таких, как AutoCAD, 3ds Max и других, используя их для создания объектов, на поверхности которых будет производиться сборка, а также для итогового наложения текстур, чтобы максимально приблизить вид готового изделия к реальности. А для создания деталей изделия и их сборки на поверхности манекена использовать разработанный авторами программный комплекс для виртуального трехмерного моделирования поведения тканых материалов.

МОДЕЛИРОВАНИЕ ДЕФОРМАЦИОННЫХ СВОЙСТВ ТКАНИ

Ткань – один из древнейших и наиболее распространенных видов текстильных материалов. Ткань представляет собой пространственную сетку, образованную переплетением в определенной последовательности двух взаимно перпендикулярных систем

нитей [1]. Нити, идущие вдоль полотна, называются основой; нити, расположенные поперек полотна, называются утком.

Основная идея компьютерного моделирования ткани заключается в том, чтобы рассматривать полотно как систему взаимодействующих частиц (узлов), которые размещены в точках пересечения продольных и поперечных нитей ткани (рис. 1) [2].

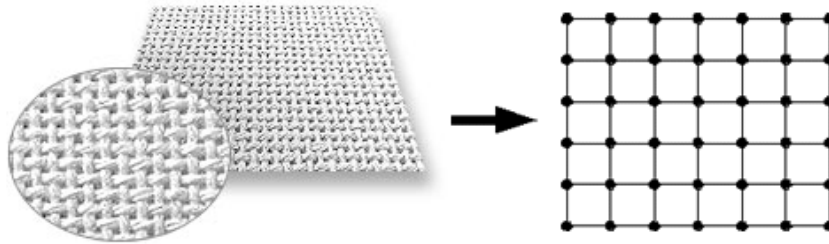


Рисунок 1 – Дискретная модель ткани

Основные взаимодействия, которые происходят на уровне нити, это растяжение-сжатие, изгиб и сдвиг. Учитывая, что ткань достаточно легкая и вес ткани в удаленных узлах оказывает пренебрежимо малое влияние на каждую рассматриваемую частицу, предположим, что на каждую внутреннюю частицу влияют 12 соседних частиц. На рисунке 2 взаимодействия растяжения-сжатия, сдвига и изгиба обозначены связями 1, 2 и 3 соответственно.

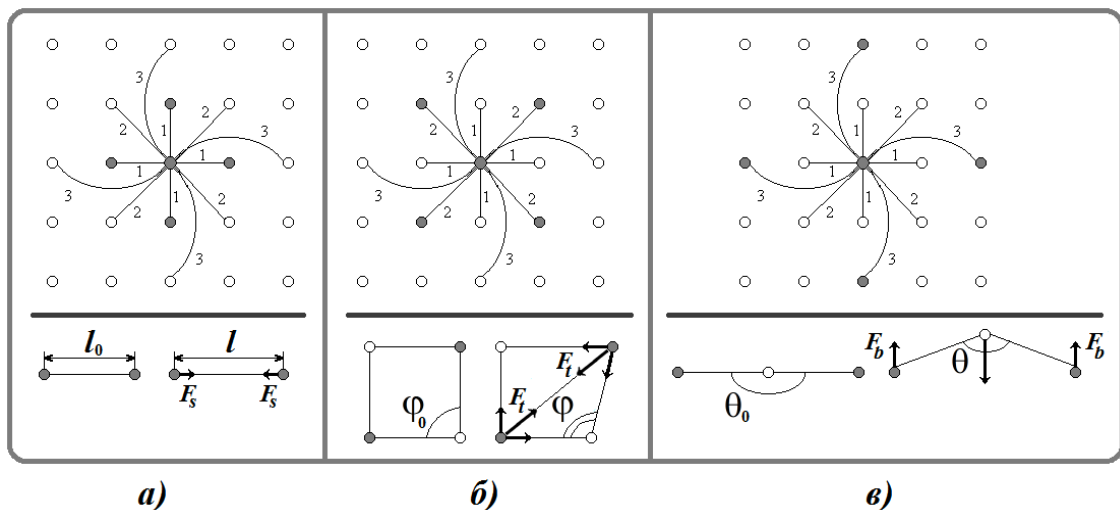


Рисунок 2 – Типы взаимодействий:

а) – растяжение-сжатие; б) – сдвиг; в) – изгиб

Взаимодействие растяжения-сжатия имеет место, когда расстояние между двумя соседними частицами (связи 1, рис. 2 (а)) отличается от расстояния между ними в состоянии покоя. В случае растяжения-сжатия направление сил очевидно, так как во взаимодействии учитываются положения только двух частиц. Значение силы между соседними частицами P_i и P_j определяется исходя из диаграммы «нагрузка-удлинение» для конкретного материала.

При этом кривые удлинения строятся отдельно для нитей основы и нитей утка.

Взаимодействие сдвига (связи 2, рис. 2 (б)) можно выразить через взаимодействие растяжения-сжатия частиц, образующих диагональ ячейки. Значение силы между соседними частицами P_i и P_j , образующими диагональ, будем определять, исходя из диаграммы

«нагрузка-удлинение», полученной при растяжении материала под углом 45° к нити основы для конкретного материала.

Последний вид взаимодействия – взаимодействие изгиба (связи 3, рис. 2 (в)) можно выразить как функцию угла, образованного тремя последовательными частицами, лежащими на одной прямой (нити). Значение силы между частицами P_i и P_j , расположенными через узел, будем определять, используя для расчетов условное значение жесткости, получаемое для проб продольного и поперечного направлений отдельно.

В итоге на частицу будут действовать три вида сил: силы, действующие на узел по прямой; силы, действующие на узел по косой (определяющие сдвиг); силы, действующие через узел (определяющие изгиб) [3]. Результирующую силу по каждому виду взаимодействий, действующую на частицу P_i , определим как сумму конкретного вида сил:

$$F_{s,i} = \sum_{j \in R_{s,i}} F_{s,ij}; \quad F_{t,i} = \sum_{j \in R_{t,i}} F_{t,ij}; \quad F_{b,i} = \sum_{j \in R_{b,i}} F_{b,ij},$$

где $R_{s,i}$, $R_{t,i}$, $R_{b,i}$ – множество индексов частиц связанных с частицей P_i взаимодействиями растяжения-сжатия, сдвига и изгиба соответственно; R_i – множество индексов всех частиц, взаимодействующих с частицей P_i .

Движение всей системы описывается обобщенными перемещениями:

$$r_i(t) = \{x_i(t), y_i(t), z_i(t)\},$$

где x_i, y_i, z_i – координаты частицы в трехмерном пространстве; t – время.

На каждом временном слое ищутся положения узлов в пространстве. Каждая из частиц обладает некоторой массой, находится в гравитационном поле, взаимодействует с окружающей средой и соседними частицами [4]. Тогда уравнение движения частицы P_i имеет следующий вид:

$$m_i r_i'' = m_i g - m_i c_i r_i' + \sum_{j \in R_i} F_{int}(r_i, r_j),$$

где m_i – масса частицы; c_i – константа демпфирования; составляющая $c_i r_i'$ представляет собой потери энергии связанные с взаимодействием частицы с окружающей средой; g – ускорение свободного падения. Последняя составляющая уравнения движения представляет собой результирующую силу взаимодействий между частицами; R_i – множество индексов узлов связанных с узлом P_i .

В силу того, что количество частиц, как правило, велико и каждое дополнительное вычисление силы требует больших временных затрат, для решения системы воспользуемся схемой с перешагиванием:

$$\begin{aligned} v_{n+1} &= v_n + hM^{-1}F(r_n, v_n), \\ r_{n+1} &= r_n + hv_{n+1}. \end{aligned}$$

Это метод второго порядка точности в сравнении с классическим явным методом Эйлера обладает большей устойчивостью, позволяя на порядок увеличить шаг интегрирования [5].

ПОЛУЧЕНИЕ КОЭФФИЦИЕНТОВ ДЕФОРМАЦИИ

Испытания материалов для получения характеристик деформации растяжения-сжатия и сдвига проводят на разрывной машине РТ-250 и ее модификациях (рис. 3 (а)).

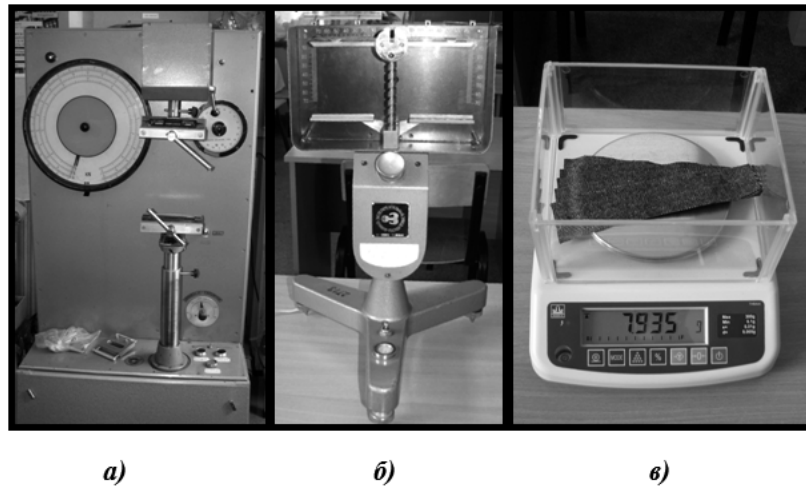


Рисунок 3 – Измерительные приборы:

а) – разрывная машина РТ-250; б) – прибор ПТ-2; в) – электронные весы

Испытания проводятся стрип-методом, подготовка и проведение эксперимента соответствует ГОСТ 3813-72. Для получения характеристик растяжения-сжатия материала отдельно проводят испытания образцов вдоль нити основы и вдоль нити утка. Для получения характеристик сдвига материала образец растягивают под углом 45° относительно нити основы. После испытания на основе полученных данных строится диаграмма «нагрузка-удлинение» (рис. 4).

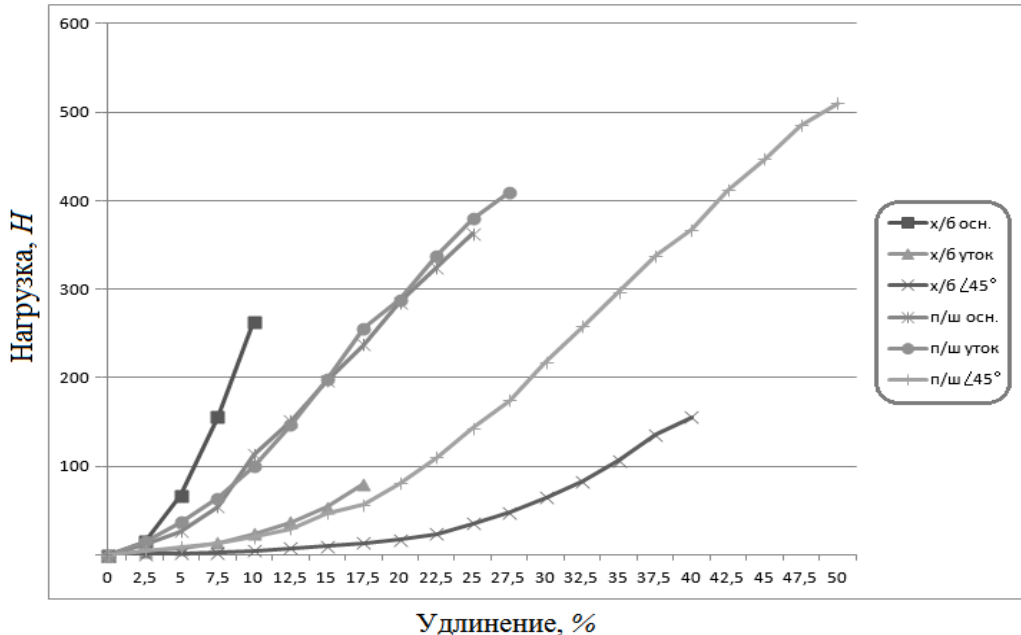


Рисунок 4 – Диаграмма «нагрузка-удлинение» для хлопчатобумажной и полушерстяной тканей

Так как конечным этапом моделирования является получение внешнего вида готового изделия, то будем считать, что все деформации, возникающие в ткани, имеют только упругую составляющую, а, следовательно, можно не учитывать эффект гистерезиса. Вычисление значений сил растяжения, сжатия или сдвига, возникающих между соседними

частицами P_i и P_j при вычислениях в программе, осуществляется в два этапа. На первом этапе следует вычислить процентное соотношение расстояния между частицами на данном шаге интегрирования и расстояния в положении равновесия и, используя данные диаграммы «нагрузка-удлинение», получить силу, возникающую в материале, по площади равном размеру пробы (200×50 мм или 100×50 мм), для которой производились измерения. На втором этапе следует рассчитать силу на единицу площади, равную площади ячейки сетки, которой представлена ткань.

Испытания материалов для получения характеристик изгиба проводят по методу консоли на приборе ПТ-2 (рис. 3 (б)). Подготовка и проведение испытания проводится согласно ГОСТ 10550-93. Для получения характеристик изгиба материала проводят отдельно испытания образцов вдоль нити основы и вдоль нити утка [6]. При проведении данного исследования также используются электронные весы высокой точности (рис. 3 (в)) для определения массы проб. После проведения испытания мы получаем условное значение жесткости $EI(H \cdot м^2)$ для проб продольного и поперечного направления.

Используя следующее выражение, можем вычислить в программе силу изгиба, возникающую между частицами P_i и P_j , расположенными через узел:

$$F_b = \frac{\left(\frac{EI}{b^2(1-\nu^2)} \Delta\theta \right)}{N},$$

где EI – жесткость при изгибе, полученная экспериментально; b – длина ребра ячейки сети; $\Delta\theta$ – угол между двумя ребрами (угол изгиба нити); ν – значение коэффициента Пуассона для данного материала; N – количество линий сетки, расположенных по ширине пробы (160×30 мм), для которой производились измерения, при данном размере сетки, представляющей тканый материал.

Следует отметить, что плюсами получения коэффициентов тканых материалов описанными выше методами является то, что данные значения могут использоваться в программе вне зависимости от размера ячейки сетки, которой представляется ткань, и уже включают в себя силу трения, возникающую между нитями данного материала при его деформировании.

КОМПЬЮТЕРНОЕ МОДЕЛИРОВАНИЕ ТКАНЫХ МАТЕРИАЛОВ

Подготовительным этапом перед началом моделирования являются: экспериментальное получение коэффициентов деформации выбранного тканого материала, построение манекена, на поверхности которого будет происходить сборка, создание деталей изделия, из которых будет производиться сборка, и создание швов, соединяющих детали. Методы получения коэффициентов дифференциальных уравнений были подробно описаны в разделе 3. Полученные коэффициенты сохраняются в базе данных, чтобы исключить повторное проведение экспериментов для данного материала.

1. Построение манекена. Построение манекена, на поверхности которого будет происходить сборка, может осуществляться в любой среде создания и редактирования трехмерной графики, например, таких, как AutoCAD, 3ds Max, Rhinoceros, Poser, Maya и других. Единственным условием является то, чтобы объект состоял из точек, которые в дальнейшем могут быть преобразованы в грани объекта. После завершения построения объекта он должен быть импортирован в AutoCAD, откуда с помощью файлов на языке LISP объект преобразуется в нужный формат для загрузки в программу моделирования тканых материалов (объект должен состоять из треугольных граней). Для примера: псевдокод

одного из файлов на языке LISP, который производит экспорт объекта, состоящего из прямоугольников, представлен в листинге 1:

```

открыть для записи файл a
присвоить переменной i значение 0
пока значение i меньше количества граней объекта
сохранить координаты первой вершины грани в r1
сохранить координаты второй вершины грани в r2
сохранить координаты третьей вершины грани в r3
сохранить координаты четвертой вершины грани в r4
записать в файл r1, r2, r3
записать в файл r3, r4, r1
увеличить значение i на 1
закрывать файл a
    
```

Листинг 1 – Псевдокод файла экспорта объекта из AutoCAD

2. Программный комплекс для моделирования. Программный комплекс разработан в среде Microsoft Visual Studio 2012 на языке C++ с использованием графической библиотеки OpenGL и механизмов написания параллельных программ OpenMP, MPI и C++ AMP. Он состоит из четырех основных модулей – модуля построения сетки частиц деталей, модуля создания соединительных швов, модуля расстановки объектов и деталей и модуля моделирования сборки изделия на поверхности манекена. Общая схема моделирования с учетом входных и выходных данных представлена на рисунке 5.

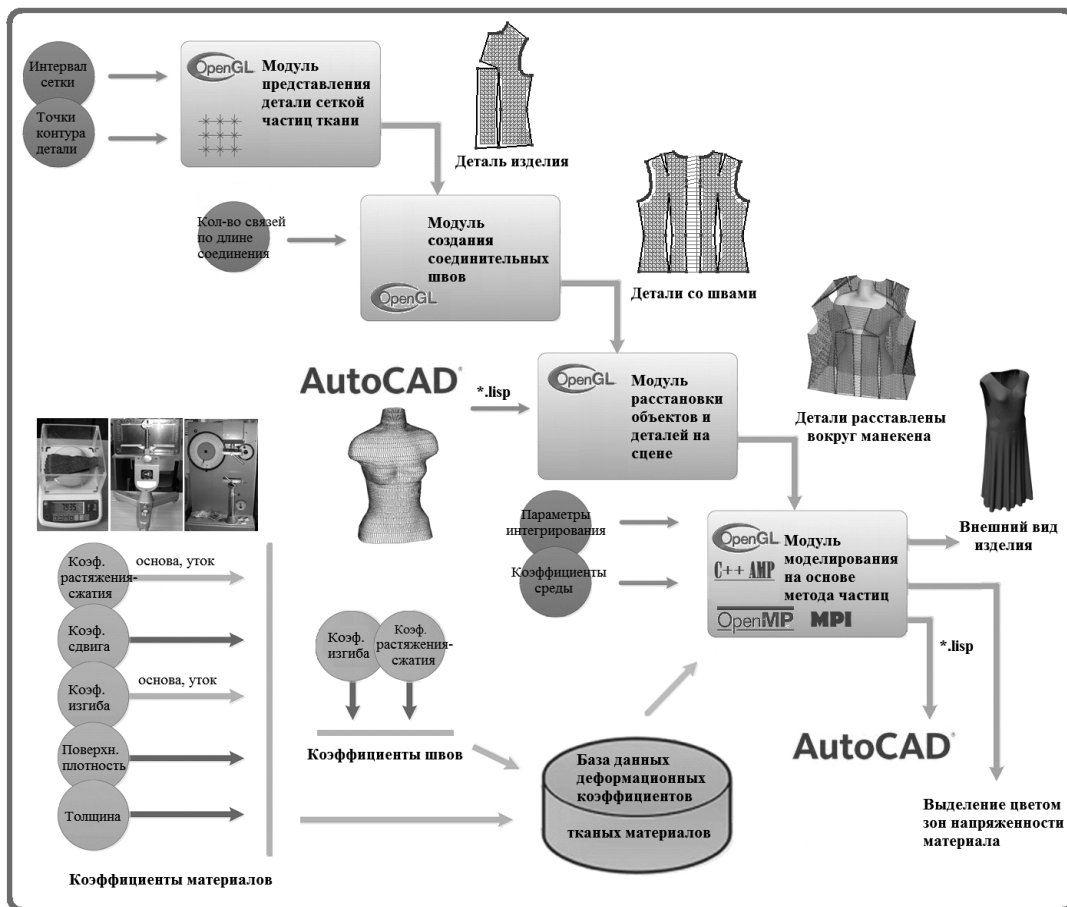


Рисунок 5 – Общая схема моделирования

2.1 Создание деталей изделия. Сначала происходит построение деталей изделия и определение местоположения основных точек выкройки, затем точки, составляющие контур детали, загружаются в первый модуль программы. При расположении контура относительно координатных осей должно учитываться направление нити основы в материале. Контур покрывается сеткой заданной пользователем величины, в узлах которой будут расположены частицы моделируемой ткани (рис. 6). Стоит отметить, что для получения реалистичных результатов сборки сетка ткани должна быть равномерной и иметь квадратную форму ячейки, чтобы действующие силы, возникающие в моделируемом изделии, были приближены по модулю и направлению к силам, возникающим в реальном материале при его деформациях.

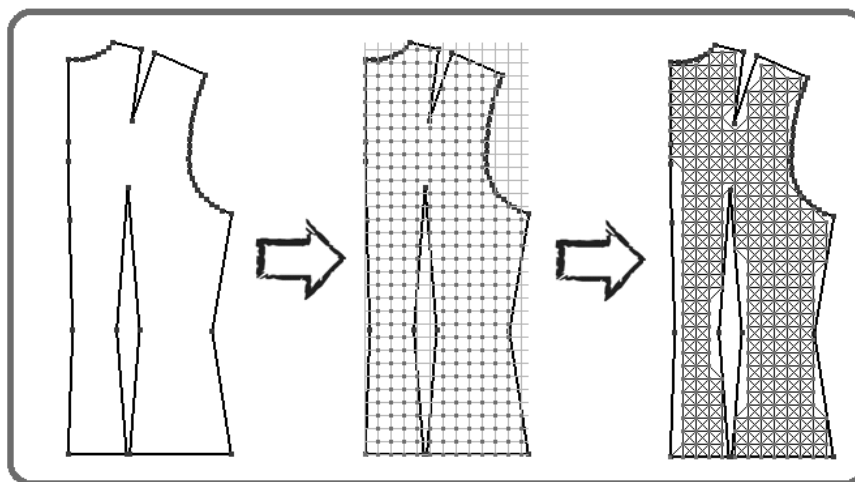


Рисунок 6 – Заполнение контура детали сеткой частиц

2.2 Создание швов. Так как равномерная сетка никогда не сможет в точности повторить контуры моделируемой детали, то в местах соединений следует представить деталь как совокупность контура и сетки ткани. Точки соединения размещаются на контуре детали на равном расстоянии друг от друга, при этом данные точки контура должны иметь жесткие связи с крайними частицами сетки ткани, чтобы в процессе сборки не изменялись размеры деталей. Следует отметить, при таком методе соединения размер ячеек сеток деталей может отличаться, как показано на рисунке 7.

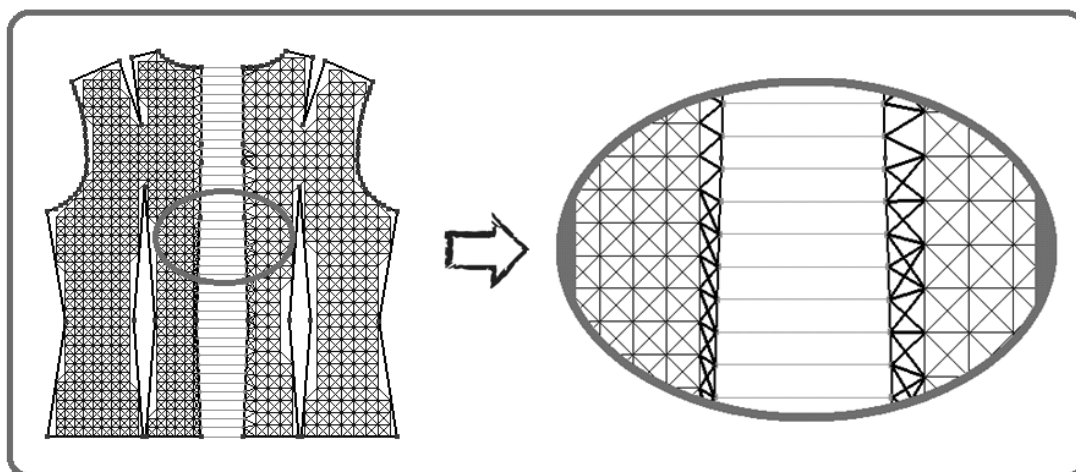


Рисунок 7 – Соединение двух деталей изделия

Для швов определяют коэффициенты растяжения-сжатия и изгиба. Определение производится таким же образом, как и для тканых материалов.

2.3 Моделирование сборки. Перед началом моделирования сборки изделия все швы должны быть созданы, а сами детали должны быть расставлены вокруг манекена.

В процессе моделирования учитываются коэффициенты среды, такие, как сила тяжести, сопротивление воздуха, сила трения об объект и прочие, что придает процессу большую схожесть с реальностью. После окончания моделирования пользователь может отобразить цветом зоны напряженности материала, что позволяет быстро оценить результаты и скорректировать размеры деталей изделия. Реалистичность получаемых изображений обеспечивается за счет использования графического API OpenGL. Для ускорения моделирования программа может использовать ресурсы многоядерных процессоров и графических видеоадаптеров.

После того, как полученные результаты оказываются удовлетворительными, пользователь может воспользоваться возможностью экспорта итоговых данных в AutoCAD, используя файл на языке LISP, а затем осуществить экспорт в 3ds Max для наложения текстур, придающих результатам большую реалистичность. Итоговый результат примерки базовой конструкции представлен на рисунке 8.

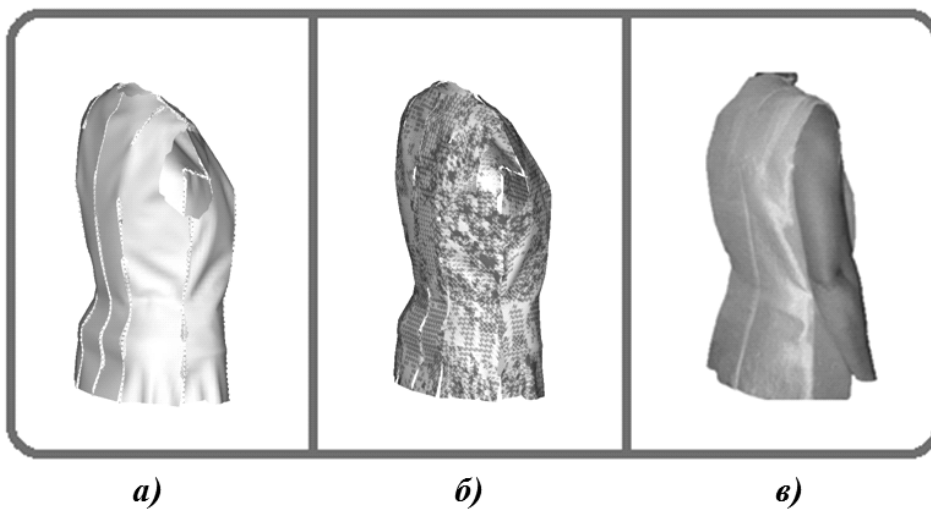


Рисунок 8 – Результаты примерки базовой конструкции:

а) – виртуальное изделие; б) – выделение цветом зон напряженности материала; в) – примерка на реальную модель

Для объективной оценки качества моделирования изделия предлагается использование следующих показателей, оказывающих существенное влияние на оценку качества посадки одежды и доступных для количественного измерения: отвесность положения краев бортов переда, боковых швов и среднего шва спинки; горизонтальность положения низа изделия; положение плечевых швов; соответствие месторасположения вытачек на выпуклость груди и наиболее выступающих точек груди. Значения допустимых отклонений для измеряемых показателей качества посадки одежды устанавливаются с использованием метода визуальной экспертной оценки. Согласно этому методу величины погрешностей (дефектов посадки) могут быть допущены только такие, которые незаметны глазу человека [7].

ЗАКЛЮЧЕНИЕ

Результатом проделанной работы является разработка универсальных алгоритмов и, как следствие, создание комплекса программных средств, включающего все этапы компьютерного моделирования сборки изделий из тканых материалов. Данный программный комплекс предоставляет возможность использования объектов, экспортированных из широко

известных графических сред, что предоставляет пользователю практически неограниченный выбор форм и размеров объектов.

Уникальность работы заключается в объединении методов экспериментального исследования деформационных свойств тканых материалов на конкретной измерительной аппаратуре и математического моделирования на компьютере с использованием полученных данных без каких-либо сложных преобразований. Еще одной новой идеей, показавшей хорошие результаты при сборке изделий из плоских заготовок, является метод создания швов, которые призваны соединять детали между собой, ориентируясь на контур детали.

Так как процесс моделирования ткани – это очень ресурсоемкая задача, то в зависимости от конфигурации компьютера программа может задействовать для выполнения вычислений как ресурсы многоядерного процессора, так и ресурсы графического ускорителя. Все это позволяет в значительной мере упростить работу исследователя по получению и обработке результатов моделирования.

Описанный программный комплекс может найти применение в легкой промышленности, дизайне, компьютерных играх и анимации, а также во многих других смежных отраслях.

СПИСОК ЛИТЕРАТУРЫ

1. Бузов Б.А., Алыменкова Н.Д. Материаловедение в производстве изделий легкой промышленности (швейное производство): учебник для студентов высших учебных заведений. – М.: Издательский центр «Академия», 2004. – 448 с.
2. Breen D.E., House D.H., Peters A.K. Cloth Modeling and Animation // Ltd. Natick, Massachusetts, 2000. – 344 p.
3. Фроловский В.Д. Избранные задачи геометрического проектирования. Параметризация сложных поверхностей. – Новосибирск: НГТУ, 2005. – 165 с.
4. Хокни Р., Иствуд Дж. Численное моделирование методом частиц: пер. с англ. – М.: Мир, 1987. – 640 с.
5. Ландовский В.В., Фроловский В.Д. Исследование методов интегрирования дифференциальных уравнений в задаче моделирования поведения ткани на основе метода частиц. – Сибирский журнал вычислительной математики, 2006. – Том 9. – С. 287-298.
6. Ландовская И.Е., Ландовский В.В., Фроловский В.Д. Моделирование изделий из тканых материалов с учетом их деформационных свойств // Современное состояние естественных и технических наук: сб. материалов XVII Международной научно-практической конференции. – М.: Издательство «Спутник+», 2014. – С. 76-81.
7. Коблякова Е.Б. Основы проектирования рациональных размеров и формы одежды. – М.: Легкая и пищевая промышленность, 1984. – 208 с.

Ландовская Ирина Евгеньевна

ФГБОУ ВПО «Новосибирский государственный технический университет», г. Новосибирск

Аспирант кафедры «Автоматизированные системы управления»

Тел.: 8 952 912 72 50

E-mail: nairy@rambler.ru

Фроловский Владимир Дмитриевич

ФГБОУ ВПО «Новосибирский государственный технический университет», г. Новосибирск

Доктор технических наук, профессор

Тел.: 8 (383) 346-11-00

E-mail: frolovskij@corp.nstu.ru

Ландовский Владимир Владимирович

ФГБОУ ВПО «Новосибирский государственный технический университет», г. Новосибирск

Кандидат технических наук

Тел.: 8 (383) 346-11-00

E-mail: landovskij@corp.nstu.ru

I.E. LANDOVSKAYA (*Post-graduate Student of the Department «Automated Control Systems»*)

V.D. FROLOVSKIY (*Doctor of Engineering Sciences, Professor*)

V.V. LANDOVSKIY (*Candidate of Engineering Sciences*)
Novosibirsk State Technical University, Novosibirsk

**DEFORMATION PROPERTIES INVESTIGATION
AND FABRIC SIMULATION AUTOMATIZATION**

The developed software environment is presented at this paper. This environment includes all phases of the computer simulation of the fabric product assembling. The methods of the experimental investigation of the fabric's deformation properties are considered. The use cases of the experimentally obtained data in the software are described. Visual comparison of the real product with its simulated prototype is provided.

Keywords: *fabric computer simulation; investigation of the fabric's deformation properties; software for computer simulation of the fabric product assembling; physical properties of the fabric; particle method; leapfrog scheme.*

BIBLIOGRAPHY (TRANSLITERATED)

1. Buzov B.A., Alymenkova N.D. Materialovedenie v proizvodstve izdelij legkoj promy'shlennosti (shvejnoe proizvodstvo): uchebnik dlya studentov vy'sshix uchebny'x zavedenij. – M.: Izdatel'skij centr «Akademiya», 2004. – 448 s.
2. Breen D.E., House D.H., Peters A.K. Cloth Modeling and Animation // Ltd. Natick, Massachusetts, 2000. – 344 p.
3. Frolovskij V.D. Izbranny'e zadachi geometricheskogo proektirovaniya. Parametrizaciya slozhny'x poverxnostej. – Novosibirsk: NGTU, 2005. – 165 s.
4. Xokni R., Istvud Dzh. Chislennoe modelirovanie metodom chastic: per. s angl. – M.: Mir, 1987. – 640 s.
5. Landovskij V.V., Frolovskij V.D. Issledovanie metodov integrirovaniya differencial'ny'x uravnenij v zadache modelirovaniya povedeniya tkani na osnove metoda chastic. – Sibirskij zhurnal vy'chislitel'noj matematiki, 2006. – Tom 9. – S. 287-298.
6. Landovskaya I.E., Landovskij V.V., Frolovskij V.D. Modelirovanie izdelij iz tkany'x materialov s uchedom ix deformacionny'x svojstv // Sovremennoe sostoyanie estestvenny'x i texnicheskix nauk: sb. materialov XVII Mezhdunarodnoj nauchno-prakticheskoy konferencii. – M.: Izdatel'stvo «Sputnik+», 2014. – S. 76-81.
7. Koblyakova E.B. Osnovy' proektirovaniya racional'ny'x razmerov i formy' odezhdy'. – M.: Legkaya i pishhevaya promy'shlennost', 1984. – 208 s.

ТЕХНОЛОГИЯ ОНТОЛОГИЧЕСКОЙ ИНТЕГРАЦИИ НА ОСНОВЕ АВТОМАТИЧЕСКОГО ФОРМИРОВАНИЯ И СРАВНЕНИЯ КОНТЕКСТОВ ПОНЯТИЙ РАЗНОРОДНЫХ ОНТОЛОГИЙ

В работе предлагается технология онтологической интеграции, основанная на автоматическом построении и сравнении контекстов понятий разнородных онтологий для оценки их семантической близости. Технология обеспечивает возможность устранения субъективности неформальных описаний элементов онтологий, отсутствие необходимости использования специализированных тезаурусов, а также ориентирована на применение для онтологий, описывающих практически любую предметную область или задачу. Технология реализует разработанные процедуры комплексного анализа корпуса текстов и алгоритмы формирования и сравнения контекстов онтологий.

Ключевые слова: онтология; интеграция; контекст; оценка семантической близости; аннотирование.

ВВЕДЕНИЕ

На сегодняшний день ввиду накопления огромного объема знаний в различных предметных областях проблема их представления и повторного использования является весьма актуальной. Одним из решений данной проблемы является использование онтологических моделей для получения формальной спецификации концептуализации. Под концептуализацией понимается согласованное видение предметной области группой специалистов (экспертов). Полученная в итоге онтологическая спецификация или онтология сочетает в себе как формализованные знания в виде системы аксиом, так и неформализованные – вербальные названия и определения терминов и отношений между ними. Эта особенность позволяет использовать ее как в задачах машинной обработки информации, так и для наглядного представления и оперирования знаниями экспертом предметной области.

Однако достаточно часто разработка онтологий ведется различными несогласованными группами экспертов, например, разными институтами, и возникает ситуация, при которой требуемую предметную область описывает не одна, а несколько семантически гетерогенных онтологий. В таких условиях неизбежно возникает задача интеграции онтологий. В общем случае интеграцию онтологий принято определять как процесс нахождения сходства двух онтологий A и B и, как результат, создание новой онтологии C , объединяющей и согласующей семантические представления исходных онтологий [1].

В ходе исследований разработана технология онтологической интеграции, основанная на автоматическом построении и сравнении контекстов понятий разнородных онтологий для оценки их семантической близости. Составление контекста осуществляется в результате проведения комплексного анализа корпуса текстов, описывающих ту же предметную область или задачу, что и обрабатываемая онтология.

Отдельное внимание в работе направлено на создание процедур синтаксического анализа текстов и разработку алгоритмов формирования и сравнения контекстов онтологий на основе результатов, полученных на этапе синтаксического анализа.

ПОСТАНОВКА ЗАДАЧИ. ПРЕИМУЩЕСТВА ПРЕДЛАГАЕМОГО ПОДХОДА

Разработка онтологии подразумевает создание спецификации в виде формального описания некоторой концептуализации, мыслимой разработчиком или их коллективом.

Частично такую концептуализацию удастся специфицировать явно в виде набора формальных аксиом. Однако некоторая часть формально не описывается, что может быть обусловлено целью разработки онтологии, а также ограничениями языка выбранной онтологической модели. Также следует заметить, что определение совокупности аксиом онтологии задает систему ограничений множеств интерпретации концептов и отношений, но при этом они остаются всего лишь абстракциями. Для их овеществления разработчик дает им осмысленные имена, а также производит их аннотирование. Тем самым некоторая (сравнительно большая) часть семантики представляется неформально. Вследствие этого при интеграции онтологий используются как формальные, так и неформальные методы для сопоставления семантики элементов [2]. Причем, как правило, формальные методы могут быть использованы лишь в дополнение к неформальным [3]. Именно поэтому развитие неформальных методов оценки семантической близости понятий позволит повысить корректность результата интеграции онтологий.

Неформальные методы оценки семантической близости состоят в выявлении различных сходств вербальных имен и определений элементов, данных разработчиком онтологии. Однако в этом случае на оценке близости понятий может сказаться субъективность определений какого-либо понятия, так как они могут быть верными во всех случаях, но по-разному сформулированы. В этом случае весьма вероятно получение негативной оценки близости при сопоставлении схожих понятий различных онтологий.

Некоторые методы сравнения также предполагают использование тезауруса, на понятиях которого определены различные лексические отношения (синонимия, гипонимия, омонимия), для выявления схожих слов в определениях или вербальных именах элементов онтологий. Однако не во всех случаях удастся найти тезаурус, соответствующий той предметной области или задаче, понятия которой описаны в онтологии. Использование же общих тезаурусов может приводить, либо к неверным оценкам в случае различных точек зрения на понятие онтологии и тезаурусе, либо к их отсутствию, так как в тезаурусе может вообще не найтись подходящего понятия. Последнее особенно вероятно для онтологий узких предметных областей или онтологий задач.

Метод сравнения понятий различных онтологий и реализуемая на его базе технология онтологической интеграции, предлагаемые в настоящей работе, основываются на том предположение, что контекст употребления одних и тех же терминов в текстовых источниках, описывающих одну предметную область, одинаковы. Сравнивая эти контексты, можно судить о семантической близости понятий, соответствующих данным терминам и определенных в разных онтологиях. Основные преимущества предлагаемого подхода заключаются в устранении субъективности неформальных описаний элементов онтологии, отсутствии необходимости использования специализированных тезаурусов, а также ориентированности на применение для онтологий, описывающих практически любую предметную область или задачу.

ОБЩАЯ ХАРАКТЕРИСТИКА ТЕХНОЛОГИИ

Реализация технологии осуществляется в два этапа: формирование контекстов понятий онтологии для более полного представления их интерпретаций и сравнение полученных контекстов для оценки семантической близости понятий.

Формирование контекстов рекомендуется проводить разработчику онтологии в качестве одной из заключительных фаз ее создания. При этом следует использовать некоторый корпус текстов, в которых с точки зрения разработчика онтологии используемые им понятия имеют правильное значение. Роль такого корпуса текстов могут играть наборы инструкций, рабочих документов, научных статей и т.п.

Сравнение контекстов производится уже непосредственно во время процесса интеграции онтологий, прошедших обработку на первом этапе, где оно, как правило,

выполняется наряду с другими методами для получения дополнительных оценок семантической близости понятий.

Под контекстом термина в данном случае будем понимать набор множеств лексем Q , семантически связанных с лексемой, обозначающей термин, в некотором корпусе текстов, относящимся к предметной области или задаче:

$$Q = (K_1, K_2, K_3, K_4, K_5), \quad (1)$$

где K_i – множества лексем; i определяет их значимость в описании интерпретации термина – чем оно больше, тем более значимым данное множество является. Сами множества состоят из лексем следующих видов:

- в K_1 попадают лексемы из тех предложений, в которых кроме лексемы, соответствующей понятию, нет других лексем-понятий онтологии, а найденная лексема-понятие употреблена не в значимой роли, например, подлежащее или дополнение;
- в K_2 попадают лексемы из тех предложений, в которых помимо искомой лексемы-понятия присутствуют и другие лексемы, соответствующие понятиям онтологии, при этом ни одна из них не является значимой;
- в K_3 попадают лексемы из предложений, в котором искомая лексема-понятие имеет значимую роль, при этом других лексем-понятий онтологии не найдено;
- в K_4 попадают лексемы из предложений, в которых искомая лексема-понятие имеет значимую роль, при этом в предложение присутствуют другие лексемы-понятия из онтологии;
- в K_5 попадают лексемы из предложений, в котором искомая лексема-понятие имеет значимую роль, а в качестве лексемы-сказуемого используется объектное отношение онтологии; либо сказуемое определено как непосредственно описывающее искомую лексему-понятие.

Алгоритм формирования контекстов понятий онтологии включает следующие основные этапы (рис. 1):

1. Обработка содержания текстов корпуса. Документы из текстового корпуса проходят многоуровневый языковой анализ, приводятся в нормализованную форму, снабженную полученной на этапе анализа информацией.

2. Индексация корпуса. При формировании контекста понятия требуется информации об употреблении соответствующей ему лексеме в тексте. Наряду с этим также необходимо располагать информацией об употреблении в тексте лексем, соответствующих и другим понятиям онтологии. Для этого перед составлением контекста строятся специальные структуры данных, которые содержат информацию об употреблении всех терминов онтологии в корпусе.

3. Формирование контекста. С помощью полученного на предыдущем шаге индекса производится включение лексем во множества контекстов понятий онтологии.

После завершения этапа формирования контекстов понятий онтологий может быть произведено их сравнение. Оно заключается в определения схожести по различным множествам контекстов понятий, входящих в интегрируемые онтологии. Схожесть различных множеств контекста в разной степени влияет на итоговую оценку.

Для оценки близости понятий онтологий по их контекстам используется формула Джаккарда, взвешенная по схожести различных множеств контекста:

$$S_{AB} = \frac{\sum_{n=1}^5 \left(i_n \cdot \frac{|K_n^A \cap K_n^B|}{|K_n^A \cup K_n^B|} \right)}{5}, \quad (2)$$

где A, B – контексты понятий двух онтологий; K_n^A, K_n^B – n -ое множество контекста; $0 \leq i_n \leq 1$ – коэффициенты значимости (определяются эмпирически).



Рисунок 1 – Логика работы алгоритма составления контекста

ПРОГРАММНАЯ РЕАЛИЗАЦИЯ ТЕХНОЛОГИИ

Программная реализация технологии была выполнена на языке Java в интегрированной среде разработки NetBeans версии 6.0 в виде программного модуля (Java-package), который в дальнейшем может быть использован в других программных продуктах. Диаграмма классов разработанного программного модуля представлена на рисунке 2.

При нормализации исходных текстов используется сторонний модуль морфологического анализа *Mcr*, описанный в работах [4, 5]. Отнесение лексем предложения к определенному множеству контекста сильно зависит от занимаемых ими синтаксических ролей, поэтому одним из важнейших этапов разработки системы стало создание синтаксического анализатора текста. При разработке синтаксического анализатора были использованы методики, описанные в работе [6].

Рассмотрим работу программного на этапе формирования контекста понятия, так как именно от его результатов зависит правильность и обоснованность дальнейшего сравнения контекстов.

Обработка текста начинается с лексического анализа, в ходе которого из исходного текста выделяются предложения, а из предложений выделяются токены, которые передаются для обработки морфологическому анализатору. Токенами в данном случае называется последовательность символов из входного потока символов, распознанная как единый объект.

Далее выполняется морфологический анализ, который заключается в поиске полученных токенов среди словоформ лексем, содержащихся в словаре анализатора, заданном изначально. В данном случае был использован грамматический словарь русского

языка А.А. Зализняка [7]. На выходе морфологический анализатор выдает набор лексем-омонимов, для словоформ которых найдено совпадение с рассматриваемым токеном. Данный набор далее будем называть юнитом.

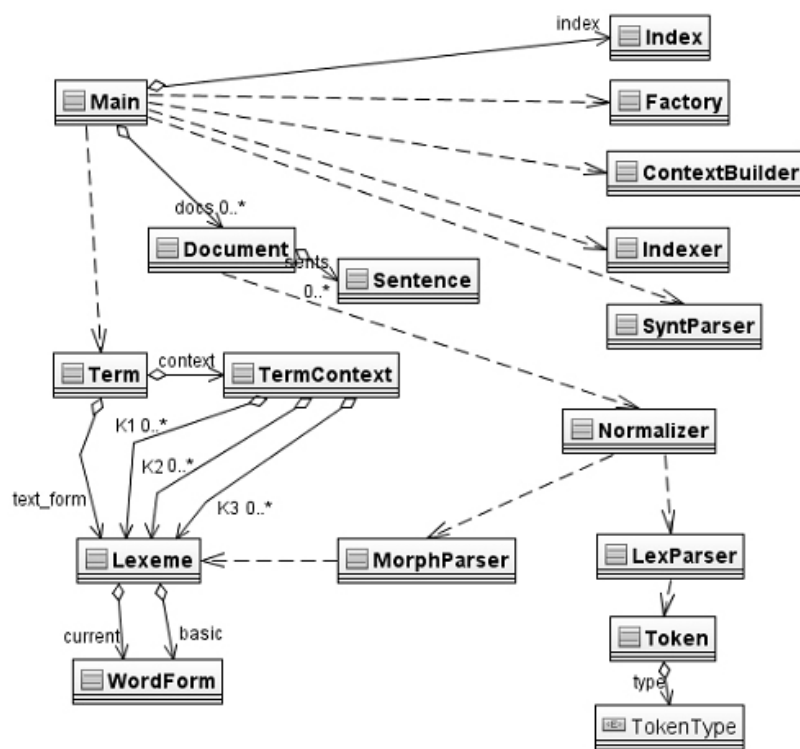


Рисунок 2 – Диаграмма классов программного модуля

Далее в ходе синтаксического анализа производится применение синтаксических правил с целью установления иерархии на юнитах, из которых состоит предложение, и формированием синтаксической группы. Под синтаксической группой будем понимать установленное n-арное отношение на юнитах. Группа может иметь вложенные подгруппы, одна из которых может быть главной. Главным юнитом считается главный юнит главной подгруппы, а главный юнит атомарной группы – единственный входящий в нее юнит. В начале анализа всем юнитам ставятся в соответствие атомарные группы – группы, состоящие из одного юнита. Далее к этому набору групп последовательно применяются синтаксические правила, в результате чего атомарные группы поглощаются более общими. Правила применяются к группам слева направо для нахождения групп, соответствующих условиям правила, перебирая при этом все комбинации лексем-омонимов в юнитах. Если найден набор групп, удовлетворяющих правилу, то они объединяются той группой, для которой выполнялся поиск. При этом из юнитов, входящих в найденные группы, удаляются все лексем-омонимы.

После обработки текста производится составление индекса. Под индексом понимается набор структур данных, служащий для быстрого доступа к информации.

Для составления контекста понятий необходимо, во-первых, для каждой соответствующей ему лексемы получать список предложений, где он используется, а во-вторых, для каждого предложения получать список понятий из онтологии, соответствующие лексем-омонимы которых в нем встречаются. Кроме того, в индекс заносится информация о синтаксической роли лексем-омонимов в предложениях. Именно эта информация используется для формирования контекстов понятий. Из вышесказанного следует, что логичным является создание прямого и обратного индекса. Прямой индекс позволяет получать информацию о

том, в каких предложениях какого документа встречается данный термин. Для его хранения используется система хэш-таблиц, содержащих ссылки на вложенные таблицы. То есть для каждого понятия создается таблица документов использования, для каждого документа – таблица предложений, а для каждого предложения – структура, содержащая адрес понятия-лексемы в предложении и данные синтаксического анализа. Обратный индекс организован похожим образом за исключением того, что цепочка хэш-таблиц строится в противоположном порядке.

Построение контекста понятия происходит следующим образом – перебираются все понятия онтологии, выбираются предложения, в которых встречаются соответствующие понятию лексемы. Затем из каждого предложения, согласно индексу, выбираются лексемы, формирующие контекст понятия с учетом других понятий онтологии, связанных с текущим.

ЗАКЛЮЧЕНИЕ

В работе предложена технология онтологической интеграции, основанная на применении разработанного метода формирования и сравнения контекстов понятий разнородных онтологий. Технология может быть эффективно использована в решении задачи оценки семантической близости онтологических описаний с целью их последующей интеграции. Отличие предлагаемого технологического решения от существующих неформальных методов установления близости понятий, предполагающих сравнение различными способами вербальных определений и имен понятий, данных разработчиком онтологии, заключается в том, что разработанная технология обеспечивает автоматическое формирование контекста термина, который можно рассматривать как его формальную аннотацию. Это позволяет разработчику онтологии избежать дополнительных трудозатрат на формулировку объемных аннотаций терминов, которые позволят более точно выявить сходство или различие понятий, а вместо этого ограничиться их короткими вариантами. Немаловажным является также отсутствие необходимости наличия при использовании технологии каких-либо тезаурусов терминов предметной области. Вместе с тем вследствие того, что контекст всегда формируется с помощью корпуса текстов по определенным правилам и независимо от разработчика онтологии, появляется возможность устранения возможной субъективности и узости аннотации понятия.

Разумеется, в полной мере проблему интеграции разнородных онтологий предложенная в работе технология не решает, так как алгоритмы составления контекста не безупречны и адекватны лишь настолько, насколько адекватны методы синтаксического разбора и анализа текста. Однако использование технологии позволяет повысить адекватность и обоснованность оценок близости понятий, полученных на основе их неформальной семантики, что в итоге положительно сказывается на интеграции онтологий в целом.

Основные области применения разработанной технологии:

- использование в программах автоматического интегрирования;
- упрощение работы эксперта, занимающегося вопросами интегрирования и составления онтологий, – составленный контекст помогает четче понять семантику термина.

*Работа выполнена при поддержке Российского фонда фундаментальных исследований.
Грант 15-07-03321.*

СПИСОК ЛИТЕРАТУРЫ

1. Бархатов А.В. Разработка и интеграция веб-ресурсов на основе их семантического представления в виде онтологий [Электронный ресурс]. – URL: <http://rrc.krasu.ru/node/1885>.

2. Маслобоев А.В., Ломов П.А. Подходы к интеграции разнородных онтологий на основе формальных и неформальных методов оценки их семантической близости // Вести высших учебных заведений Черноземья, 2010. – № 4(22). – С. 42-46.
3. Кудрявцев Д.В. Практические методы отображения и объединения онтологий // Труды 11-й Национальной конференции по искусственному интеллекту с международным участием (КИИ-08), семинар «Знания и Онтологии *ELSEWHERE*», 29 сент.-3 окт. 2008, г. Дубна, Россия. – М.: URSS, 2008. – Т. 3. – С.164-173.
4. Выдрин Д.В., Поляков В.Н. Реализация электронного словаря на основе n-грамм // Труды III Международной научно-практической конференции «Искусственный интеллект-2002». – Кацевелли: Изд-во «Институт проблем искусственного интеллекта», 2002. – Т. 2. – С.79-84.
5. Выдрин Д.В., Громов С.А., Поляков В.Н. Метод сравнения библиографических описаний, представленных в различных форматах // Обработка текста и когнитивные технологии № 9. VII Международная конференция, Варна. – М: Изд-во «Учеба», 2004. – С. 166-172.
6. Автоматическая обработка текста: синтаксический анализ [Электронный ресурс]. – URL: <http://aot.ru/docs/synan.html>.
7. Зализняк А.А. Грамматический словарь русского языка. – М.: АСТ-Пресс, 2009. – 720 с.

Маслобоев Андрей Владимирович

ФГБУН Институт информатики и математического моделирования технологических процессов Кольского научного центра РАН, г. Апатиты
Кандидат технических наук, доцент, старший научный сотрудник
Тел.: 8 953 306 57 23
E-mail: masloboev@iimm.ru

Ломов Павел Андреевич

ФГБУН Институт информатики и математического моделирования технологических процессов Кольского НЦ РАН, г. Апатиты
Кандидат технических наук, научный сотрудник
Тел.: 8 952 296 76 33
E-mail: lomov@iimm.ru

A.V. MASLOBOEV (*Candidate of Engineering Science, Associate Professor, Senior Researcher*)

P.A. LOMOV (*Candidate of Engineering Science, Research Associate*)
Institute of Informatics and Mathematical Modelling Process
Kola Scientific Center of Russian Academy of Sciences, Apatity'

**SEMANTIC INTEGRATION TECHNOLOGY BASED ON AUTOMATIC FORMATION
AND COMPARISON OF HETEROGENEOUS ONTOLOGY CONCEPT CONTEXTS**

In this paper a technology for ontological integration based on automatic formation and comparison of heterogeneous ontology concept contexts for their semantic adjacency assessment has been proposed. The technology provides possibilities for ontology elements informal descriptions subjectivity elimination, necessity absence of special-purpose thesaurus usage and application oriented for ontologies, which describe practically any knowledge domain or problem. The technology implements designed procedures for text frame complex analysis and algorithms for ontology contexts formation and comparison.

Keywords: *ontology; integration; context; semantic adjacency assessment; annotation.*

BIBLIOGRAPHY (TRANSLITERATED)

1. Barxatov A.V. Razrabotka i integraciya veb-resursov na osnove ix semanticheskogo predstavleniya v vide ontologij [E'lektronny'j resurs]. – URL: <http://rrc.krasu.ru/node/1885>.

2. Masloboev A.V., Lomov P.A. Podxody' k integracii raznorodny'x ontologij na osnove formal'ny'x i neformal'ny'x metodov ocenki ix semanticheskoy blizosti // Vesti vy'sshix uchebny'x zavedenij Chernozem'ya, 2010. – № 4(22). – S. 42-46.
3. Kudryavcev D.V. Prakticheskie metody' otobrazheniya i ob''edineniya ontologij // Trudy' 11-j Nacional'noj konferencii po iskusstvennomu intellektu s mezhdunarodny'm uchastiem (KII-08), seminar «Znaniya i Ontologii *ELSEWHERE*», 29 sent.-3 okt. 2008, g. Dubna, Rossiya. – M.: URSS, 2008. – T. 3. – S.164-173.
4. Vy'drin D.V., Polyakov V.N. Realizaciya e'lektronnoho slovarya na osnove n-gramm // Trudy' III Mezhdunarodnoj nauchno-prakticheskoy konferencii «Iskusstvenny'j intellekt-2002». – Kacevelli: Izd-vo «Institut problem iskusstvennogo intellekta», 2002. – T. 2. – S.79-84.
5. Vy'drin D.V., Gromov S.A., Polyakov V.N. Metod sravneniya bibliograficheskix opisaniy, predstavlenny'x v razlichny'x formatax // Obrabotka teksta i kognitivny'e texnologii № 9. VII Mezhdunarodnaya konferenciya, Varna. – M: Izd-vo «Ucheba», 2004. – S. 166-172.
6. Avtomaticheskaya obrabotka teksta: sintaksicheskij analiz [E'lektronny'j resurs]. – URL: <http://aot.ru/docs/synan.html>.
7. Zaliznyak A.A. Grammaticheskij slovar' russkogo yazyka. – M.: AST-Press, 2009. – 720 s.

О ВОЗМОЖНОСТЯХ ИСПОЛЬЗОВАНИЯ ДИСКРЕТНОЙ ФОРМУЛЫ АКАДЕМИКА КОТЕЛЬНИКОВА

В работе отмечено, что при описании нелинейностей посредством дискретных вариантов формулы отсчетов возникают особенности: 1) восстановленный сигнал (формальное описание) обладает существенно искаженными производными и поэтому его использование в дифференциальных уравнениях, описывающих динамику систем автоматического управления, может быть ограниченным; 2) при построении используются равноотстоящие узлы, что приводит к необходимости применения значительного числа узлов и в конечном итоге увеличивает количество и время вычислений в микропроцессорных элементах. В работе отмечены отдельные преимущества обобщенных выражений для дискретных вариантов формулы Котельникова.

Ключевые слова: модель; восстановление сигнала; управление; интерполяция; аппроксимация; теорема отсчетов; моделирование функций; дискретизация; частотный спектр; преобразование дискретных данных в аналоговую форму.

АКТУАЛЬНОСТЬ

Восстановление сигналов $f(x)$ проводится по-разному. Распространено использование формулы академика Котельникова [1] в современной трактовке [2]:

$$F(x) = \sum_{-\infty}^{\infty} f(k) \frac{\sin \pi(x-k)}{\pi(x-k)}. \quad (1)$$

К особенности применения (1) можно отнести условие того, что $f(x)$ должна быть функцией ограниченного спектра. В реальности сигналы ограничены во времени и поэтому могут не иметь ограниченного спектра. Кроме того, не просто установить, является ли реальный сигнал сигналом с ограниченным спектром. Ситуация несколько усложняется при необходимости проведения процессов восстановления сигналов с использованием микропроцессорной техники и дискретизации формулы (1):

$$F(x) = \sum_{k=-N}^{k=+N} f(k) \frac{\sin \pi(x-k)}{\pi(x-k)}, \quad (2)$$

где моменты k определяются массивом значений:

– на интервале $-N \leq x \leq +N$ с $(2N+1)$ -им отсчетом:

$$\left(\begin{matrix} f(-N) \\ x = -N \end{matrix} \right), \left(\begin{matrix} f(-N+1) \\ x = -N+1 \end{matrix} \right), \left(\begin{matrix} \dots \\ \dots \end{matrix} \right), \left(\begin{matrix} f(-1) \\ x = -1 \end{matrix} \right), \left(\begin{matrix} f(0) \\ x = 0 \end{matrix} \right), \left(\begin{matrix} f(1) \\ x = 1 \end{matrix} \right), \dots, \left(\begin{matrix} f(N-1) \\ x = N-1 \end{matrix} \right), \left(\begin{matrix} f(N) \\ x = N \end{matrix} \right); \quad (3)$$

– на интервале $-a \leq x \leq +a$ с $(2N+1)$ отсчетом:

$$\left(\begin{matrix} f(a \frac{-N}{N}) \\ x = a \frac{-N}{N} \end{matrix} \right), \left(\begin{matrix} f(a \frac{-N+1}{N}) \\ x = a \frac{-N+1}{N} \end{matrix} \right), \left(\begin{matrix} \dots \\ \dots \end{matrix} \right), \left(\begin{matrix} f(-a \cdot \frac{1}{N}) \\ x = -a \cdot \frac{1}{N} \end{matrix} \right), \left(\begin{matrix} f(0) \\ x = 0 \end{matrix} \right), \left(\begin{matrix} f(a \cdot \frac{1}{N}) \\ x = a \cdot \frac{1}{N} \end{matrix} \right), \dots, \left(\begin{matrix} f(a \cdot \frac{N-1}{N}) \\ x = a \cdot \frac{N-1}{N} \end{matrix} \right), \left(\begin{matrix} f(a \frac{N}{N}) \\ x = a \frac{N}{N} \end{matrix} \right). \quad (4)$$

Здесь основными неудобствами являются:

1) кривая (2) «извивается» вокруг исходного сигнала (функции) практически при любом числе слагаемых в этом соотношении (рис. 1).

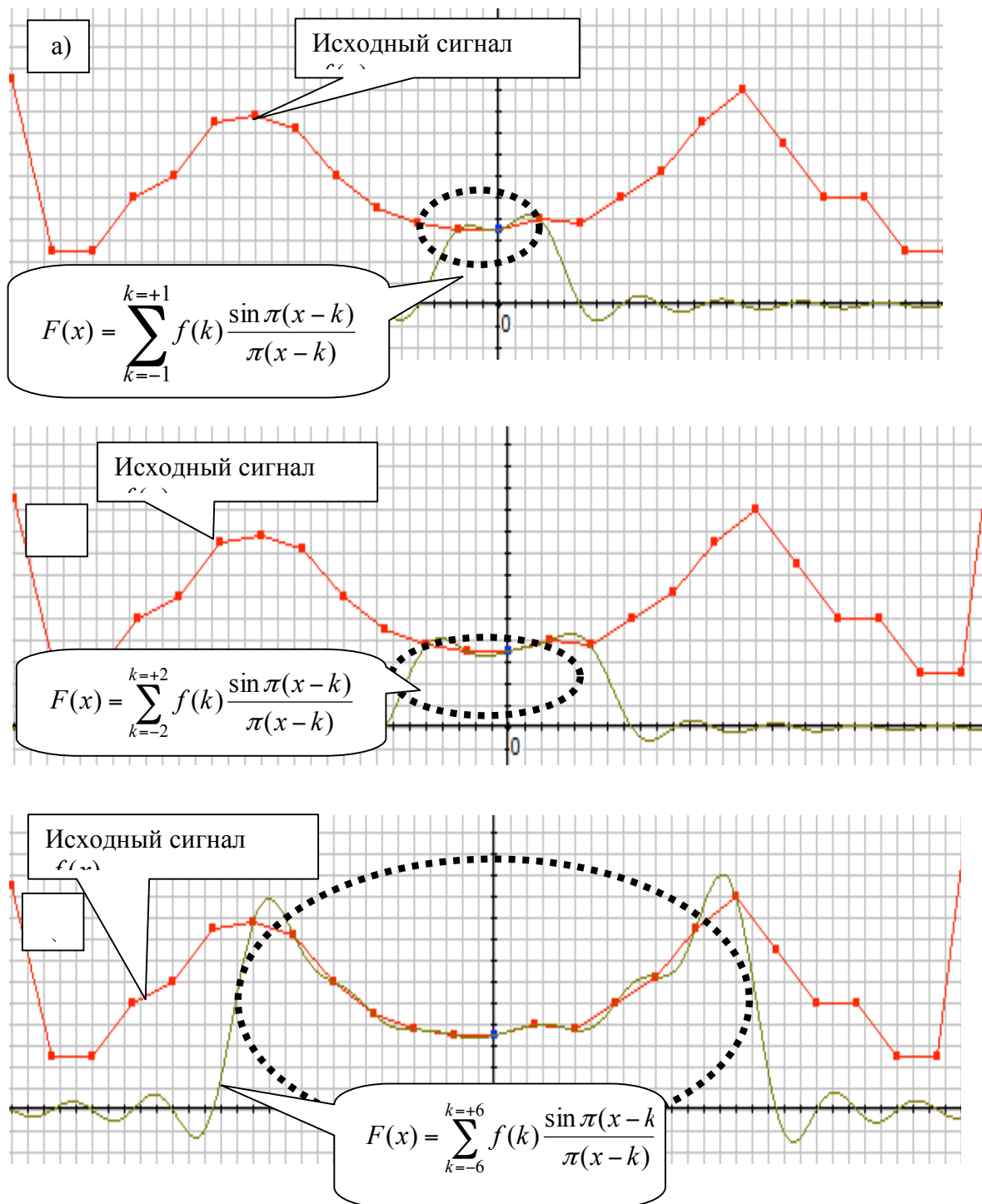


Рисунок 1 – Типичный пример «извивающегося» представления сигнала соотношением (2):

а) на первых трех «колокольчиках» $\frac{\sin \pi(x-k)}{\pi(x-k)}$; б) на первых пяти «колокольчиках»

$\frac{\sin \pi(x-k)}{\pi(x-k)}$; в) на первых тринадцати «колокольчиках» $\frac{\sin \pi(x-k)}{\pi(x-k)}$ (в границах

пунктирных овалов)

Такой восстановленный сигнал обладает существенно искаженными производными и уже только поэтому его использование в дифференциальных уравнениях, описывающих динамику систем автоматического управления (САУ), должно быть крайне осторожным или даже неприменимым;

2) при построении используются равноотстоящие узлы (3) или (4), что имеет, во-первых, риски отсутствия требуемых значений в массивах значений исходной кривой; во-вторых, приводит к потребности применения значительного числа узлов (рис. 2), что в конечном итоге увеличивает количество и время вычислений, погрешность восстановления сигнала и, в-третьих, ведет к нерегулируемой окончательной погрешности приближения.

В связи с этим становятся актуальными вопросы некоторой трансформации формулы (2) для преодоления отмеченных ограничений.

В работе предложены мероприятия по снятию условия использования равноотстоящих узлов, мероприятия по применению значительно меньшего числа узлов и намечены возможности регулирования погрешности приближения (восстановления).

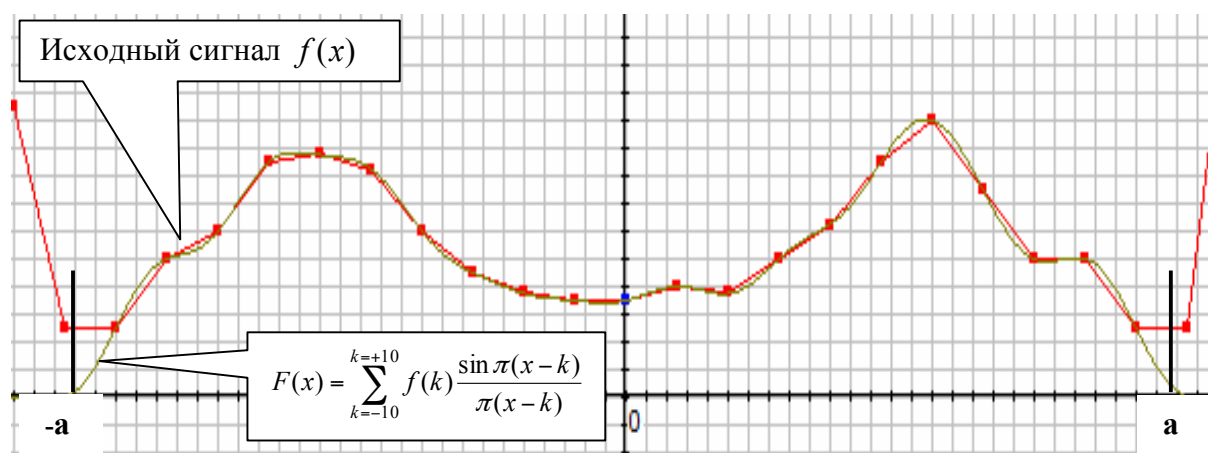


Рисунок 2 – Восстановление исходного сигнала 21-им «колокольчиком» $\frac{\sin \pi(x-k)}{\pi(x-k)}$ формулы (2) на интервале $[-a, +a]$ по числу отсчетов, равного 21 ($2N + 1$ и в данном случае $N = 10$)

ТРАНСФОРМАЦИЯ ФОРМУЛЫ

Регистрация сигнала с датчиков или в датчиках структур управления может проводиться по-разному: а) в моменты изменения состояния объекта управления (субстанционного, энергетического, структурно-информационного или их определенных композиций) или его компонентов для формирования подходящих управляющих воздействий; б) непрерывно и независимо от изменения или «неизменения» состояния объекта управления для создания целостной картины функционирования всех элементов системы; в) в моменты времени в зависимости от некоторого вычисленного обобщенного функционала, обусловленного определенной комбинацией характеристик или параметров эффективности системы управления в целом и т.п.

Для моделирования структур управления существенно важным является не столько то, что отмечена какая-то конкретная последовательность изменения событий, состояний или ситуаций, сколько возможность задания любой последовательности моментов времени, в которые сигнал может быть зафиксирован, измерен, сохранен и воспроизведен. В этом плане можно отметить некоторые интуитивные соображения.

Во-первых, роль аналого-цифровых преобразователей (АЦП) не должна сводиться к приборам, задающим отсчеты для восстановления сигналов. Их основное предназначение –

это представление исходного сигнала множеством точек (измеренных значений регулируемой величины), из которых и будут выбраны требуемые точки для подходящего восстановления сигналов.

Во-вторых, представление (2), как и представление (1), – это математическая модель сигнала; особенности математической модели должны напоминать (быть похожими, подобными, идентичными) особенности восстанавливаемого сигнала.

В-третьих, безусловно, общая (суммарная) «сложность» сигналов

$$\left\{ \frac{\sin \pi(x-k)}{\pi(x-k)} \right\}, \quad (5)$$

определяющих структуру формулы, описывающей исходную кривую, должна быть адекватна «сложности» исходного сигнала и не зависеть от используемого числа точек отсчетов.

В-четвертых, чем больше используется узлов, то есть колокольчиков (5) в разложении (2), тем больше извивается восстановленный сигнал вдоль исходной кривой, тем больше у такого восстановленного сигнала существенно искаженных производных и тем больше будет нежелательных особенностей у интегральных кривых дифференциальных уравнений, которыми описывается динамика систем автоматического управления (САУ).

Исходя из сказанного, можно предположить, что желаемые и нежелательные характерные черты формуле Котельникова придает количество используемых гармоник (5), которые увязаны с отсчетами (3) или (4), то есть с наличием в формуле (2) связи коэффициентов $F(k)$ с гармониками $\frac{\sin \pi(x-k)}{\pi(x-k)}$. Чтобы избавиться от этой обусловленности ориентируемся на логику преобразований Котельникова [1].

Для любой $F(x)$, удовлетворяющей условиям Дирихле (конечное число максимумов, минимумов и точек разрыва первого рода на любом конечном отрезке) и интегрируемости в пределах $(-\infty, +\infty)$, справедливо представление рядом:

$$\begin{aligned} F(t) &= \sum_{n=-\infty}^{\infty} f\left(\frac{n}{2f_{\Gamma P}}\right) \frac{\sin(2\pi f_{\Gamma P}t - n\pi)}{2\pi f_{\Gamma P}t - n\pi} = \sum_{n=-\infty}^{\infty} f\left(\frac{n}{2f_{\Gamma P}}\right) \frac{\sin(2\pi f_{\Gamma P}t - n\pi)}{\pi(2f_{\Gamma P}t - n)} = \\ &= \sum_{n=-\infty}^{\infty} f\left(\frac{n}{2f_{\Gamma P}}\right) \frac{\sin 2\pi f_{\Gamma P}\left(t - \frac{n}{2f_{\Gamma P}}\right)}{2\pi f_{\Gamma P}\left(t - \frac{n}{2f_{\Gamma P}}\right)} = \sum_{n=-\infty}^{\infty} \frac{f\left(\frac{n}{2f_{\Gamma P}}\right)}{\omega_{\Gamma P}} \cdot \frac{\sin \omega_{\Gamma P}\left(t - \frac{n}{2f_{\Gamma P}}\right)}{\left(t - \frac{n}{2f_{\Gamma P}}\right)}, \end{aligned} \quad (6)$$

где n – целое; $f(t)$ – функция, состоящая из частот от 0 до $f_{\Gamma P}$; $\omega_{\Gamma P} = 2\pi f_{\Gamma P}$.

Заменяя t на x и полагая $x_n = \frac{n}{2f_{\Gamma P}}$, соотношение (6) формально представляется в виде:

$$F(x) = \sum_{n=-\infty}^{\infty} f(x_n) \cdot \frac{\sin \omega_{\Gamma P}(x - x_n)}{\omega_{\Gamma P}(x - x_n)}. \quad (7)$$

Вид соотношения (7) наглядно указывает на два варианта избавления от связи $F(k)$ с гармониками:

1) осуществить переход:

$$\begin{aligned}
 F(x) &= \sum_{n=-\infty}^{\infty} f(x_n) \cdot \frac{\sin \omega_{\text{Гр}}(x-x_n)}{\omega_{\text{Гр}}(x-x_n)} \Rightarrow F(x) = \sum_{n=-\infty}^{\infty} f(x_n) \cdot \frac{\sin \omega_n(x-x_n)}{\omega_n(x-x_n)} \Rightarrow \\
 \Rightarrow F(x) &= \sum_{n=-\frac{N}{2}}^{\frac{N}{2}} f(x_n) \cdot \frac{\sin \omega_n(x-x_n)}{\omega_n(x-x_n)} \Rightarrow F(x) = \sum_{n=0}^N f(x_n) \cdot \frac{\sin \omega_n(x-x_n)}{\omega_n(x-x_n)}; \quad (8)
 \end{aligned}$$

2) осуществить переход от (8) к форме первой теоремы академика Котельникова [1, С. 763]:

$$F(x) = \sum_{n=0}^N f(x_n) \cdot \frac{\sin \omega_n(x-x_n)}{\omega_n(x-x_n)} \Rightarrow F(x) = \sum_{n=0}^N D_n \cdot \frac{\sin \omega_n(x-x_n)}{\omega_n(x-x_n)}. \quad (9)$$

Имеем две формулы, которые можно испытать на приемлемость:

$$F(x) = \sum_{i=0}^n f(x_i) \cdot \frac{\sin \omega_i(x-x_i)}{\omega_i(x-x_i)}, \quad (10)$$

$$F(x) = \sum_{i=0}^n D_i \cdot \frac{\sin \omega_i(x-x_i)}{\omega_i(x-x_i)}, \quad (11)$$

в которых n – это не все точки, полученные с АЦП, и не все точки, являющиеся отсчетами формулы (2).

Формируя соотношения (10) и (11), можно принципиально исходить из условия:

$$n \ll 2N + 1.$$

По самому обобщению (11) условие использования равноотстоящих узлов (3) или (4) становится необязательным, как и то, что необязательно использовать все точки исходного сигнала. Для определенности вариант (10) будем именовать аналогом функции Котельникова, а (11) – свободной функцией (Котельникова).

СТРУКТУРА ПРОГРАММЫ МОДЕЛИРОВАНИЯ

Для экспериментирования с формулами (10) и (11) предложенная программа моделирования имеет тривиальную структуру (рис. 3), у которой головным окном является панель ввода (рис. 4).

Исходная функция может быть задана двумя способами – по точкам и с помощью формулы, моделирующей исходную нелинейность. При этом координаты точек могут вводиться вручную или считываться автоматически при щелчке на системе координат. Координаты заносятся в массивы. При этом в массиве *tpoints* хранятся «условные» координаты, а в массиве *rpoints* – «фактические», т.е. те, по которым программа рисует график. Это связано с тем, что поле, в котором отображается график, имеет свою систему отсчета. Кроме того, используется функция *sort*, которая сортирует координаты точек в обоих массивах в порядке возрастания x (рис. 5).

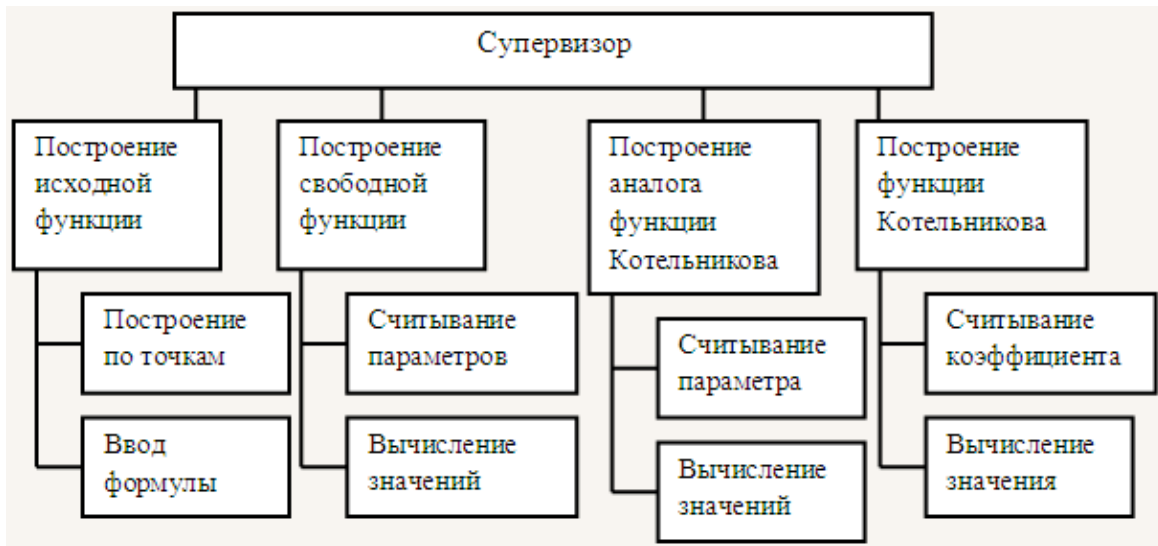


Рисунок 3 – Инструментарий программного эксперимента

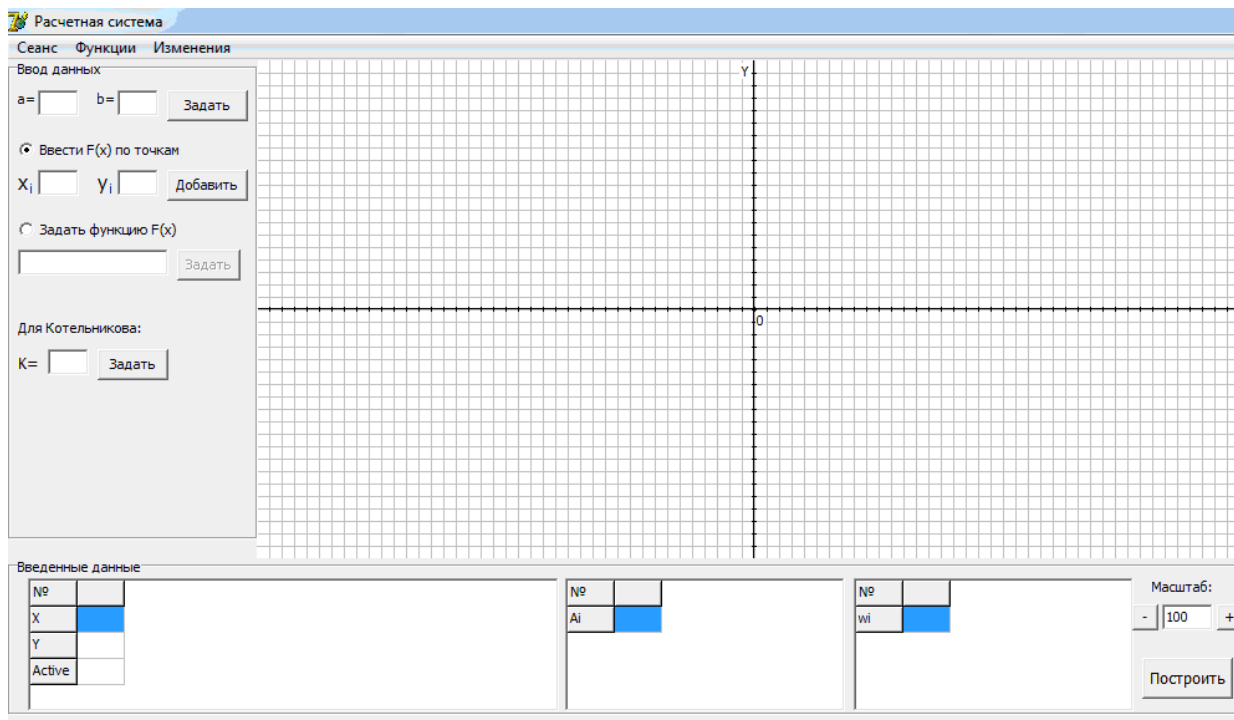


Рисунок 4 – Головное окно супервизора

```

    If(strtfloat(edit1.text)>=((a-400)/mash))and(strtfloat(edit1.text)<=((b-400)/mash))
then begin
    c:=form1.StringGrid1.ColCount-1;
    f:=true;
    For i:=1 to c-1 do
    If mpoints[i].x=strtfloat(edit1.text) then f:=false;
    If f then begin
    mpoints[c].x:=strtfloat(edit1.text);
    mpoints[c].y:=strtfloat(edit2.Text);
    if checkbox1.Checked=true then mpoints[c].flag:=true else mpoints[c].flag:=false;
    rpoints[c].x:=400+mpoints[c].x*mash;
    rpoints[c].y:=200-mpoints[c].y*mash;
    rpoints[c].flag:=mpoints[c].flag;
    sort;
    form1.StringGrid1.ColCount:=form1.StringGrid1.ColCount+1;
    edit1.Text:="";
    edit2.text:="";
    end;
    end
    else showmessage('Координаты выходят за границы диапазона');

```

Рисунок 5 – Задание исходной функции

При задании функции по формуле координаты x заносятся в массивы автоматически, а для вычисления координаты y (рис. 6) используется функция *latscalculate* из прикрепленного модуля *MyLib* (рисунок 7).

При построении свободной функции сначала получают значения параметров для построения на панели ввода, появляющейся автоматически при выборе точки (рис. 8).

```

str:=Edit7.Text;
c:=0;
For i:=(trunc(a)div mash)+1 to (trunc(b)div mash)+1 do begin
inc(c);
mpoints[c].x:=i-41;
rpoints[c].x:=mpoints[c].x*mash+400;
mpoints[c].y:=fun(mpoints[c].x,mpoints[c].x,str);
rpoints[c].y:=200-mpoints[c].y*mash;
rpoints[c].flag:=false;
mpoints[c].flag:=false;

```

Рисунок 6 – Задание функции по формуле

```

Function fun(x,y:extended; str:string):extended;
var Func:boolean;
Begin
result:=latscalculate(x,y,str,func);
end;

```

Рисунок 7 – Использование функции *latscalculate*

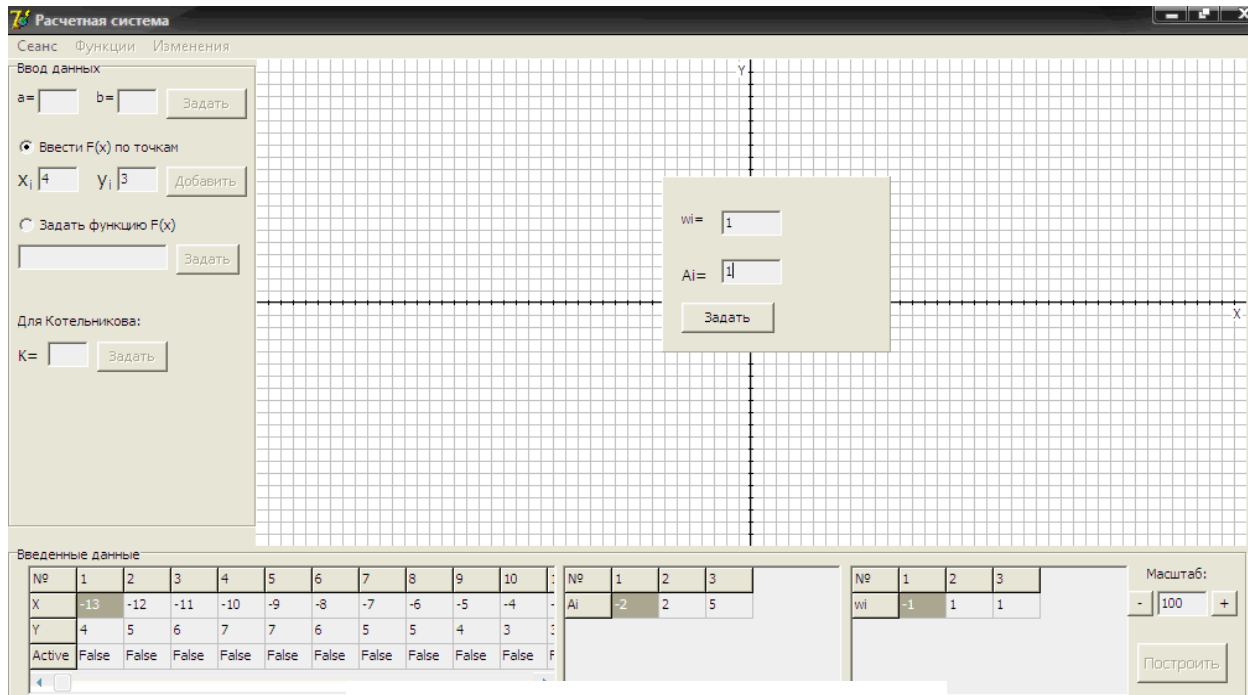


Рисунок 8 – Панель ввода

При нажатии кнопки «Задать» значения параметров помещаются в специальный массив $rabmn$: $\langle rabmn[aind].wi:=strtofloat(Edit12.Text); rabmn[aind].ai:=strtofloat(Edit13.Text); \rangle$. В этот же массив заносятся вычисленные координаты выбранной точки с использованием *getsin* (рис. 9).

Аналог функции Котельникова строится с использованием коэффициентов ω_i , которые считываются из панели ввода. Для вычисления значения функции Котельникова также используется отдельная функция (рис. 10). Для ее построения необходимо ввести значение коэффициента k в соответствующее поле (рис. 8, 11). После нажатия кнопки «Задать» будет считан коэффициент и выведена информация о количестве точек N , на которых будет строиться функция ($N=2*k+1$):

$\langle NCot:=strtoint(edit6.Text); Label8.Caption:='N='+inttostr(NCot*2+1); \rangle$.

Значения функции вычисляются обычным образом (рис. 12).

```
function getsin(X:extended):extended;
var i:integer;
    a:array[1..5000] of extended;
    f:extended;
begin
    result:=0;
    for i:=1 to aind do
        If X<>rabmn[i].x then begin
            f:=rabmn[i].wi*(X-rabmn[i].x);
            result:=result+rabmn[i].ai*sin(f)/f;
        end;
    end;
end;
```

Рисунок 9 – Использование функция *getsin* для вычисления значения функции

```

function getanalog(X:extended):extended;
var i:integer;
    f:extended;
begin
    result:=0;
    for i:=1 to aind do
        If X<>rabmn[i].x then begin f:=rabmn[i].wi*(X-rabmn[i].x);
            result:=result+rabmn[i].y*sin(f)/f;
        end; end;
end;
    
```

Рисунок 10 – Вычисление значения функции Котельникова

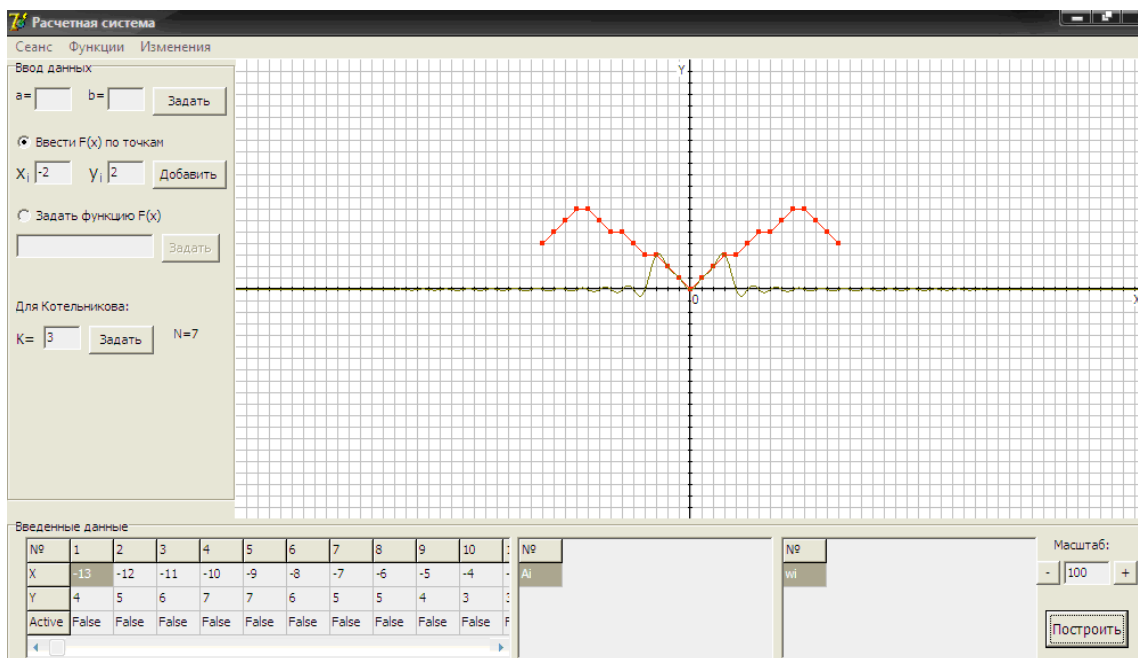


Рисунок 11 – Моделирование функций

```

function getcot(X:extended):extended;
var i,j:integer;
    f,y:extended;
begin
    result:=0;
    for i:=-NCot to NCot do
        If X<>i then begin
            f:=pi*(X-i);
            For j:=1 to c do If mpoints[j].x=i then y:=mpoints[j].y;
            result:=result+y*sin(f)/f;
        end;
    end;
end;
    
```

Рисунок 12 – Вычисление значения функции

ПРОГРАММНЫЕ ЭКСПЕРИМЕНТЫ

Несмотря на то, что описанный программный инструмент позволяет продуктивно работать с обеими формулами (10, 11), будем моделировать общий вид (11). Для этого используем подход интерактивной аппроксимации [3], когда качество аппроксимации

оценивается не формальной процедурой, а непосредственно человеком-исследователем, и когда последующий выбор узлов обуславливается предыдущими действиями и направляется исключительно человеком, участвующим в процедуре аппроксимации. Другими словами, человек, оценивая «гладкость» кривой, определяет последующие действия по продолжению конструирования свободной функции (11), когда потребность накапливания гармоник (колокольчиков)

$$D_i \cdot \frac{\sin \omega_i(x - x_i)}{\omega_i(x - x_i)} \quad (12)$$

определяется исключительно человеком-исследователем на каждом шаге восстановления (аппроксимации) посредством задания конкретных ω_i, x_i, D_i ($i = 0, 1, \dots, n$ – количество узлов).

Действительно, на массиве точек исходного сигнала (исходной кривой на рис. 2) можно получить не 21 (рис. 2) и 13 (рис. 1 (в)) «колокольчиков», как в случае с представлением (2), а значительно меньшее число гармоник (12) для выражения (11), например, в последовательности рисунков 13-15 с окончательной формулой описания на интервале [-6, +6]:

$$f(x) = 4,2 \cdot \frac{\sin(0,85(x + 6))}{0,85(x + 6)} + 4,4 \cdot \frac{\sin(0,85(x - 6))}{0,85(x - 6)} + 3 \cdot \frac{\sin(0,95x)}{0,95x},$$

состоящей из 3-х гармоник.

Регулирование погрешности очевидно добавлением к (11) новых гармоник (12) в интервалах наибольшего отклонения от исходного сигнала (по результатам анализа функции ошибки).

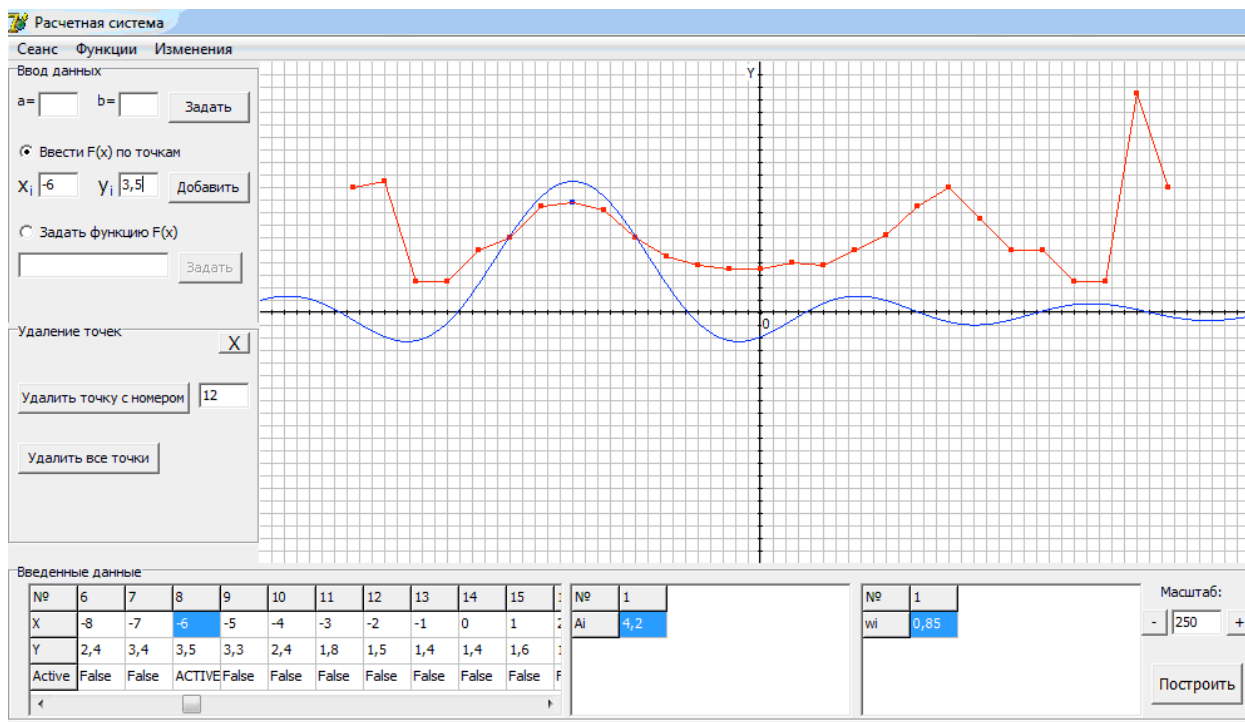


Рисунок 13 – Формирование аналитического описания (11) заданием первой гармоники (12) (в точке №8) с параметрами: $x_1 = -6$, $\omega_1 = 0,85$, $D_1 = 4,2$ и выражением (11) для первого шага

интерактивной аппроксимации в виде $f(x) = 4,2 \cdot \frac{\sin(0,85(x + 6))}{0,85(x + 6)}$

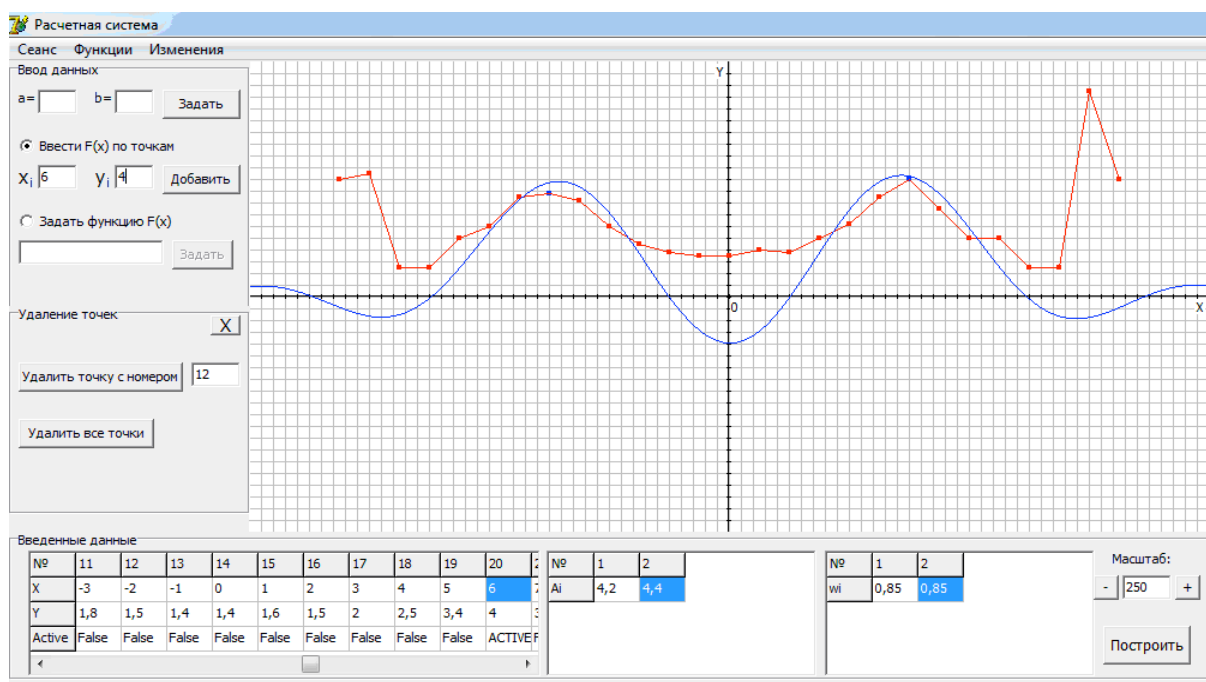


Рисунок 14 – Формирование аналитического описания (11) заданием второй гармоники (12) (в точке №20) с параметрами: $x_2 = 6$, $\omega_2 = 0,85$, $D_2 = 4,4$ и выражением (11) для второго шага в виде:

$$f(x) = 4,2 \cdot \frac{\sin(0,85(x + 6))}{0,85(x + 6)} + 4,4 \cdot \frac{\sin(0,85(x - 6))}{0,85(x - 6)}$$

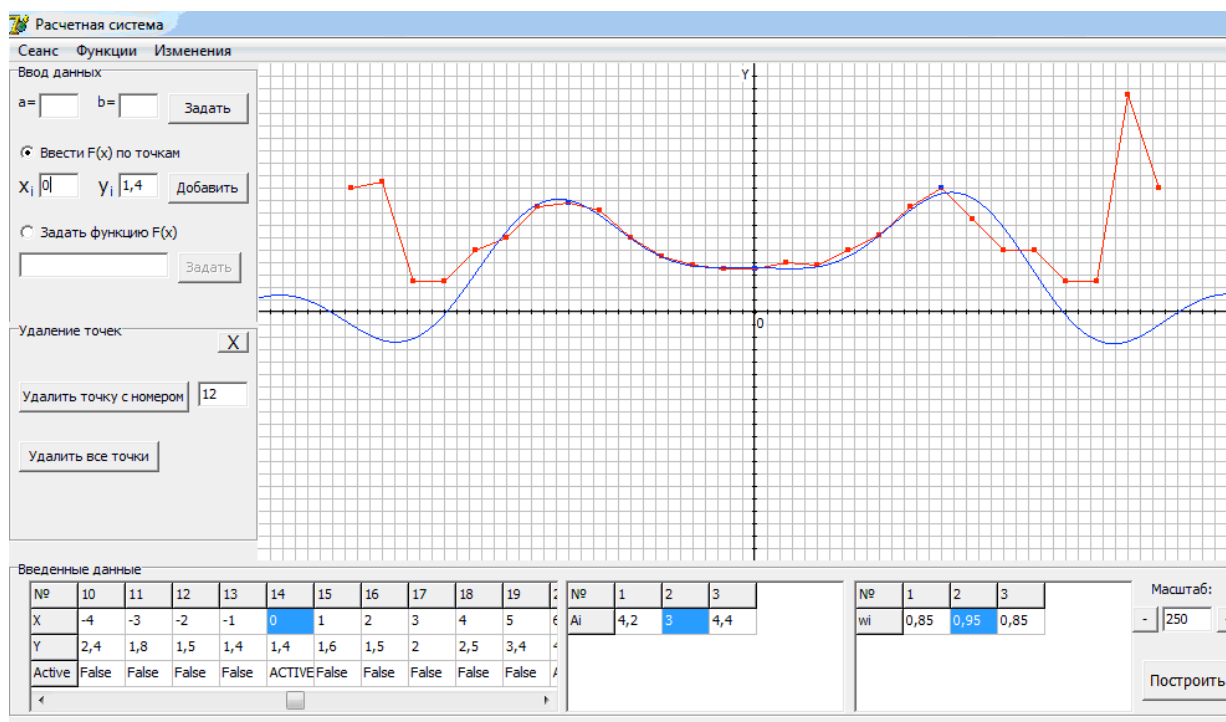


Рисунок 15 – Формирование аналитического описания (11) заданием третьей гармоники (12) (в точке №14) с параметрами: $x_3 = 0$, $\omega_3 = 0,95$, $D_3 = 3$ и выражением (11) для третьего шага в виде:

$$f(x) = 4,2 \cdot \frac{\sin(0,85(x + 6))}{0,85(x + 6)} + 4,4 \cdot \frac{\sin(0,85(x - 6))}{0,85(x - 6)} + 3 \cdot \frac{\sin(0,95x)}{0,95x}$$

РЕЗУЛЬТАТЫ

1. Отмечены некоторые неудобства использования «дискретной формулы» академика Котельникова для описания непрерывных функций, предложена к использованию ее трансформация и отмечены возможные ее преимущества при построении аналитических выражений

2. Предложена структура программы моделирования позволяет регулировать погрешность добавлением новых гармоник (12).

3. Сокращение числа используемых узлов может вести к уменьшению возможного искажения производных.

4. Дискретные варианты формулы акад. В.А. Котельникова могут оказаться полезными при моделировании динамики систем управления.

Работа выполнена при поддержке «Госунiversитет – УНПК» по теме «Разработка программной системы поддержки процесса управления в предаварийных состояниях для восстановления нормальной работы», приказ №7-н/26 от 23.10.13.

СПИСОК ЛИТЕРАТУРЫ

1. Котельников В.А. О пропускной способности «эфира» и проволоки в электросвязи: материалы к Первому Всесоюзному съезду по вопросам технической реконструкции дела связи и развития слаботочной промышленности. Сборник Всесоюзного энергетического комитета по радиосекции. – М.: Управление связи РККА, 1933 // Успехи физических наук, 2006. – № 7. – С. 762-770.
2. Хемминг Р.В. Численные методы. – М.: Наука, 1972. – 400 с.
3. Раков В.И. Моделирование и информационные системы в интерактивной аппроксимации: учебное пособие. Часть 1. Методика описания. – М.: Изд. РАЕ, 2012. – 112 с.

Раков Владимир Иванович

ФГБОУ ВПО «Госунiversитет – УНПК», г. Орел

Доктор технических наук, профессор

Тел.: 8 906 569 20 20 (кафедра)

E-mail: rakov2010vi@mail.ru

Чернобровкина Юлия Владимировна

ФГБОУ ВПО «Госунiversитет – УНПК», г. Орел

Студент кафедры «Информационные системы»

Тел.: 8 906 569 20-20 (кафедра)

E-mail: nikeli2009@yandex.ru

I.V. RAKOV (*Doctor of Engineering Sciences, Professor*)

Yu.V. CHERNOBROVKINA (*Student of the Department «Information Systems»
State University – ESPC, Orel*)

ABOUT THE POSSIBILITIES OF USING THE DISCRETE FORMULA OF ACADEMICIAN KOTELNIKOV

In work it is noted that at the description of non linearities by means of discrete options of a formula of counting there are features: 1) approximating function possesses significantly distorted derivatives and therefore its use in the differential equations describing dynamics of systems of automatic control, can be limited; 2) at construction equidistant knots are used, it results in need of application of considerable number of knots, at last it increases quantity and time of calculations in microcontrollers. In work the new option of a discrete formula is offered and separate advantages are noted it.

Keywords: *model; signal reshaping; control; interpolation; approximation; theorem of counting; modeling of functions; discretization; frequency content; desampling.*

BIBLIOGRAPHY (TRANSLITERATED)

1. Kotel'nikov V.A. O propusknoj sposobnosti «e'fira» i provoloki v e'lektrosvyazi: materialy' k Pervomu Vsesoyuznomu s'ezdu po voprosam texnicheskoj rekonstrukcii dela svyazi i razvitiya slabotochnoj promy'shlennosti. Sbornik Vsesoyuznogo e'nergeticheskogo komiteta po radiosekcii. – M.: Upravlenie svyazi RKKa, 1933 // Uspexi fizicheskix nauk, 2006. – № 7. – S. 762-770.
2. Xemming R.V. Chislenny'e metody'. – M.: Nauka, 1972. – 400 s.
3. Rakov V.I. Modelirovanie i informacionny'e sistemy' v interaktivnoj approksimacii: uchebnoe posobie. Chast' 1. Metodika opisaniya. – M.: Izd. RAE, 2012. – 112 s.

УДК 681.5:004.4

А.Ф. АНТИПИН

К ВОПРОСУ О СЕМАНТИЧЕСКОМ АНАЛИЗЕ ПРОГРАММ АВТОМАТИЗИРОВАННЫХ СИСТЕМ УПРАВЛЕНИЯ

В статье рассматривается авторский способ анализа программного кода автоматизированных систем управления, основанный на концепции семантического анализа структуры многомерного интервально-логического регулятора, и условия его применимости. Данный способ позволяет выявлять смысловые ошибки в коде до начала тестирования и отладки.

Ключевые слова: многомерный интервально-логический регулятор; система производственных правил; автоматизированная система управления; семантический анализ.

ВВЕДЕНИЕ

Информационные и коммуникационные технологии с каждым днем все больше входят в нашу жизнь в тех или иных формах. Люди тесно связаны с ними на работе и в жизни. Интернет, социальные сети, электронная почта и т.п. стали неотъемлемой ее частью. Связующим звеном этих и других технологий являются программы – алгоритмы, записанные на том или ином языке. Разнообразие технологий напрямую влияет на число доступных средств их создания, но только не на их качество. И если раньше считалось, что программы, прежде всего, должны быть синтаксически правильно написаны, то сегодня во главу угла ставится смысл заложенных в них алгоритмов.

Тестирование и отладка программ не позволяют в полной мере подтвердить их семантическую корректность, т.е. правильность заложенных в них алгоритмов. Таким образом, в процессе разработки программ возникает необходимость в автоматизации некоторых функций программистов, отвечающих за проверку и анализ семантики (или смысла) программного кода, во избежание ошибок, связанных с человеческим фактором.

Целью семантического анализа программ является выявление семантических ошибок в коде, что позволит увеличить эффективность программ, снизить стоимость и сократить сроки их разработки и сопровождения, что особенно актуально для автоматизированных систем управления (АСУ), разработка которых представляет собой достаточно длительный и дорогостоящий процесс.

Таким образом, разработка эффективных методов и способов проверки и анализа программного кода АСУ на предмет обнаружения ошибок, связанных с человеческим фактором, способствует ускорению процесса отладки и более качественной работе системы автоматизации.

СПОСОБ АНАЛИЗА ПРОГРАММНОГО КОДА АСУ

В процессе отладки и анализа программ разного уровня не редки ситуации, когда в их коде встречаются операторы, которые полностью идентичны и при этом следуют либо друг за другом, либо через n операторов. Такие операторы следует трактовать как повторяющиеся, однако нет однозначного ответа на вопрос, что с ними следует делать. Прежде всего следует проанализировать содержимое (структуру) этих операторов и только потом принимать решение об их дальнейшем использовании, объединении или удалении.

Кроме того, возможны следующие ситуации в программном коде:

- переменная многократно инициализируется на некотором участке (фрагменте) программы без использования ее значения в каком-либо выражении;
- antecedentes или консеквенты продукционных правил условных операторов полностью идентичны друг другу;
- параметры операторов цикла полностью идентичны друг другу;
- переменные, значения которых изменяются в теле операторов цикла, одновременно используются в качестве их параметров, в предусловиях или постусловиях и пр.

Все перечисленные ситуации объединяет одно – отсутствие в языках программирования строгих правил, которых необходимо придерживаться для создания семантически верного программного кода.

Автором предлагается использовать в процессе разработки программ способ анализа кода, основанный на концепции семантического анализа структуры многомерного нечеткого интервально-логического регулятора (МИЛР). Отличительной особенностью МИЛР является задание на начальном этапе проектирования системы автоматического регулирования интервалов значений, которые могут принимать переменные-параметры в процессе эксплуатации, что дает возможность просчитать все возможные ситуации и выявить повторяющиеся, взаимодополняющие и противоречащие друг другу операторы. Семантический анализ программ МИЛР выполняется автоматически при помощи специального программного обеспечения.

Проблемы реализации систем управления на базе нечеткой логики рассматриваются в работах Заде, Мамдани и Сугено, предложивших алгоритмы работы нечетких регуляторов [2, 7], В.В. Круглова и А.А. Ускова [3, 4], Б.Г. Ильясова [5], А.И. Каяшева и Е.А. Муравьевой [6] и др., однако вопросы семантического анализа не нашли в них должного отражения.

Основные положения концепции семантического анализа структуры МИЛР, разработанной автором, следующие [8]:

- если условная часть (антецедент) продукционного правила не содержит термы переменных МИЛР, то это правило защиты или аварийное правило. В МИЛР может быть всего одно такое правило, предназначенное для контроля над выходом значений непрерывных величин за пределы указанного диапазона;
- если продукционное правило или его консеквент (заключение) не содержит термы переменных МИЛР, то это пустое правило;
- если правила полностью соответствуют друг другу, то это повторяющиеся правила;
- если консеквенты правил идентичны, но их antecedentes – нет, то это взаимодополняющие правила. Их можно и нужно объединять при возможности;
- если antecedentes продукционных правил идентичны, но значения их консеквентов не равны друг другу и не содержат одинаковых порядковых номеров термов, то это противоречащие правила. В противном случае это будут взаимодополняющие правила.

Рассмотрим следующий пример (примеры программ, приведенные в статье, написаны на языке программирования Pascal):

```
if X1 = 10 then Y1 := 10;  
if X1 = 10 then Y1 := 20;
```

- если значения "10" и "20" переменной Y_1 лежат в числовых пределах одного терма, то эти правила дополняют друг друга и их можно объединить, иначе они будут противоречить друг другу;
- если правило A не равно правилу B , то:
- если antecedentes правил не содержат одинаковых номеров термов для каждого из используемых блоков, то правила A и B – нормальные, т.е. удовлетворяют структуре МИЛР;

– если правила A и B – не нормальные, но значения их консеквентов не равны друг другу и не содержат одинаковых порядковых номеров термов, то это противоречащие правила. В противном случае это будут взаимодополняющие правила. Например:

```
if X1 = 10 then Y1 := 10;
if X1 = 30 then Y1 := 20;
```

Если значения "10" и "30" непрерывной величины X_i не лежат в числовых пределах одного терма, то указанные правила являются нормальными, т.е. удовлетворяют структуре МИЛР. В противном случае, если значения "10" и "20" переменной Y_i лежат в числовых пределах одного терма, то эти правила дополняют друг друга и их можно объединить. Иначе это будут противоречащие правила.

Далее рассмотрим собственно сам способ анализа программного кода АСУ. Для этого введем понятия полной, частично полной и неполной системы продукционных правил (СПП).

Предположим, что значение переменной x может меняться в диапазоне $[0; 100]$. Тогда для системы правил S :

```
if x < 70 then
  begin <блок операторов 1> end
else if x < 80 then
  begin <блок операторов 2> end
else
  begin <блок операторов 3> end;
```

термы переменной x соответствуют интервалам: $[0; 70)$, $[70; 80)$ и $(80; 100]$.

Данная СПП является полной, т.е. она описывает все возможные действия при всех возможных вариантах значений переменной x .

В общем виде полная СПП S_n для переменной x , представленной совокупностью n термов T_i , имеет вид:

```
if T1 then
  begin <блок операторов 1> end
else if T2 then
  begin <блок операторов 2> end
...
else
  begin <блок операторов n> end;
```

Система правил $S_{\text{чп}}$ для переменной x является частично полной, если ее можно считать полной, но при этом в ней отсутствует проверка истинности некоторых термов. Например, в системе правил S :

```
if T1 then
  begin <блок операторов 1> end
else if T3 then
  begin <блок операторов 2> end
...
else
  begin <блок операторов n - 1> end;
```

отсутствует проверка истинности терма T_2 , но, если нахождение переменной x в пространстве значений данного терма не влияет на значения других переменных АСУ, то такая система правил является частично полной.

Если система правил не включает в себя все возможные действия при всех возможных вариантах значений переменной x , то она является неполной и обозначается как S_n .

Решить, является ли некоторая система правил S полной, частично полной или неполной, может только эксперт.

Чтобы лучше понять, насколько важно правильно установить тип СПП, рассмотрим следующий пример. В резервуар с вмонтированным датчиком уровня наливают бензин со скоростью, которая зависит от значения переменной l , хранящей текущее значение уровня бензина в резервуаре, и представленной в виде совокупности термов T , как на рисунке 1.

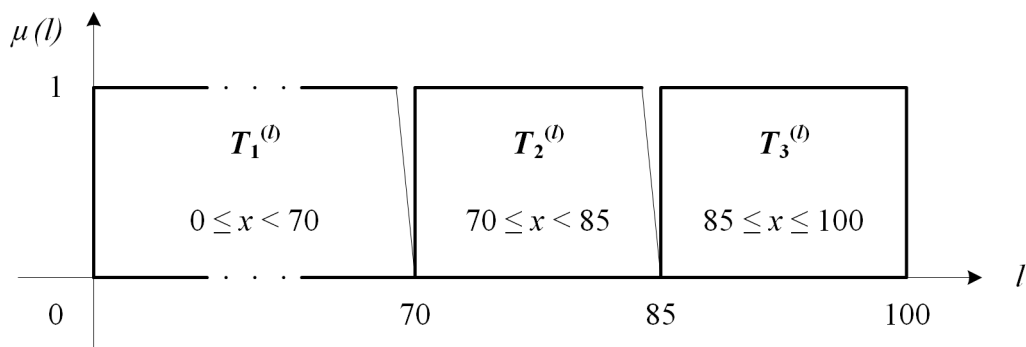


Рисунок 1 – Интерпретация переменной l в виде совокупности термов T

Упрощенно система правил S для переменной l имеет вид:

```
if T1 then
    begin <открыть подачу бензина на 100%> end
else if T2 then
    begin <открыть подачу бензина на 75%> end
else
    begin <остановить подачу бензина> end;
```

На первый взгляд, такая система правил кажется полной, однако в ней нет ни одного правила, описывающего последовательность действий в случае, если подача бензина в резервуар не прекратилась при достижении уровня l отметки в 85%.

Таким образом, если в АСУ не предусмотрен сигнализатор верхнего уровня в резервуаре, то защита от переливов отсутствует, что является потенциальным источником опасности. В таком случае СПП будет неполной.

Чтобы избежать такой ситуации, можно, например, разбить терм T_3 на два терма T_3 и T_4 , установив тем самым сигнализацию верхнего и аварийного уровней. СПП S будет иметь вид:

```
if T1 then
    begin <открыть подачу бензина на 100%> end
else if T2 then
    begin <открыть подачу бензина на 75%> end
else if T3 then
    begin <остановить подачу бензина> end
else
    begin <какие-либо защитные действия> end;
```

Частично полная СПП для данного примера может иметь вид:

```
if T1 then
    begin <открыть подачу бензина на 100%> end
else if T3 then
    begin <остановить подачу бензина> end;
```

Здесь подача бензина в резервуар останавливается при достижении уровня в 85%, но начинается только при уровне, лежащем в диапазоне $[0; 70)$.

Таким образом, предлагаемый автором способ анализа кода программ АСУ заключается в выполнении следующих действий:

1. Представить АСУ в виде МИЛР, а все ее параметры (переменные) интерпретировать эквивалентной совокупностью аргументов двузначной логики (термами).
2. Выделить в программном коде АСУ СПП для каждой из ее переменных, термы которых содержатся в антецедентах правил.
3. Установить тип каждой из выделенных СПП. Если имеются неполные системы правил, то их необходимо дополнить до полных или как минимум до частично полных.
4. Выполнить семантический анализ МИЛР, согласно положениям концепции, приведенной ранее.

Шаги 1 и 2 можно объединить; шаги 3 и 4 можно выполнять в любой последовательности и неоднократно до тех пор, пока не будут выявлены и устранены все ошибки.

УСЛОВИЯ ПРИМЕНИМОСТИ ОПИСАННОГО СПОСОБА

Для того, чтобы осуществить автоматический анализ программ с использованием этого способа, необходимо как минимум [1, 7]:

1. Определить диапазоны значений, которые могут принимать переменные в процессе работы программы. Если значение какой-либо переменной в процессе работы выйдет за пределы указанных диапазонов, то возникнет аварийная ситуация, для которой следует определить соответствующий набор (последовательность) действий. Данный шаг позволяет выявить в коде программы ошибки, связанные с неправильной инициализацией переменных значениями, лежащими вне их рабочих диапазонов.

2. Внутри диапазонов значений для каждой из переменных выделить интервалы (термы), при попадании в которые значения любой из переменных происходит отработка какого-либо продукционного правила. Эти интервалы могут быть единичными, что позволяет отмечать внутри диапазонов переменных некоторые пороговые значения. Данный шаг обусловлен тем фактом, что значения числовых переменных в большинстве случаев лежат в пределах 5-7 термов, 1-3 из которых обозначают различные рабочие состояния, а остальные – верхние и нижние переходные (пороговые) значения.

Здесь необходимо отметить одну особенность, для чего рассмотрим простой пример. Имеется задвижка трубопровода для подачи воды в резервуар, у которой имеется два положения – «открыто» и «закрыто». В программе положением задвижки управляет некая логическая переменная z , которая может принимать значения `true` и `false`, соответствующие положениям «открыто» и «закрыто» этой задвижки. Известно, что при заполнении резервуара (переменная v в программе) на 90% задвижка закрывается, а при снижении ниже 10% – открывается. Простейшая система правил для управления задвижкой будет иметь вид:

```
if v >= 90 then z := false;
if v < 10 then z := true;
```

Переменная v в данном случае может интерпретироваться следующей совокупностью термов, показанной на рисунке 2, из которого следует, что в системе правил остается незадействованным терм T_2 .

Действительно, в системе отсутствует правило для значения переменной v , расположенной в промежутке [10, 90).

Терм T_2 обозначает особый вид действия – бездействие, и, помимо прочего, служит для связки двух задействованных в программе термов. Подобные термы могут достаточно сильно усложнить анализ, поэтому их появления следует избегать путем замены логики правил.

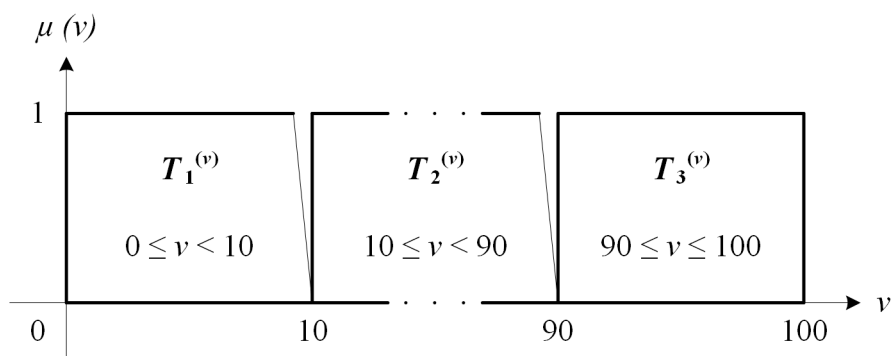


Рисунок 2 – Интерпретация переменной v в виде совокупности термов T

3. Составить таблицу взаимосвязей между всеми переменными программы для того, чтобы снизить максимальное число продукционных правил R , вычисляемое по формуле:

$$R = K(X_1) \cdot K(X_2) \cdot \dots \cdot K(X_n),$$

где $K(X_1), (X_2), \dots, (X_n)$ – число разнообразных интервалов (термов) или значений, которые могут принимать переменные X_1, X_2, \dots, X_n , анализируемые в antecedentes продукционных правил.

Если взаимосвязей между этими параметрами не существует, то выражение для расчета R примет вид:

$$R = K(X_1) + K(X_2) + \dots + K(X_n).$$

Рассмотрим пример. В МИЛР задействовано 3 входных параметра – x_1, x_2 и x_3 , интерпретированные тремя термами каждый, и 2 выходных параметра – y_1 и y_2 . Значения x_1 и x_2 влияют на значение y_1 , а x_2 и x_3 – на значение y_2 . В данном случае суммарное число продукционных правил R МИЛР лежит в диапазоне $[9; 27]$, согласно приведенным выше выражениям. Взаимосвязи параметров МИЛР отражены на рисунке 3(а).

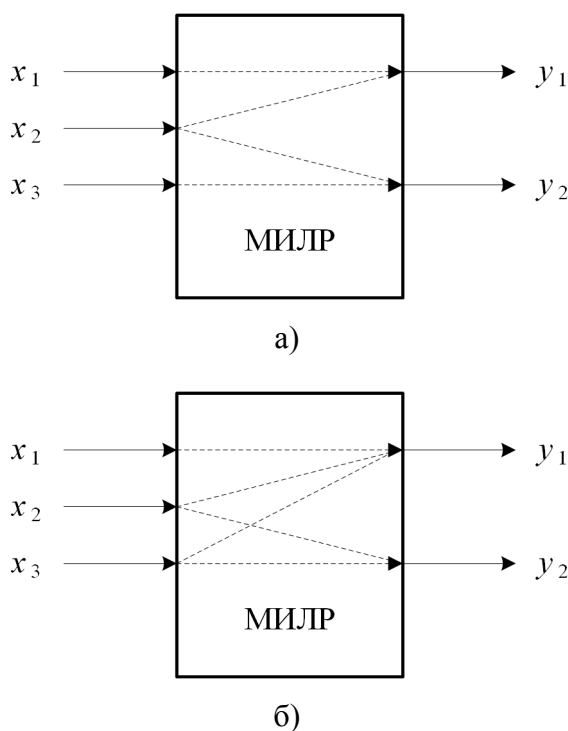


Рисунок 3 – Пример взаимосвязей в МИЛР:

а) без поглощения; б) с поглощением продукционных правил

Здесь анализируемые входы одних выходных параметров не перекрывают анализируемые входы других выходных параметров МИЛР, как в случае, представленном на рисунке 3 (б), где система правил для y_1 включает в себя правила для y_2 . Данное явление называется поглощением [9].

Можно выделить так называемые уровни связности входных и выходных параметров МИЛР в зависимости от того, с каким числом выходных и входных параметров соответственно они связаны. Это способствует более гибкой настройке взаимосвязей внутри программы и, как следствие, более точному анализу.

4. Составить СПП для всех переменных, значения которых формируются в результате вычислений логических функций, аргументами которых служат взаимосвязанные переменные. При этом требуется разбивать сложные составные правила на ряд простых. Например, составное правило

```
if ((x > 1) or (y < 5)) and (z = 6) then a := false;
```

можно разбить на следующие правила:

```
if (x > 1) and (z = 6) then a := false;
```

```
if (y < 5) and (z = 6) then a := false;
```

Анализ СПП можно выполнить, используя соответствующую программу системы автоматизированной разработки МИЛР, благодаря которой можно выявить смысловые ошибки в коде [1], после чего их нужно устранить. При необходимости анализ можно проводить неограниченное число раз.

Таким образом, в результате выполнения перечисленных выше операций получается семантически верный программный код.

ЗАКЛЮЧЕНИЕ

Предлагаемый автором способ анализа программного кода АСУ не отменяет известные и широко применяемые методы верификации, такие, как тестирование и отладка, а всего лишь дополняет их. Данный способ позволяет повысить надежность и эффективность АСУ в целом, а также снизить стоимость и сократить сроки разработки и сопровождения, т.к. способствует улучшению процесса отладки и более качественной работе АСУ.

Применение описанного способа, безусловно, сопряжено с целым рядом сложностей, как, например, разделение правил или установка взаимосвязей, что особенно тяжело сделать для крупных АСУ, код которых включает в себя десятки тысяч строк. В связи с этим необходимо выделить внутри программы небольшие блоки (подпрограммы), в каждом из которых следует определить ряд входных и выходных параметров с указанием диапазонов значений по аналогии с процессом программирования промышленных контроллеров.

СПИСОК ЛИТЕРАТУРЫ

1. Антипин А.Ф. Вопросы автоматизации семантического анализа программ // Автоматизация, телемеханизация и связь в нефтяной промышленности, 2014. – № 7. – С. 26-30.
2. Sugeno M. On Stability of Fuzzy Systems Expressed by Fuzzy Rules with Singleton Consequents // IEEE Trans. Fuzzy Systems, 1999. – № 7. – P. 201-224.
3. Круглов В.В., Усков А.А. Интеллектуальные системы управления на основе методов нечеткой логики. – Смоленск: Смоленская городская типография, 2003. – 177 с.
4. Усков А.А. Системы с нечеткими моделями объектов управления. – Смоленск: Смоленский филиал АНО ВПО ЦС «Российский университет кооперации», 2013. – 153 с.
5. Васильев В.И., Ильясов Б.Г. Интеллектуальные системы управления. Теория и практика. – М.: Радиотехника, 2009. – 392 с.
6. Каяшев А.И., Муравьева Е.А., Антипин А.Ф. Система автоматического управления элементами дистилляции на базе многомерного логического регулятора // Вестник УГАТУ.

Серия «Управление, вычислительная техника и информатика», 2010. – Т. 14. – № 4. – С. 126-131.

7. Антипин А.Ф. Особенности настройки взаимосвязей и составления системы производственных правил в интервально-логическом регуляторе // Информационные системы и технологии, 2015. – № 1. – С. 5-13.
8. Антипин А.Ф. О проверке программ автоматизированных систем управления // Современная техника и технологии, 2015. – № 1. – С. 62-66.
9. Мустафина С.А., Степашина Е.В. Редукция кинетических схем сложных химических процессов на основе теоретико-графового подхода // Вестник Казанского технологического университета, 2014. – Т. 17. – № 10. – С. 17-20.

Антипин Андрей Федорович

Стерлитамакский филиал ФГБОУ ВПО «Башкирский государственный университет», г. Стерлитамак

Кандидат технических наук, доцент кафедры прикладной информатики и программирования

Тел.: 8 917 781 76 02

E-mail: andrejantipin@mail.ru

A.F. ANTIPIN (*Candidate of Engineering Science,
Associate Professor of the Department of Applied Informatics and Programming*)
Sterlitamak Branch of Bashkir State University, Sterlitamak

**TO THE QUESTION ABOUT SEMANTIC ANALYSIS OF PROGRAMS
OF AUTOMATED CONTROL SYSTEMS**

The paper discusses the author's method of analysis of code of automated control systems, based on the concept of semantic analysis of the structure of multidimensional interval-logic controller, and the conditions for its applicability. This method allows to detect semantic errors in the code before testing and debugging.

Keywords: *multidimensional interval-logic controller; system of condition-action rules; automated control system; semantic analysis.*

BIBLIOGRAPHY (TRANSLITERATED)

1. Antipin A.F. Voprosy' avtomatizacii semanticheskogo analiza programm // Avtomatizaciya, telemexanizaciya i svyaz' v neftyanoj promy'shennosti, 2014. – № 7. – S. 26-30.
2. Sugeno M. On Stability of Fuzzy Systems Expressed by Fuzzy Rules with Singleton Consequents // IEEE Trans. Fuzzy Systems, 1999. – № 7. – P. 201-224.
3. Kruglov V.V., Uskov A.A. Intellektual'ny'e sistemy' upravleniya na osnove metodov nechetkoj logiki. – Smolensk: Smolenskaya gorodskaya tipografiya, 2003. – 177 с.
4. Uskov A.A. Sistemy' s nechetkimi modelyami ob'ektov upravleniya. – Smolensk: Smolenskij filial ANO VPO CS «Rossijskij universitet kooperacii», 2013. – 153 с.
5. Vasil'ev V.I., Il'yasov B.G. Intellektual'ny'e sistemy' upravleniya. Teoriya i praktika. – M.: Radiotekhnika, 2009. – 392 с.
6. Kayashev A.I., Murav'eva E.A., Antipin A.F. Sistema avtomaticheskogo upravleniya e'lementami distillyacii na baze mnogomernogo logicheskogo regul'yatora // Vestnik UGATU. Seriya «Upravlenie, vy'chislitel'naya tekhnika i informatika», 2010. – Т. 14. – № 4. – S. 126-131.
7. Antipin A.F. Osobennosti nastrojki vzaimosvyazey i sostavleniya sistemy' produkcijny'x pravil v interval'no-logicheskom regul'yatore // Informacionny'e sistemy' i tekhnologii, 2015. – № 1. – S. 5-13.
8. Antipin A.F. O proverke programm avtomatizirovanny'x sistem upravleniya // Sovremennaya tekhnika i tekhnologii, 2015. – № 1. – S. 62-66.
9. Mustafina S.A., Stepashina E.V. Redukciya kineticheskix sxem slozhny'x ximicheskix processov na osnove teoretiko-grafovogo podxoda // Vestnik Kazanskogo tekhnologicheskogo universiteta, 2014. – Т. 17. – № 10. – S. 17-20.

УДК 658.512.6

В.Н. ВОЛКОВ, Р.А. ЛУНЁВ, А.А. СТЫЧУК,
А.Е. ЯСТРЕБКОВ, А.С. БЫЧКОВА, А.Б. НЕЧАЕВА

АНАЛИЗ ВОЗМОЖНОСТЕЙ СОВМЕСТНОГО ИСПОЛЬЗОВАНИЯ СОЦИАЛЬНЫХ СЕТЕЙ И ГЕОИНФОРМАЦИОННЫХ СЕРВИСОВ ПРИ ОКАЗАНИИ ЭЛЕКТРОННЫХ УСЛУГ НАСЕЛЕНИЮ

В данной статье авторы освещают решение задач об обмене и обеспечении непротиворечивости геоинформации. Проведен анализ совместного использования социальных сетей и геоинформационных сервисов при работе с геоинформацией при оказании электронных услуг населению. Описано совмещение подходов социальных сетей к сбору, хранению и распространению информации и способов представления адресных данных геоинформационных сервисов, что позволяет в полной мере добиться решения всех задач по работе с геоинформацией.

Ключевые слова: *геосоциальный сервис; электронная услуга; регламент электронной услуги; электронное правительство; обслуживание населения; инструментальное средство; веб-сервис; облачные технологии; социальная сеть; геоинформационный сервис; геоинформация.*

На сегодняшний день одним из важнейших направлений развития России в настоящее время является создание информационного общества, важным принципом которого является широкое распространение и доступность для населения электронных услуг. Все большую популярность в предоставлении услуг населению приобретают веб-сервисы.

Предоставление электронных услуг посредством создания и предоставления доступа к специализированным сервисам сопряжено с необходимостью унификации интерфейсов взаимодействия с пользователем с целью минимизировать затрачиваемое пользователем время на изучение нового для него интерфейса [12]. При оказании большого количества услуг требуется осуществлять работу с информацией, привязанной к географическим координатам объекта на местности, с геоинформацией.

В современном мире повсеместно производится работа с геоинформацией – транспортные бюро, отмечающие маршруты своего проезда, кадастровые бюро, почтовые отделения, ЖРЭУ и др. Данный вид информации требует постоянной актуализации, т.к. она создается и используется параллельно самыми различными ведомствами каждый день, из-за чего старые данные быстро устаревают или могут конфликтовать друг с другом. Решение задач об обмене и обеспечении непротиворечивости геоинформации возможно лишь при наличии правильной инфраструктуры циркуляции геоданных, которая должна решать следующие классы задач:

1. Управление геоданными и метаданными.
2. Доставка геоданных пользователям.
3. Обмен геоданными внутри организации.
4. Возможность регулирования прав доступа к геоданным.
5. Возможность поиска геоданных по заданным критериям.
6. Доступ к геоданным при помощи удобного пользовательского интерфейса.

Для обеспечения своевременной доставки геоданных для пользователей и поддержания актуальности этих данных необходимо использовать динамические средства доставки данных – информационные системы и различные интернет-ресурсы, в которых геоинформация будет подлежать легкому изменению в отличие от хранения ее на бумаге или даже электронных носителях.

Создание информационной системы для работы с геоинформацией даст пользователям возможность удобно вносить в эту систему необходимые данные, держать данные в актуальном виде, устранять конфликты этих данных, быстро обмениваться ими

между другими пользователями и осуществлять гибкий поиск. Кроме того, в информационной системе автоматизируются громоздкие процессы по актуализации геоданных и разграничению прав доступа, понижая сложность этих процессов и уменьшая количество ошибок в них.

Растущая потребность рынка в геоинформации, особенно оперативной, требует предоставлять доступ к ней через сеть Интернет, причем так, чтобы она была понятна пользователям и удобна в использовании. Для этого необходимо организовать доступ к данным с помощью удобных, простых и интуитивно понятных пользовательских интерфейсов, которые будут предоставлены пользователю в зависимости от решаемой в данный момент задачи.

Таким образом, автоматизация процессов сбора, обмена и хранения информации позволяет решить следующий круг задач:

1. Создание геоданных и метаданных и поддержание их в актуальном состоянии.
2. Стабильная доставка геоданных пользователям в реальном времени независимо от используемого ими ПО и платформ.
3. Организация эффективного обмена геоданными внутри организации независимо от ее структуры и места положения.
4. Управление правами доступа к данным.
5. Организация поиска геоданных пользователями.
6. Возможность настройки пользовательского интерфейса для доступа к геоданным под конкретные задачи.

В настоящее время одними из лучших средств для обмена информацией между пользователями являются социальные сети – пользователи с легкостью могут найти друг друга в такой системе, а затем обмениваться личными сообщениями, новостями или другой необходимой информацией. Данный подход давно подтвердил свою эффективность и является самым популярным методом распространения информации в сети Интернет на сегодняшний день. Кроме того, он продолжает набирать обороты.

Среднее количество пользователей одной только социальной сети ВКонтакте насчитывает 54,6 млн человек в России за зиму-лето 2014-2015 гг. (рис. 1), что показывает, что данная социальная сеть является самым популярным интернет-ресурсом в России за указанный период времени.

Колоссальный объем информации размещается в данных социальных сетях различными авторами. Приведенная аналитическими компаниями статистика показывает лишь объем обмена сообщениями, что покрывает далеко не всю систему – в социальных сетях есть и другие способы распространения информации – группы, встречи, загружаемые файлы (например, фотографии, видео), комментарии и др.

Вся приведенная информация подтверждает эффективность данного метода распространения информации среди пользователей, что делает этот метод и структуру наиболее предпочтительным для распространения геоинформации. Многие современные социальные сети уже позволяют обмениваться геоданными с использованием различных геоинформационных сервисов, таких, как Яндекс.Карты, Google Maps, Open Street Maps и пр., которые позволяют наглядно визуализировать представляемые геоданные.

Однако данный подход не позволяет структурировать эту информацию для дальнейшего использования, поиска, группировки и сортировки этой информации. Эта информация носит лишь демонстративный характер, такой же, как если бы пользователь просто приложил изображение к своему сообщению. В таком сообщении работа с метаданными геоинформации сведена к минимуму.

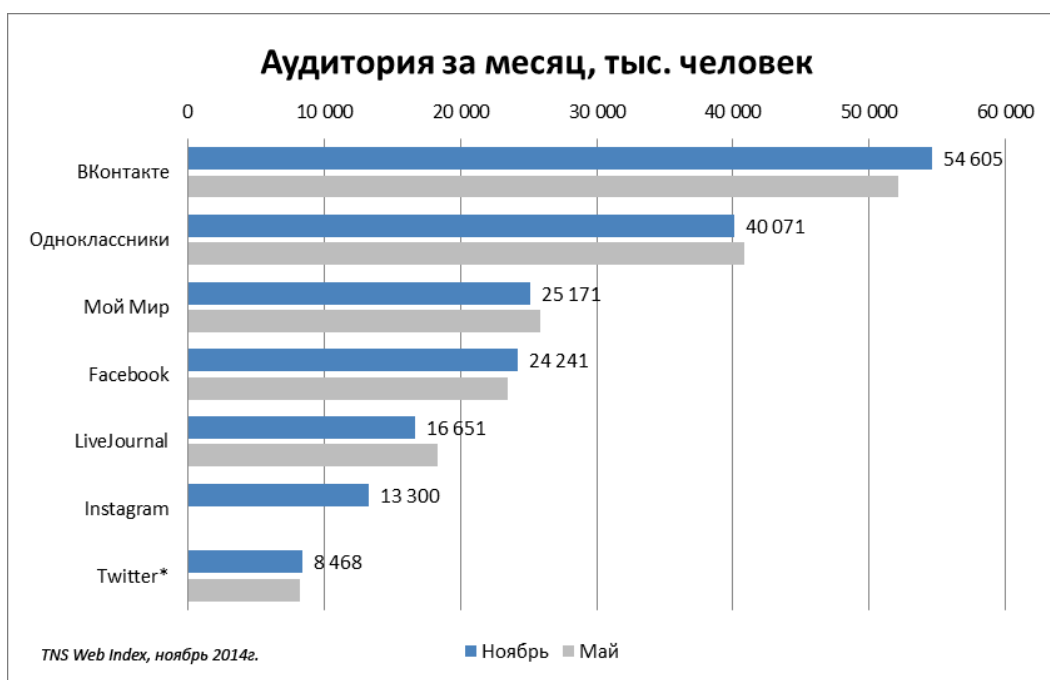


Рисунок 1 – Аудитория социальных сетей в России за зиму-лето 2014-2015 гг.

Основным способом общения и обмена информацией в социальных сетях является обмен личными сообщениями (рис. 2, 3).

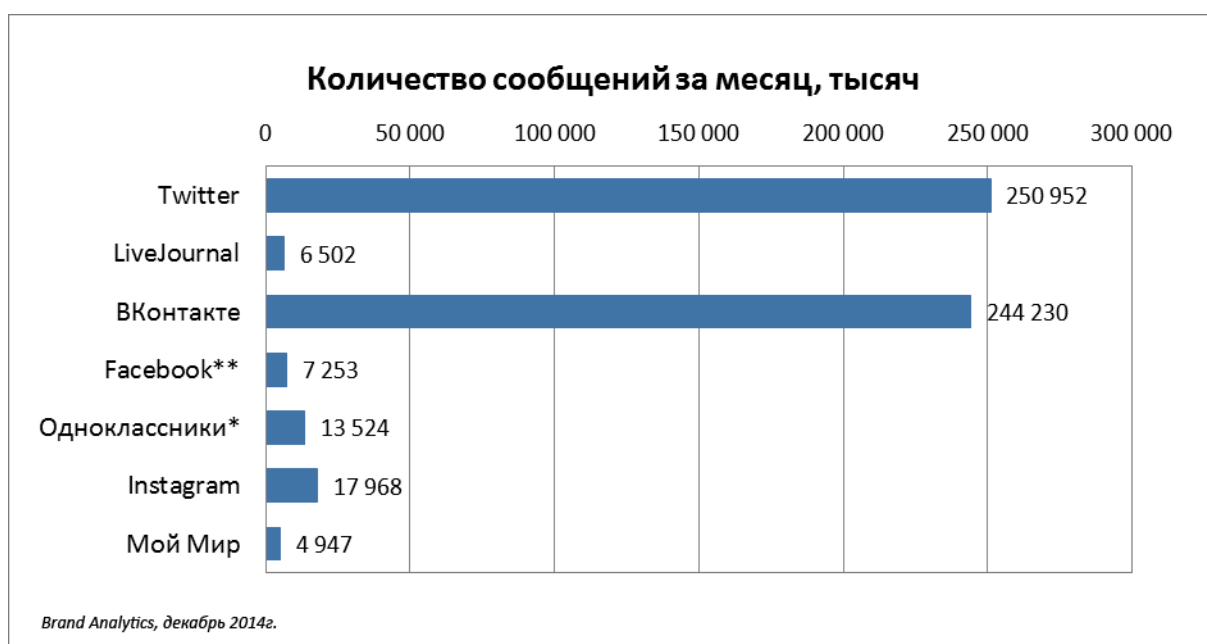


Рисунок 2 – Количество сообщений за месяц в социальных сетях в России за зиму-лето 2014-2015 гг.

Сами геоинформационные сервисы в основном носят собирательный и демонстративный характер – они либо самостоятельно регистрируют необходимые для отображения данные, либо принимают эти данные по заявкам пользователей, которые обрабатываются порой долгое время. Пока эти заявки обрабатываются, геоданные устаревают, становятся неактуальными и уже могут конфликтовать с другими новыми данными. Таким образом, одно из требований к работе с геоинформационными данными

соблюдается в таких системах не полностью – эти системы добавляют и изменяют геоинформацию, но просто не в состоянии держать ее в актуальном виде. Главным преимуществом данных систем является удобный пользовательский интерфейс для представления этих данных – представление данных на карте.

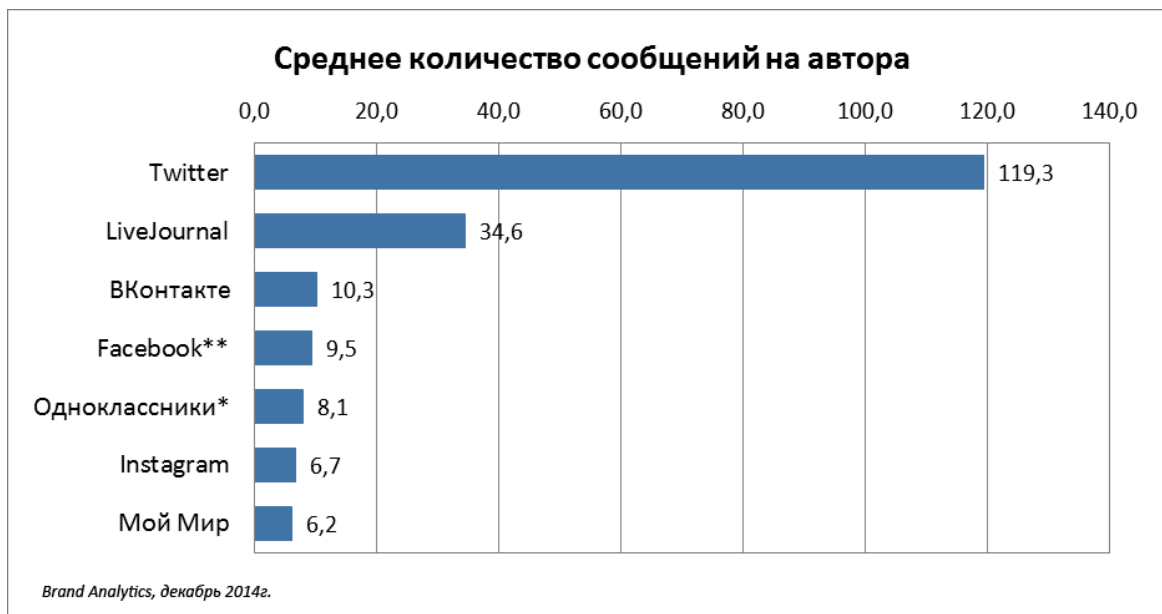


Рисунок 3 – Среднее количество сообщений на автора в России за зиму-лето 2014-2015 гг.

Компании, предоставляющие геоинформацию в виде карт местности, дают открытый доступ к географической информации – они позволяют получить доступ о различного рода местах, координатах, улицах, городах. Однако за эти данными обычно скрывается нечто большее, чем просто адреса и координаты. В таких системах не всегда данные о каком-либо месте содержат актуальную логическую информацию о том, что конкретно в этом месте находится – магазин, университет, кафе и т.д. Одной из главных задач в таких системах является актуализация геоинформации и лишь потом логической информации, связанной с этими геоданными.

Требование в привязке логической информации к геоданным вызвало появление таких геоинформационных сервисов, как Lokata и 2GIS. Создатели геоинформационного сервиса 2GIS используют карты местности, предоставляемые распространителями OpenStreetMap, на которых отмечают различную логическую информацию о местах на карте с возможностью подробного поиска. Размещается информация о ресторанах, гостиницах, автосервисах и других местах, пользователю дается возможность осуществлять поиск не только по адресу, но и по таким полям, как, например, «средняя стоимость обеда» в случае, если ищутся рестораны, «рейтинг компании», «стоимость номера» в случае, если ищутся гостиницы, и т.д.

Проблема данного сервиса, однако, остается такой же, как и у других геоинформационных сервисах – разработчики данной системы самостоятельно обновляют контент. Они осуществляют централизованный сбор геоинформации и сами решают, какую информации необходимо им показывать на своих картах местности. Внесение любой новой информации обычным пользователем требует подачи соответствующей заявки, что делает данную информацию закрытой для изменения и может сделать эту информацию неактуальной после истечения какого-то времени, пока эта заявка будет обрабатываться.

Существующие решения не давали возможности пользователям обмениваться необходимой геоинформацией в режиме реального времени либо решали эту задачу в узком

спектре как, например, подсистема Яндекс.Пробки, отмечающая на картах проблемные в данный момент места для проезда по городу.

Необходимость в распространении геоданных в режиме реального времени вызвала появление таких систем, как Foursquare. Это социальная сеть с функцией геопозиционирования, предназначенная в основном для работы с мобильными устройствами. Пользователи отмечают в различных заведениях с помощью мобильного устройства и оставляют отзывы о различных заведениях. У данного геосервиса есть четкий круг применения, который в основном используется для создания отзывов каким-либо заведениям.

Также путем развития обмена геоданными в режиме реального времени стало создание полноценного геосоциального сервиса AlterGeo, который позволяет пользователям в зависимости от их текущего местоположения находить поблизости и выбирать места для отдыха, делиться с друзьями рекомендациями и фотографиями о конкретных местах на карте. Данный геосоциальный сервис уже не специализируется на конкретной задаче и конкретных местах, он охватывает любые геоданные и любую связанную с ними логическую информацию. Пользователи сами выбирают, какой информацией хотят поделиться. Кроме того, пользователи могут оставлять комментарии к событиям своих друзей и окружающих в одной из соответствующих лент. Человек, чье событие прокомментировали, и другие пользователи могут просматривать такие сообщения и отвечать на них. Таким образом, комментарии становятся своего рода мини-чатами, каждый из которых, как правило, привязан к конкретным местам. С помощью встроенной функции кросс-постинга круг обмена такой информацией выходит за рамки этого сервиса, поскольку позволяет легко транслировать ее на свои страницы в популярных социальных сетях.

Совмещение подходов социальных сетей к сбору, хранению и распространению информации и способов представления адресных данных геоинформационных сервисов позволяет в полной мере добиться решения всех задач по работе с геоинформацией. Однако в существующей системе AlterGeo отсутствует как таковая возможность структурирования получаемой от других пользователей геоинформации – пользователи не могут объединяться в группы или сообщества, в которых будет доступна строго необходимая им информация. Вся создаваемая информация сразу является открытой. Пользователь не может самостоятельно разбивать на группы публикуемую геоинформацию и устанавливать для нее списки приватности. Кроме того, в данной системе используется маркетинговый подход к социализации пользователей, на котором основывается вся работа системы – пользователям за различные социальные операции в системе выдаются специальные очки, которыми они могут пользоваться для получения различных привилегий, скидок в заведениях, для создания рейтинга и так далее.

Таким образом, можно сказать, что существующая система решает много проблем, связанных со сбором, хранением и обменом геоинформации, но не решает проблему управления геоданными и метаданными в полной мере из-за отсутствия возможности структуризации геоинформации и проблему возможности гибкого регулирования пользовательских прав, в какой-то мере связанную с проблемой отсутствия структуризации информации.

Используя подход данной геосоциальной сети и дав пользователю возможность не только самостоятельно создавать, описывать и обмениваться с другими геоданными, но и дав возможность пользователю разбивать, сортировать, группировать и показывать эту информацию в нужном ему виде, можно получить автоматизированную систему, которая будет удовлетворять всем требованиям по работе с геоданными:

1. Управление геоданными и метаданными. Пользователь системы создает необходимые отметки на картах местности, которые он сам описывает и изменяет в режиме реального времени. Пользователь может разбить информацию по группам и предоставлять

ее другим структурировано, например, в каталогах или созданных группах как в других социальных сетях, добавлять геоданным различную информацию для поиска и фильтрации отображения информации на карте. Данные хранятся в распределенном виде у пользователей, а не централизованно в одной закрытой системе, что позволяет самими пользователями держать их в актуальном виде.

2. Доставка геоданных пользователям. После создания соответствующей геоинформации пользователь в режиме реального времени может оповестить других пользователей о ней путем публикации ее на своей странице или с помощью системы личных сообщений или уведомлений.

3. Обмен геоданными внутри организации. Поскольку геоинформация подлежит любого вида группировке и сортировке, ей будет удобно обмениваться между различными организациями точно также, как могли бы обмениваться два обычных пользователя в системе.

4. Возможность гибкого регулирования пользовательских прав. Группируя геоданные, пользователь сможет в системе отметить людей, которые смогут их просмотреть.

5. Возможность самостоятельного поиска геоданных. Возможность соответствующей группировки и сортировки данных дает возможность искать необходимую как геоинформацию, так и логическую информацию, связанную с этими данными.

6. Удобство пользовательского интерфейса. Геосоциальные сервисы совмещают в себе способы предоставления геоданных от геоинформационных сервисов (карты местности) с принципами социальных сетей для их сбора, хранения и обмена между пользователями. Это позволяет не только удобно вносить любые необходимые изменения с помощью предоставляемых интерфейсов, но и в дальнейшем работать с этими данными, искать, сортировать и группировать их с помощью уже давно известных и популярных интерфейсных решений, применяющихся в других социальных сетях.

Создание соответствующей геосоциальной сети позволит решать широкий класс задач по работе с геоданными, а не конкретные узкие специализированные задачи, такие, как, например, отображение пробок на улице. Пользователи смогут мгновенно (так же, как и в социальных сетях) оповестить друг друга о каком-либо месторасположении и поделиться логической информацией, связанной с ним. Поскольку вся вводимая геоинформация подлежит группировке, пользователям будет легко найти то, что именно им необходимо в системе не только по адресным данным, но и по другим логическим данным, как это делается в геоинформационных сервисах. Различные компании и организации могут быть представлены в такой системе в виде сообщества, которое публикует различную геоинформацию для своих пользователей, а благодаря настройкам приватности они также смогут легко обмениваться этими данными друг с другом, дав возможность только определенному кругу людей просматривать имеющуюся информацию.

Хранимые геоданные будут распределены в системе между ее пользователями, что позволит самим пользователям вносить контент на карты местности или создавать собственные карты местности, на которых они смогут показать только интересующие их места – собираться данная информация системой будет лишь для выдачи данных в форме поиска. Благодаря возможности группировать информацию по собственно созданным картам местности или каталогам у пользователей системы появляется возможность получить все необходимые геоданные от системы или своих друзей в нужном настраиваемом формате с использованием системы фильтрации поиска.

Сделав данную систему высокоинтегрируемой с другими социальными сетями, данная геосоциальная сеть не будет пытаться бороться с этими системами за пользователей, а даст им новые инструменты для работы с геоданными и обмена ими.

Одной из главных проблем, которая может встретиться на пути создания и использования данного сервиса, является возможность внесения пользователем в систему

дезинформации нарочно или случайно. Это явление является главной причиной создания в геоинформационных сервисах централизованного подхода к сбору и хранению геоданных – данные проверяются до того, как будут предоставлены пользователям. Массовость использования создаваемого сервиса пользователями других социальных сетей позволит преодолеть данный порог дезинформации, как это было преодолено в социальных сетях, таким образом позволив полностью применить распределенный подход к сбору и хранению геоданных вместо централизованного подхода и устранить этот недостаток.

Широкий подход к управлению геоинформацией самими пользователями позволит им использовать созданные средства независимо от вида геоинформационных данных (пробки, магазины, заведения, чрезвычайные ситуации в городе, строительство и ЖКХ), взаимодействуя с другими пользователями в режиме реального времени и решая самый разнообразный класс задач от обмена комментариями о любимых местах города как в других геосервисах до информирования пользователей о чрезвычайных ситуациях.

СПИСОК ЛИТЕРАТУРЫ

1. Ветцель И. Перспективные архитектуры систем для работы с геоданными [Электронный ресурс]. – URL: <http://www.futureussia.ru/upload/iblock/63a/606-erdas-apollo.pdf>.
2. Ветцель И. Геоинформационные данные для обеспечения жизнедеятельности человека [Электронный ресурс]. – URL: <http://cyberleninka.ru/article/n/geoinformatsionnye-dannye-dlya-obespecheniya-zhiznedeyatelnosti-cheloveka>.
3. Струков Д. Проект федерального эпидемиологического геопортала Российской Федерации. Группа компаний «Центр пространственных исследований» [Электронный ресурс]. – URL: <http://www.gosbook.ru/node/55705>.
4. Кравцова, Н.А. Структурная модель сущностей и типов данных системы мониторинга объектов учета // Известия Орловского государственного технического университета, 2009. – № 6. – С. 11-14.
5. Кравцова Н.А., Константинов И.С., Фролов А.И. Модель хранения данных в адаптивной автоматизированной системе административного мониторинга // Информационные системы и технологии, 2010. – С. 66-73.
6. Кравцова Н.А., Фролов А.И. Формализация процесса генерации диалоговых интерфейсов сбора данных в автоматизированной адаптивной системе административного мониторинга // Информационные системы и технологии, 2011. – С. 12-20.
7. Жучков Е.В. Геоинформационный сервис карты города // Прикладная математика, управление и информатика: сборник трудов международной молодежной конференции: в 2 т. – Белгород, 3-5 октября 2012 г. – Белгород: ИД «Белгород», 2012. – Т. 2., С. 99-101.
8. Жучков Е.В. Перспективы использования геосоциальных сервисов в реалиях современного бизнеса // Информационные технологии в науке, образовании и производстве (ИТНОП): материалы VI Международной научно-технической конференции, 22-23 мая 2014 г. – Орел: ФГБОУ ВПО «Госуниверситет – УНПК», 2014 [Электронный ресурс]. – URL: <http://irsit.ru/files/article/467.pdf>. – 7 с.
9. Жучков Е.В. Перспективы использования геосоциальных сервисов в реалиях современного бизнеса // Сборник трудов международного конкурса научных работ студентов SmartIT. – Орел: ФГБОУ ВПО «Госуниверситет – УНПК», 2014. – С. 52-57.
10. Жучков Е.В. Аспекты энергоэффективности мобильных устройств при технической реализации геосоциального сервиса // Теплотехника и информатика в образовании, науке и производстве: сборник докладов IV Всероссийской научно-практической конференции студентов, аспирантов и молодых ученых (ТИМ'2015) с международным участием, посвященной 95-летию основания кафедры и университета (Екатеринбург, 26-27 марта 2015 г.). – Екатеринбург: УрФУ, 2015. – С. 219-221.
11. Жучков Е.В. «Глобальная деревня», или Тенденции развития глобальной сети // Материалы VII Международной студенческой электронной научной конференции

«Студенческий научный форум» [Электронный ресурс]. – URL: www.scienceforum.ru/2015/1354/13290.

12. Лунев Р.А., Стычук А.А., Митин А.А. Требования к составу функций веб-сервиса оказания электронных услуг населению // Информационные системы и технологии, 2015. – № 1/87. – С. 49-58.
13. Лунев Р.А. Геосоциальный сервис как электронная услуга населению / Р.А. Лунев, А.А. Стычук, В.Н. Волков, А.А. Митин // Информационные системы и технологии, 2015. – № 3/89. – С. 65-70.

Волков Вадим Николаевич

ФГБОУ ВПО «Госуниверситет – УНПК», г. Орел
Кандидат технических наук, доцент, доцент кафедры «Информационные системы»
Тел.: 8 (4862) 43-49-56
E-mail: vadimvolkov@list.ru

Лунёв Роман Алексеевич

ФГБОУ ВПО «Госуниверситет – УНПК», г. Орел
Кандидат технических наук, доцент, директор НОЦ ФиПИТ
Тел.: 8 920 287 79 85
E-mail: rolu@yandex.ru

Стычук Алексей Александрович

ФГБОУ ВПО «Госуниверситет – УНПК», г. Орел
Кандидат технических наук, доцент, заместитель директора ресурсного центра информатизации образования по научно-методической работе
Тел.: 8 (4862) 43-49-56
E-mail: stichuck@yandex.ru

Ястребков Артём Евгеньевич

ФГБОУ ВПО «Госуниверситет – УНПК», г. Орел
Аспирант кафедры «Информационные системы»
Тел.: 8 (4862) 43-49-56
E-mail: nerlin57@gmail.com

Бычкова Анастасия Сергеевна

ФГБОУ ВПО «Госуниверситет – УНПК», г. Орел
Студент кафедры «Информационные системы»
Тел.: 8 (4862) 43-49-56
E-mail: anastasiya@skb-it.ru

Нечаева Анастасия Борисовна

ФГБОУ ВПО «Госуниверситет – УНПК», г. Орел
Студент кафедры «Информационные системы»
Тел.: 8 (4862) 43-49-56
E-mail: nechaevastasya@yandex.ru

V.N. VOLKOV (*Candidate of Engineering Sciences, Associate Professor,
Associate Professor of the Department «Information Systems»*)

R.A. LUNYOV (*Candidate of Engineering Sciences, Associate Professor,
Director of Research and Education Center «Fundamental and Applied Information Technologies»*)

A.A. STY'CHUK (*Candidate of Engineering Sciences, Associate Professor,
Deputy Director of Resource Center of Informatization of Education on Scientific and Methodological Work*)

A.E. YaSTREBKOV (*Post-graduate Student of the Department «Information Systems»*)

A.S. BY'ChKOVA (*Student of the Department «Information Systems»*)

A.B. NEChAEVA (*Student of the Department «Information Systems»
State University – ESPC, Orel*)

THE ANALYSIS OF OPPORTUNITIES OF SHARING OF SOCIAL NETWORKS AND GEOINFORMATION SERVICES WHEN RENDERING ELECTRONIC SERVICES TO THE POPULATION

Authors cover the solution of tasks on an exchange and ensuring consistency of geoinformation in this article. The analysis of sharing of social networks and geoinformation services during the work with geoinformation is carried out when rendering electronic services to the population. Combination of approaches of social networks to collecting, storage and distribution of information and ways of representation of address these geoinformation services that allows to achieve fully the solution of all tasks of work with geoinformation is described.

Keywords: *geosocial service; electronic service; regulations of electronic service; electronic government; population service; tool means; web service; cloudy technologies; social network; geoinformation service; geoinformation.*

BIBLIOGRAPHY (TRANSLITERATED)

1. Vetcel' I. Perspektivny'e arxitektury' sistem dlya raboty' s geodanny'mi [E'lektronny'j resurs]. – URL: <http://www.futurerussia.ru/upload/iblock/63a/606-erdas-apollo.pdf>.
2. Vetcel' I. Geoinformacionny'e danny'e dlya obespecheniya zhiznedeyatel'nosti cheloveka [E'lektronny'j resurs]. – URL: <http://cyberleninka.ru/article/n/geoinformatsionnye-dannye-dlya-obespecheniya-zhiznedeyatel'nosti-cheloveka>.
3. Strukov D. Proekt federal'nogo e'pidemiologicheskogo geoportala Rossijskoj Federacii. Gruppa kompanij «Centr prostranstvenny'x issledovaniy» [E'lektronny'j resurs]. – URL: <http://www.gosbook.ru/node/55705>.
4. Kravcova, N.A. Strukturnaya model' sushhnostej i tipov danny'x sistemy' monitoringa ob'ektov ucheta // Izvestiya Orlovskogo gosudarstvennogo texnicheskogo universiteta, 2009. – № 6. – S. 11-14.
5. Kravcova N.A., Konstantinov I.S., Frolov A.I. Model' xraneniya danny'x v adaptivnoj avtomatizirovannoj sisteme administrativnogo monitoringa // Informacionny'e sistemy' i texnologii, 2010. – S. 66-73.
6. Kravcova N.A., Frolov A.I. Formalizaciya processa generacii dialogovy'x interfejsov sbora danny'x v avtomatizirovannoj adaptivnoj sisteme administrativnogo monitoringa // Informacionny'e sistemy' i texnologii, 2011. – S. 12-20.
7. Zhuchkov E.V. Geoinformacionny'j servis karty' goroda // Prikladnaya matematika, upravlenie i informatika: sbornik trudov mezhdunarodnoj molodezhnoj konferencii: v 2 t. – Belgorod, 3-5 oktyabrya 2012 g. – Belgorod: ID «Belgorod», 2012. – T. 2., S. 99-101.
8. Zhuchkov E.V. Perspektivy' ispol'zovaniya geosocial'ny'x servisov v realiyax sovremennogo biznesa // Informacionny'e texnologii v nauke, obrazovanii i proizvodstve (ITNOP): materialy' VI Mezhdunarodnoj nauchno-texnicheskoy konferencii, 22-23 maya 2014 g. – Orel: FGBOU VPO «Gosuniversitet – UNPK», 2014 [E'lektronny'j resurs]. – URL: <http://irsit.ru/files/article/467.pdf>. – 7 c.
9. Zhuchkov E.V. Perspektivy' ispol'zovaniya geosocial'ny'x servisov v realiyax sovremennogo biznesa // Sbornik trudov mezhdunarodnogo konkursa nauchny'x rabot studentov SmartIT. – Orel: FGBOU VPO «Gosuniversitet – UNPK», 2014. – S. 52-57.
10. Zhuchkov E.V. Aspekty' e'nergoeffektivnosti mobil'ny'x ustrojstv pri texnicheskoy realizacii geosocial'nogo servisa // Teplotexnika i informatika v obrazovanii, nauke i proizvodstve: sbornik dokladov IV Vserossijskoj nauchno-prakticheskoy konferencii studentov, aspirantov i molody'x ucheny'x (TIM'2015) s mezhdunarodnym uchastiem, posvyashhennoj 95-letiyu osnovaniya kafedry' i universiteta (Ekaterinburg, 26-27 marta 2015 g.). – Ekaterinburg: UrFU, 2015. – S. 219-221.
11. Zhuchkov E.V. «Global'naya derevnya», ili Tendencii razvitiya global'noj seti // Materialy' VII Mezhdunarodnoj studencheskoj e'lektronnoj nauchnoj konferencii «Studencheskij nauchny'j forum» [E'lektronny'j resurs]. – URL: www.scienceforum.ru/2015/1354/13290.
12. Lunev R.A., Sty'chuk A.A., Mitin A.A. Trebovaniya k sostavu funkciy veb-servisa okazaniya e'lektronnyx uslug naseleniyu // Informacionny'e sistemy' i texnologii, 2015. – № 1/87. – S. 49-58.
13. Lunev R.A. Geosocial'ny'j servis kak e'lektronnaya ushuga naseleniyu / R.A. Lunev, A.A. Sty'chuk, V.N. Volkov, A.A. Mitin // Informacionny'e sistemy' i texnologii, 2015. – № 3/89. – S. 65-70.

Н.Г. ПЕНЬКОВ, С.В. ЕРЕМЕНКО, А.Г. СТАВЦЕВА

МОДЕЛИРОВАНИЕ УЧЕТА И ПРИЕМА УЧЕТНЫХ БЛАНКОВ В СИСТЕМЕ ЭЛЕКТРОННОГО ДОКУМЕНТООБОРОТА

В статье приведена экспериментальная оценка качества обработки информации реальной информационной системы корпоративного уровня. Рассмотрен алгоритм «Автоматическое обнаружение ошибок и восстановление достоверности значений показателей» на примере учета и приема учетных бланков в системе электронного документооборота. Предложена модель для анализа показателей, оценки качества и эффективности работы информационной системы.

Ключевые слова: электронный документооборот; автоматизированная информационная система; информационно-вычислительный центр; модель; параметры; производительность; индекс; учетные бланки.

Документ является основным способом представления информации на любом современном предприятии. Неоспорима важность сохранности и умелого использования информационных ресурсов предприятия для успешного ведения бизнеса. Способность принять верное решение и вовремя отреагировать на ситуацию, гибко реагировать на все изменения рынка зависит не только от таланта и опыта руководителей. Эффективность управления предприятием зависит и от того, насколько разумно в нем организовано управление документооборотом. Фактически малоэффективное использование накопленной информации (или, еще хуже, ее утрата) может привести к потере всего бизнеса. Ведь вовремя не полученная информация или документ – это, прежде всего, потерянные деньги, время и упущенные возможности. Вследствие этого на любом предприятии, где ведется активная работа с различными документами, рано или поздно встает проблема систематизации, обработки и безопасного хранения значительных объемов информации. Важную роль в оптимизации деятельности предприятия любого размера и профиля деятельности играют современные системы электронного документооборота, которым и посвящена данная статья.

Работы по подготовке, вводу, контролю и анализу документов, ошибок в документах и их исправлению выполняются в соответствии с технологическими рабочими инструкциями [1]. В контролируемых документах контрольные суммы подсчитываются в соответствии с регламентом их представления в информационно-вычислительном центре (ИВЦ) только по тем строкам, в которых они отсутствуют. Ниже приводится описание функционирования алгоритма и программы «Автоматическое обнаружение ошибок и восстановление достоверности значений показателей» в реальной автоматизированной информационной системе (АИС) на уровне технологического процесса обработки данных (ТПОД) [2].

Рассмотрим последовательность получения исходных данных для экспериментальной оценки качества обработки информации реальной информационной системой (ИС) корпоративного уровня.

На первом этапе производятся учет и прием первичных документов – заполненных учетных бланков (УБ), поступающих от подчиненных предприятий в ИВЦ организации. Взята выборка объемом 101 пачка УБ. Дефектной считалась та пачка, которая поступала с опозданием после срока, установленного организацией. Каждая пачка регистрировалась случайным событием в ведомости дефектов отдельной строкой. На данном этапе также выявлялись дефекты по полноте – отсутствие значений показателей в УБ. Объем выборки в данном случае составил 250 УБ. Дефекты по достоверности на данном этапе не проявились.

На втором этапе (прием УБ после их индексирования (кодирования) в ИВЦ) путем анализа УБ и журнала регистрации приема УБ от предприятий методом случайных чисел была взята выборка объемом 164 пачки УБ за определенный период. Техническими условиями по плану-графику установлено время кодирования – 200 УБ за рабочую смену, поэтому как дефектные идентифицировались те пачки УБ, время кодирования которых превысило установленное. На данном этапе дефектов по полноте обнаружено не было, а дефекты по достоверности не выявлялись.

На третьем этапе (компьютерный ввод УБ и обработка документации) была взята выборка объемом 200 УБ, дефекты по своевременности и полноте не обнаружены. Дефекты по достоверности регистрировались отдельной строкой в ведомости.

На четвертом этапе обработки (выдача выходных (результатных) документов абонентам ИС) взята выборка объемом 4806 УБ, в которой были выявлены только дефекты по достоверности – 10 ошибочных символов.

Для кластеризации дефектов и получения статистических оценок по классам дефектов относительно времени и стоимости могут быть использованы пакеты прикладных программ статистического анализа данных. Кластеризация дефектов выполняется по программе кластер-анализа данных. В результате формируются таблица, описывающая пошаговый процесс объединения кластеров, и дендрограмма классификации дефектов (рис. 1).

В каждой строке таблицы ее элементы означают: 1-й – порядковый номер шага объединения; 2-й – отклонение между объединяемыми на данном шаге кластерами (евклидово расстояние); 3-й и 4-й – соответственно разности между средним значением каждой переменной по времени и стоимости и средним взвешенным значением образованного на данном шаге кластера; 5-й – суммарный вес реализаций в кластере, полученном на данном шаге.

Дендрограмма на рисунке 1 соответствует машинной распечатке и имеет вид древовидной горизонтальной схемы. Слева от нее размещены три столбца цифр: 1-й столбец означает номера строк (реализаций) вводимых в компьютер данных из ведомости; 2-й – содержит коды дефектов, участвующих в кластеризации; 3-й указывает последовательность объединения кластеров (дефектов) на каждом шаге. Дендрограмма показывает, что в соответствии с кластеризацией на шаге 497 завершено объединение дефектов в классе по достоверности – дефекты с кодами 72, 74, 77, 85, 86 и др. На шаге 504 сформировался класс дефектов по полноте – коды 12, 22, на шаге 514 – класс дефектов по своевременности – код 3. Внутри класса своевременности модификации дефектов отсутствуют, внутри класса дефектов по полноте имеются две разновидности дефектов, а по классу дефектов достоверности – 10 модификаций.

Оценка параметров ИС выполняется по программе статистического анализа данных «STADIA» [3]. На внутримашинном уровне исходные данные могут быть представлены следующими массивами: массив А – данные по дефектам достоверности; массив В – по полноте; массив DE – по своевременности. Для удобства анализа и оценки уровня качества полученные на ЭВМ основные статистические данные представляются в унифицированной форме карты данных распределения дефектов, например по своевременности (табл. 1). Эти карты наглядно отображают результаты обработки экспериментальных данных, необходимые для расчета показателей комплексной оценки качества.

В нашей конкретной задаче среднеквадратические отклонения существенно отличаются по достоверности, полноте и своевременности как параметров времени, так и от параметров стоимости. Если среднее выборочное по достоверности равно 3,084 мин, по полноте – 11,536 мин, то по своевременности – 2548,30 мин или около 42 ч, т.е. факторы-дефекты по своевременности доминируют в общей структуре факторов-дефектов, снижающих качество обработки УБ. Среднее время опоздания документа составляет 42 ч.

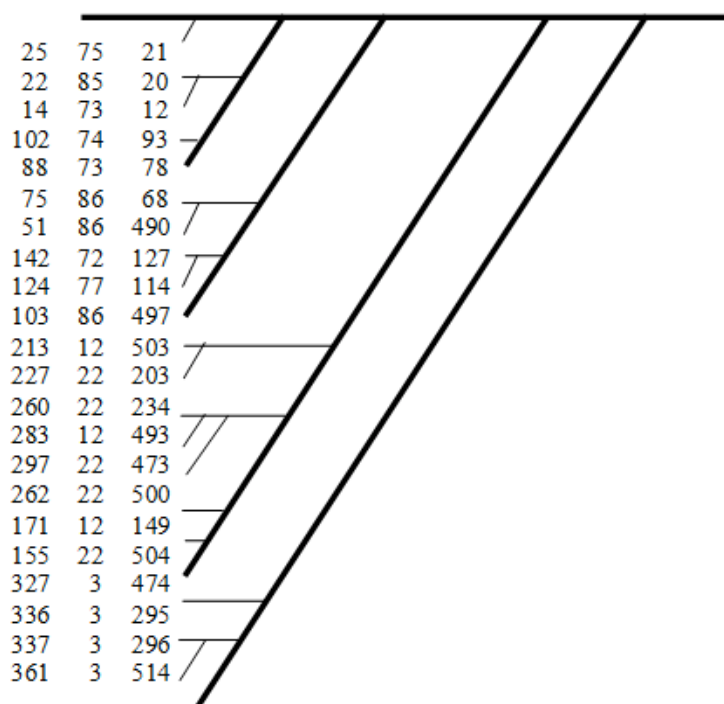


Рисунок 1 – Фрагмент дендрограммы кластеризации дефектов автоматизированной обработки учетных бланков

Таблица 1 – Карта данных распределения дефектов по своевременности на первом и третьем этапах

№ п/п	Границы интервалов, мин	Абсолютная частота, пачка	Относительная частота	Накопленная частота
1	492-984	8	0,04	0,04
2	984-1476	13	0,07	0,11
3	1476-1968	19	0,1	0,21
4	1968-2460	29	0,15	0,36
5	2460-2952	44	0,23	0,59
6	2952-3444	30	0,15	0,74
7	3444-3936	26	0,13	0,87
8	3936-4428	15	0,08	0,95
9	4428-4920	9	0,04	0,99
10	4920-5412	2	0,01	1

Характеристики, значения которых являются исходными для дальнейших расчетов:

- объем выборки: пачек документов – 269, среднее число документов в пачке – 46, документов – 12374;
- среднее выборочное пачки: $x = 2548,30769$;
- среднеквадратическое отклонение: $\sigma = 1003,28$;
- границы доверительного интервала среднего выборочного: нижняя – 27,11, верхняя – 3016,32;
- вероятность дефекта (пачки документов): $\rho = 195/269 = 0,724907$;
- показатель своевременности $\rho_s = 1 - 0,724907 = 0,275093$;
- оценка математического ожидания дефекта: $X = x\rho = 2548,30769 \cdot 0,724907 = 1847,286$ мин.

Графически распределение дефектов по их видам отображается в виде гистограмм. На рисунке 2 представлена гистограмма распределения дефектов по времени. По оси абсцисс отложена ширина интервалов по времени (часы), а по оси ординат – число опоздавших (дефектных) документов (десятки документов), попавших в соответствующий временной интервал. Проверка показала соответствие экспериментальных данных нормальному закону распределения. Как видно из графика, время опоздания документов колеблется в пределах 8-90 ч, что свидетельствует о низком уровне ритмичности, слабой технологической дисциплине и недостаточной диспетчеризации решения информационно-вычислительных задач.

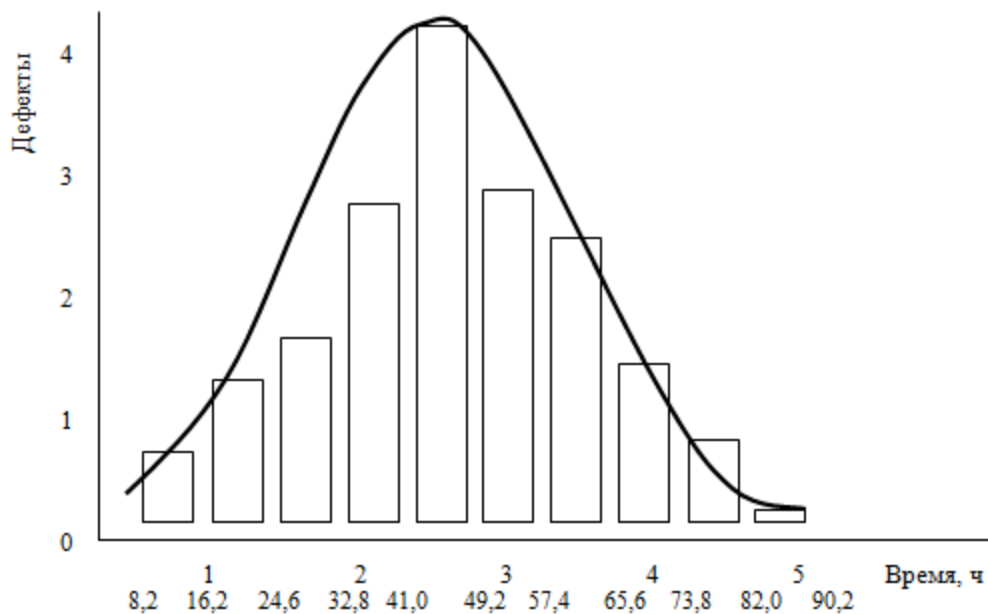


Рисунок 2 – Гистограмма распределения дефектов по своевременности

В соответствии с моделями регрессии обобщенных показателей на основе полученных оценок могут быть составлены две матрицы фиксированных данных – по производительности и себестоимости обработки УБ. Регрессионный анализ зависимости обобщенных показателей от факторов-дефектов обработки выполняется по программе статистического анализа данных [4]. В нашем случае могут быть применены модели множественной линейной или нелинейной регрессий. Для решения практических задач диспетчеризации и управления по таким параметрам, как производительность ИС, себестоимость обработки документов, целесообразно привлекать линейные модели [5]. В результате реализации модели множественной линейной регрессии определяются коэффициенты весомости по факторам-дефектам.

На основе полученных оценок по достоверности, полноте и своевременности производится расчет значений матриц фиксированных по производительности и себестоимости. Например, в таблицах 2 в графах 2, 3, 4 даны предсказывающие переменные (достоверность, полнота, своевременность) в человеко-днях, в графе 5 – прогнозируемая переменная (производительность) – в документо-днях; в графе 1 указывается процент снижения переменных, указанных в графах 2-5.

Данные матриц вводятся в ЭВМ и в результате реализации программы регрессионного анализа получают графики зависимости обобщенных показателей по производительности и себестоимости от дефектов обработки. Например, на графике регрессии по производительности по оси абсцисс отмечено снижение дефектов в процентах, а

по оси ординат соответствующие значения прогнозируемой переменной – производительности ИС (рис. 3). Программа путем нанесения точек указанных значений (табл. 2) и соответствующего уравнения множественной линейной регрессии строит графическую модель прогнозирования обобщенных показателей при снижении числа дефектов по соответствующим значениям процентов.

В практических задачах оперативного управления качеством ИС на основе графиков можно получать экспресс-оценки прогнозируемого снижения (увеличения) по производительности и/или себестоимости. Например, прогнозируемое значение обобщенного показателя по производительности ИС при условии снижения числа дефектов на 5% будет равно ориентировочно 207 документов в день (рис. 3), что вполне согласуется с данными, приведенными в таблице 2.

Для определения обобщенных показателей качества по программе регрессионного анализа могут быть получены коэффициенты регрессии и оценочные величины по производительности (табл. 3): свободный член – 348,14; коэффициент множественной корреляции – 0,9982; коэффициент детерминации 0,9965; F – статистика для проверки нулевой гипотезы: $H_0 : b_1 : b_2 = \dots = 0 - F_0 = 499,7966$; уровень значимости (P -значение) гипотезы H_0 – 0,0000; среднеквадратическое отклонение ошибки – 0,58649.

По полученным данным можно составить уравнение множественной линейной регрессии по производительности ИС:

$$Y^P = 348,14 - 0,18425\bar{x}_1 - 0,072831\bar{x}_2 - 0,038581\bar{x}_3.$$

Здесь свободные члены регрессии и коэффициенты регрессии в правой части уравнения обозначают соответственно базовый показатель производительности и коэффициенты весомости по достоверности, полноте и своевременности. Подставляя значения показателей по достоверности, полноте и своевременности, можно определить фактические или прогнозируемые значения показателей по производительности.

Таблица 2 – Матрица фиксированных данных по производительности

%	Достоверность, человек-день	Полнота, человек-день	Своевременность, человек-день	Производительность, документо-день
0	246,14	794,01	4162,76	200,79
1	243,68	786,07	4121,14	202,05
2	241,22	778,13	4079,52	203,37
3	238,76	770,19	4037,90	204,72
4	236,30	762,25	3996,28	206,07
5	233,84	754,31	3954,66	207,45
...				
19	199,40	643,15	3371,98	228,81
20	196,94	635,21	3330,36	230,53

Для проверки адекватности регрессионной модели обратимся к содержанию полученных оценок (табл. 3). Что касается коэффициентов множественной корреляции и детерминации, уровней значимости нулевой гипотезы, а также среднего квадратического отклонения ошибки, можно предположить, что качество линейного прогноза высокое. Все коэффициенты регрессии укладываются в 95%-ные доверительные границы. Это означает, что их истинные значения при нулевой гипотезе не могут быть отвергнуты при 5%-ном уровне значимости.

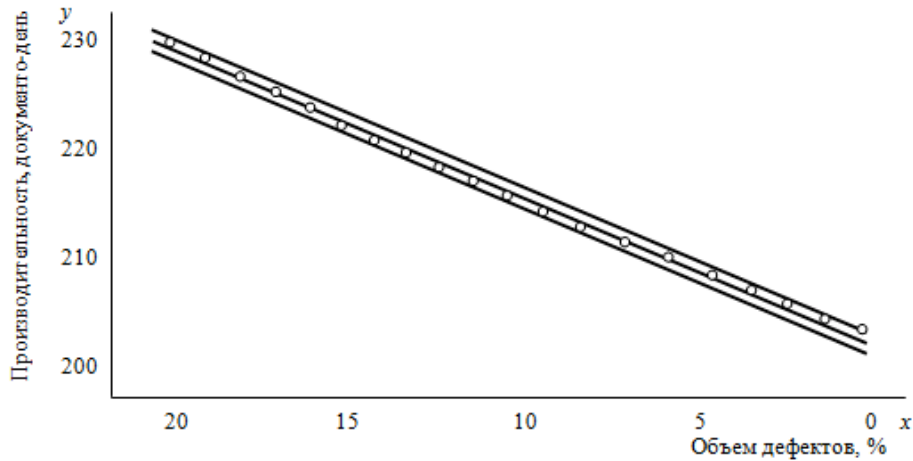


Рисунок 3 – Зависимость производительности ИС от снижения трудоемкости устранения дефектов обработки

Таблица 3 – Коэффициенты регрессии по производительности и данные по их оценке

Признак	Коэффициент регрессии	Стандартное отклонение коэффициента	Уровень значимости нулевой гипотезы (Р-значение)	Т-значение	95%-ные доверительные границы	
					верхняя	нижняя
Достоверность	-0,18425	0,46758	0,3494	0,3940	0,8069	-1,1755
Полнота	-0,072831	0,15696	0,3244	0,4640	0,4055	-0,2599
Своевременность	-0,038581	0,04561	0,2050	0,8458	0,0581	-0,1352

Более четкое заключение можно дать на основе анализа остатков (табл. 4). Анализ параметров показывает, что остатки незначительны как в абсолютном, так и в относительном измерении. Визуальную проверку адекватности модели удобнее всего выполнить по графикам зависимости нормированных остатков от процентов снижения дефектов по производительности. На рисунке 4 по времени наибольший «выброс» наблюдается по 1% и 19% соответственно значениям относительной погрешности нормированных остатков – 0,730 и 1,001 (табл. 4, графа 6, строки 1 и 19). Вместе с тем на графике отсутствует ярко выраженный криволинейный тренд. Наблюдается случайный разброс, свидетельствующий о том, что модель едва ли можно или целесообразно улучшить. Сравнительная смещенность, например, распределения на рисунке 4, довольно мала (близость к оси), что не имеет принципиального значения относительно номинальных величин прогнозируемой переменной по производительности (табл. 4, граф 2).

Таблица 4 – Данные анализа остатков регрессии по производительности

Значение, %	Номинальное значение производительности	Значение прогноза производительности	Остаток	Остаток/SY	Остаток /SS	Относительная погрешность
1	202,0	201,6	0,428	0,049	0,730	0,002118
2	203,4	203,0	0,352	0,040	0,600	0,001732
3	204,7	204,6	0,122	0,014	0,208	0,0005957
4	206,1	206,0	0,076	0,009	0,130	0,0003688
5	207,4	207,6	-0,124	-0,014	-0,212	0,0005987

6	208,8	209,0	-0,130	-0,015	-0,222	0,0006232
7	210,3	210,6	-0,290	-0,033	-0,495	0,001381
8	211,1	211,9	-0,256	-0,029	-0,437	0,001211
9	213,1	213,3	-0,202	-0,023	-0,345	0,0009491
10	214,6	214,9	-0,313	-0,035	-0,533	0,001456
11	216,1	216,3	-0,218	-0,025	-0,372	0,001011
12	217,6	217,9	-0,289	-0,033	-0,492	0,001327
13	219,2	219,3	-0,135	-0,015	-0,230	0,0006142
14	220,7	220,9	-0,165	-0,019	-0,281	0,0007467
15	222,3	222,3	0,019	0,002	0,033	0,00008649
16	223,9	223,9	0,039	0,004	0,067	0,0001743
17	225,5	225,2	0,263	0,030	0,449	0,001167
18	227,2	227,1	0,042	0,005	0,071	0,0001829
19	228,8	228,2	0,587	0,067	1,001	0,002565
20	230,0	229,8	0,198	0,022	0,338	0,0008614

Примечание: SY, SS – стандартное отклонение прогнозируемой переменной и ошибки соответственно

Свободный член регрессионного уравнения при условии нулевого значения предсказываемой переменной отображает по существу возможно достижимое значение, т.е. в нашем случае базовое значение по производительности.

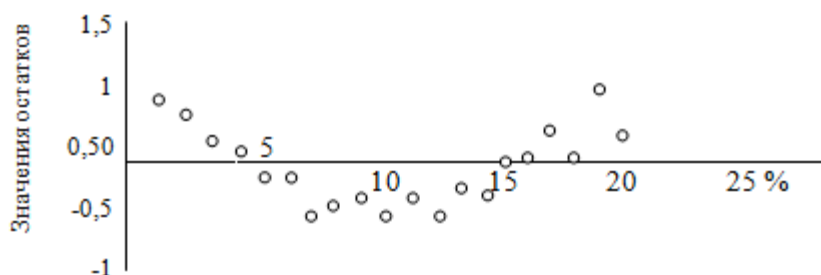


Рисунок 4 – График зависимости нормированного остатка от процента снижения дефектов по производительности ИС

Анализ показателей и оценку качества и эффективности обработки УБ целесообразно начать по принципу «от общего к частному», т.е. от обобщенных показателей к единичным поданным карты анализа и оценки качества ИС (табл. 5). Относительный уровень производительности ИС, равный 0,58, свидетельствует о том, что используются только 58% потенциальных возможностей ИС. Если действия учтенных дефектов обуславливают производительность на уровне 200,79 документов в день, то при условии устранения 100% дефектов производительность может подняться до ее базового значения 348,14 док/день.

Относительный уровень себестоимости обработки составляет 0,32, т.е. ниже, чем соответствующий показатель по производительности. Подобное расхождение можно объяснить сравнительно большим удельным весом в стоимости капитальных вложений в ИС, оборудование, аренду здания и т.д. На величину себестоимости отрицательно действуют факторы-дефекты, при условии устранения которых себестоимость может быть снижена от фактического значения 11,24 руб/док до базового значения 3,61 руб/док.

Удельный вес каждого класса дефектов можно отобразить посредством диаграммы Парето (рис. 5), где по оси абсцисс отложены классы факторов-дефектов по своевременности, полноте и достоверности, а по оси ординат – объемы дефектов в

процентном отношении. На диаграмме видно, что наибольшая доля дефектов приходится на своевременность, а затем на полноту и достоверность. Экспоненты расходятся по всем классам дефектов, но наибольшее расхождение наблюдается по полноте и своевременности. Расхождение по полноте можно объяснить повышенной стоимостью исправления указанного вида дефектов из-за необходимости использования междугородных переговоров с предприятиями по отсутствию в УБ значений показателей и привлекаемых для этого дополнительных финансовых затрат на оплату телефонных услуг. Расхождение по своевременности можно объяснить, в частности, различием в должностных окладах некоторых категорий сотрудников, например, индексировщиков документов.

Таблица 5 – Карта анализа и оценки качества ИС

№ п/п	Показатель	Значение показателя			
		Фактическое		Базовое	Относительное
		Единичное	Групповое		
1	Достоверность	—	0,9992275	0,99999996	0,99992278
	– этап ввода	0,99846	—	—	—
	– этап выдачи	0,999995	—	—	—
2	Полнота	—	0,9866784	0,9999993	0,986679
	– этап приема	0,97344	—	—	—
	– этап выдачи	0,9999169	—	—	—
3	Своевременность	—	0,5715997	0,9999804	0,5716109
	– этап приема от предприятий	0,6930694	—	—	—
	– этап приема от индексировщиков	0,0238096	—	—	—
4	Интегральные	—	0,8525018	0,9999932	0,8527375
	Обобщенные	—	—	—	—
5	– производительность	—	200,79	348,14	0,58
	– себестоимость	—	11,24	3,61	0,32

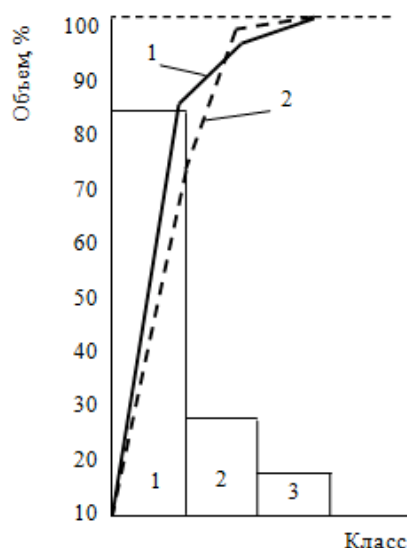


Рисунок 5 – Распределение дефектов на диаграмме Парето:
1 – трудоемкость дефектов; 2 – стоимость дефектов

СПИСОК ЛИТЕРАТУРЫ

1. Еременко В.Т., Фисенко В.Е., Фисун А.П. Методы и модели оценки надежности распределенных систем обмена данными: монография. – Орел: Изд-во «Госуниверситета – УНПК», 2014. – 197 с.
2. Еременко В.Т. Моделирование информационных потоков в сетях передачи данных интегрированных АСУ / С.И. Афонин, В.Т. Еременко, Т.М. Парамохина, Л.В. Кузьмина, Д.А. Плащенков // Информационные системы и технологии. – 2011, № 6 – С. 35-42.
3. Еременко В.Т., Тютякин А.В. Методологические аспекты выбора профилей сбора и обработки данных в системах неразрушающего контроля и диагностики технических объектов // Контроль. Диагностика, 2013. – № 1. – С. 24-31.
4. Еременко В.Т., Полянский И.С., Беседин И.И. Методологические аспекты синтеза оптимальной древовидной структуры в системах сбора и обработки информации // Вестник компьютерных и информационных технологий, 2013. – № 11. – С. 15-21.
5. Еременко В.Т., Парамохин В.М. Метод формирования тестовых комплектов для протоколов безопасности в системах обработки данных // Информационные системы и технологии, 2015. – № 2. – С. 131-137.

Пеньков Николай Геннадьевич

ФГБОУ ВПО «Госуниверситет – УНПК», г. Орел

Аспирант кафедры «Электроника, вычислительная техника, информационная безопасность»

Тел.: 8 906 664 24 42

E-mail: corsa1r_89@mail.ru

Еременко Сергей Владимирович

ФГБОУ ВПО «Госуниверситет – УНПК», г. Орел

Аспирант кафедры «Электроника, вычислительная техника, информационная безопасность»

Тел.: 8 920 287 57 71

E-mail: sv-5771@mail.ru

Ставцева Алена Геннадьевна

ФГБОУ ВПО «Госуниверситет – УНПК», г. Орел

Студентка кафедры «Электроника, вычислительная техника, информационная безопасность»

Тел.: 8 953 611 60 91

E-mail: studentkalenka@mail.ru

N.G. PEN'KOV (*Post-graduate Student of the Department
«Electronics, Computer Engineering, Information Security»*)

S.V. ERYOMENKO (*Post-graduate Student of the Department
«Electronics, Computer Engineering, Information Security»*)

A.G. STAVCEVA (*Student of the Department «Electronics, Computer Engineering, Information Security»
State University – ESPC, Orel*)

MODELING OF THE ACCOUNT AND THE RECEIVING ACCOUNTING FORMS IN THE ELECTRONIC DOCUMENT MANAGEMENT SYSTEM

The article describes the experimental evaluation of the quality of the actual data processing enterprise-level information systems. The algorithm «Automatic error detection and recovery reliability indicator values» for example, accounting records and receiving forms in the electronic document management system. A model for the analysis, evaluation of the quality and effectiveness of the information system.

Keywords: *electronic document management; automated information system; data-computing center; model; options; performance; index; accounting forms.*

BIBLIOGRAPHY (TRANSLITERATED)

1. Eremenko V.T., Fisenko V.E., Fisun A.P. *Metody' i modeli ocenki nadezhnosti raspredelenny'x sistem obmena danny'mi: monografiya.* – Orel: Izd-vo «Gosuniversiteta – UNPK», 2014. – 197 s.
2. Eremenko V.T. *Modelirovanie informacionny'x potokov v setyax peredachi danny'x integrirovanny'x ASU* / S.I. Afonin, V.T. Eremenko, T.M. Paramoxina, L.V. Kuz'mina, D.A. Plashhenkov // *Informacionny'e sistemy' i tehnologii.* – 2011, № 6 – S. 35-42.
3. Eremenko V.T., Tyutyakin A.V. *Metodologicheskie aspekty' vy'bora profilej sbora i obrabotki danny'x v sistemax nerazrushayushhego kontrolya i diagnostiki texnicheskix ob'ektov* // *Kontrol'. Diagnostika,* 2013. – № 1. – S. 24-31.
4. Eremenko V.T., Polyanskij I.S., Besedin I.I. *Metodologicheskie aspekty' sinteza optimal'noj drevovidnoj struktury' v sistemax sbora i obrabotki informacii* // *Vestnik komp'yuterny'x i informacionny'x tehnologij,* 2013. – № 11. – S. 15-21.
5. Eremenko V.T., Paramoxin V.M. *Metod formirovaniya testovy'x komplektov dlya protokolov bezopasnosti v sistemax obrabotki danny'x* // *Informacionny'e sistemy' i tehnologii,* 2015. – № 2. – S. 131-137.

УДК 631.17/629.7.05

В.В. ВЛАСОВ, Н.А. ВЛАСОВА, А.В. ДЕМИДОВ,
Н.В. КАНАТНИКОВ, М.В. СМОЛЯКОВ

К ВОПРОСУ О ПРИМЕНЕНИИ БЕСПИЛОТНЫХ ЛЕТАТЕЛЬНЫХ АППАРАТОВ В СФЕРЕ ТОЧНОГО ЗЕМЛЕДЕЛИЯ

В статье приведен краткий обзор и анализ возможностей применения беспилотных летательных аппаратов для решения задач точного земледелия. Рассмотрена концепция точного земледелия, возможности и проблемы применения беспилотных летательных аппаратов в сфере точного земледелия, а также анализ полей с помощью NDVI-индекса.

Ключевые слова: точное земледелие; беспилотный летательный аппарат; мониторинг; NDVI-индекс.

ВВЕДЕНИЕ

Развитие сельского хозяйства по всему миру переходит на новый этап – точное земледелие, которое предполагает использование новейших технологий, таких, как системы глобального позиционирования ГЛОНАСС, инфракрасные датчики, данные аэрофотосъемки и спутникового мониторинга, а также специальные программы для агроменеджмента на базе геоинформационных систем (ГИС). Использование данных технологий помогает при планировании высева, средств защиты растений от сорняков и вредителей, прогнозировании урожая и расчет норм внесения удобрений. С помощью карт агрохимобследования и урожайности, на которых участки поля идентифицированы с помощью ГНСС-приемников, определяется норма удобрений, что позволяет максимизировать прибыль владельцам фермерских хозяйств и сократить издержки на удобрения, а также влечет к улучшению состояния полей. Все это в свою очередь приводит к сокращению негативного воздействия сельхозпроизводства на окружающую среду за счет более точной оценки потребностей культуры в азотных удобрениях, росту производительности и сокращению затрат на внесение азотных удобрений, а также улучшению планирования сельскохозяйственных операций благодаря более совершенному тайм-менеджменту.

О КОНЦЕПЦИИ ТОЧНОГО ЗЕМЛЕДЕЛИЯ

Концепция точного земледелия позволяет контролировать, что происходит с каждым растением, что приводит к появлению возможности более точного применения химикатов. Проблема обычных небольших сельхозугодий заключается в невозможности осуществления оперативного контроля площади посевных полей. Большинство оценок, производимых в таких случаях, делаются наземным путем при помощи выезда на поля экспертной группы, что приводит к невозможности точной оценки всего масштаба текущего состояния поля. Поэтому для ускорения этого процесса необходимо использовать оперативную аэрофотосъемку, использующую летающих роботов – беспилотные летательные аппараты, которые будут применять пестициды и фунгициды только там, где это действительно необходимо, и в меньших количествах, за счет чего, таким образом, будет предотвращено заражение пищи и окружающей среды химикатами.

Такие дефекты при посеве, как проплешины, гибель урожая после засухи или затопления и других факторов, требуют оперативного контроля, что может предоставить только беспилотная аэрофотосъемка.

Благодаря использованию БПЛА в сельском хозяйстве появляется возможность оперативного оконтуривания сельхозугодий за счет геоданных, полученных в ходе проведения аэрофотосъемки. Иными словами, происходят выделение и оцифровка границ полей по растровым аэрофотоснимкам, при этом растровый снимок, который подвергается векторизации, должен быть правильно откорректирован и должен обладать приемлемым разрешением, в противном случае качество векторизации или оцифровки полей по снимку будет неудовлетворительным. На рисунке 1 отображено применение информационных технологий для поиска и отображения контуров объектов, в данном случае – границ сельхозугодий.



Рисунок 1 – Векторизация и поиск контуров полей

ВОЗМОЖНОСТИ ПРИМЕНЕНИЯ БПЛА В СФЕРЕ ТОЧНОГО ЗЕМЛЕДЕЛИЯ

Беспилотные летательные аппараты, находящие применение до недавнего времени в основном в военной сфере, находят свое место в гражданских областях. Сельское хозяйство без большого объема качественных данных превращается в сложную работу по планированию. Около половины таких материалов, как пестициды, фунгициды и гербициды для растениеводства, оказываются просто бесполезными, так как тратятся в большем количестве, чем нужно, или же находятся не там, где предполагалось опыление, например, в канавах между, а не под самими растениями. Последствия подобной ситуации могут быть самыми плачевными вплоть до полной потери урожая.

Аэрофотосъемочный беспилотный летательный аппарат запускается, взлетает и садится в полуавтоматическом режиме (так называемый автопилот) по загруженному маршруту и, пролетая по заранее спланированному в ГИС маршруту, выполняет цифровую съемку местности, результатом которой являются снимки высокого разрешения на запрограммированных точках по глобальным географическим координатам. Выполнив аэрофотосъемочный маршрут, БПЛА приземляется в точку взлета и передает полученные данные в ходе аэрофотосъемки на наземную станцию. Для каждого снимка получается полный набор цифровой информации – географические координаты центральной точки снимка, высота съемки, полный набор телеметрических данных для переноса и использования в общепринятых ГИС системах, таких, как ArcView и MapInfo. Таким образом, все фотографии являются геопривязанными с помощью одно- или двухчастотного ГНСС приемника; появляется возможность сшивки фотоснимков в единый ортофотоплан

местности. Аэрофотосъемка с БПЛА может заменить спутниковые снимки высокого разрешения для сельского хозяйства. Контроль за посадками с использованием снимков может внести качественные изменения в мониторинг сельскохозяйственных работ, поможет предотвратить лишние затраты на удобрения и воду, к тому же некоторые потребители часто готовы больше платить за продукцию, при выращивании которой использовалась аэрофотосъемка, тем самым отдавая предпочтение экологичному сельскому хозяйству и делая ставку на органическое земледелие.

Применение БПЛА в сельском хозяйстве дает возможности:

- создания электронных карт полей;
- инвентаризации сельхозугодий;
- оценки объема работ и контроль за их выполнением;
- ведения оперативного мониторинга состояния посевов (БПЛА позволяет быстро и эффективно строить карты по всходам);
- определения индекса NDVI (Normalized Difference Vegetation Index – нормализованный вегетационный индекс);
- оценки всхожести сельскохозяйственных культур;
- прогнозирования урожайности сельскохозяйственных культур;
- проверки качества пропашности;
- вести экологический мониторинг сельскохозяйственных земель.

Научный и технический прогрессы позволяют сегодня широко использовать в земледелии современные технологии во время планирования и выполнения агротехнологий. Такими технологиями, несомненно, являются беспилотные летательные аппараты. Точное земледелие включает в себя большое количество элементов, которые делятся на три основных этапа:

- 1) сбор информации о хозяйстве, поле, культуре;
- 2) анализ информации и принятие решения;
- 3) выполнение решений – проведение агротехнологической операции.

Перед тем, как приступить к земледелию, необходимо измерить поля по факту (если они не были измерены), чтобы составить точный план затрат на обрабатываемые площади. В результате измерений поля составляется электронная карта поля.

Преимущества электронной карты поля заключаются в следующем:

- в возможности ведения учета и контроля всех сельскохозяйственных операций за счет базирования на точных данных: площади полей, расстоянии дорог, населенных пунктов и т.д.;
- в помощи проведения полного анализа условий, влияющих на рост растительности на данном поле;
- в оптимизировании процесса производства с целью получения максимального дохода, а также рационального использования в производстве ресурсов;
- в ведении паспортных данных о сельскохозяйственных угодьях с учетом привязки к году урожая;
- в просмотре и анализе тематических карт агрохимического мониторинга полей, возделываемой культуры, вносимых удобрений, урожайности, экономической эффективности культуры и пр.;
- в учете и анализе последствий при различных неблагоприятных погодных условиях и других показателей посредством беспилотной авиации (площади полеглости посевов, вымерзших участков посевов, стадии созревания, засоренность полей);
- в формировании статистических справок и отчетов.

После получения электронной карты поля возможно проводить агрохимическое обследование полей и вносить дополнительную информацию (карты содержания основных элементов N, P, K, Ca, Mg, S, Ph, гумус) о поле в существующую базу данных.

NDVI-ИНДЕКС ДЛЯ ОЦЕНКИ СОСТОЯНИЯ ПОЛЕЙ

Оценка зеленой фитомассы с высоты является перспективным направлением использования NDVI-индекса в сельском хозяйстве. Как известно, отражение растительного покрова в красной и ближней инфракрасной областях электромагнитного спектра тесно связано с его зеленой фитомассой. Для количественной оценки состояния растительности широко применяется вегетационный индекс NDVI (Normalized Difference Vegetation Index), характеризующий плотность растительности и позволяющий растениеводам оценить всхожесть растений и продуктивность сельхозугодий. Индекс рассчитывается как разность значений отражения в ближней инфракрасной и красной областях спектра, деленная на их сумму. В результате значения NDVI меняются в диапазоне от -1 до 1 . Для зеленой растительности отражение в красной области всегда меньше, чем в ближней инфракрасной, за счет поглощения света хлорофиллом, поэтому значения NDVI для растительности не могут быть меньше 0 .

Использование инфракрасной камеры дает дополнительную информацию для анализа состояния сельскохозяйственных полей. Цвета на снимке (рис. 2) отображают состояние растительности. Желтый цвет соответствует мало содержанию азотистых удобрений в почке. Зеленый – нормальное созревание растений. Сдвоенные горизонтальные и вертикальные полосы означают технологические полосы работы сельскохозяйственной техники. Интенсивные цвета на поле, отличающиеся от зеленого, предполагают наличие плотной массы сорняков.

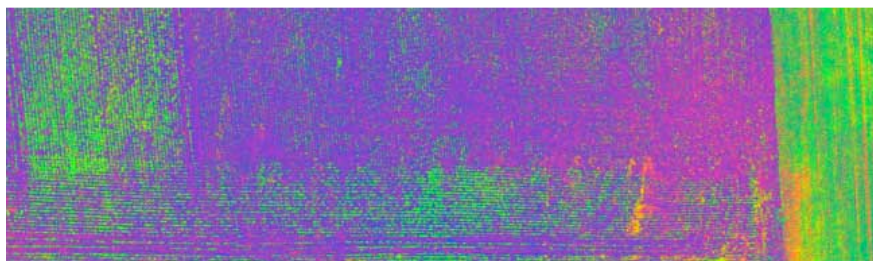


Рисунок 2 – Отображение поля в инфракрасном спектре

Для оценки влажности почвы данных аэрофотосъемки может оказаться недостаточно, поэтому необходимо наличие ИК-камеры на борту БПЛА; также нужно проводить одновременно с полетом анализ почвы влагомером до глубины $1,5$ метров. В случае необходимости построения детальной картины изменения влажности почвы необходимо производить аэрофотосъемки несколько раз в сутки, так как влага на поле меняется в течение дня.

ЗАКЛЮЧЕНИЕ

В настоящее время в сельском хозяйстве широкое применение получают технологии точного земледелия. Они базируются на новом взгляде на сельское хозяйство, в котором сельскохозяйственное поле, неоднородное по рельефу, агрохимическому содержанию питательных веществ, нуждается в применении на каждом участке наиболее эффективных агротехнологий. Технологии точного земледелия направлены на повышение продуктивности, уменьшение себестоимости продукции и сохранение окружающей среды. БПЛА для сельского хозяйства могут решить многие вопросы быстрее и дешевле, чем спутниковый мониторинг. Это означает, что своевременное проведение работ и

аэрофотосъемка в тот же день помогут принять оперативные меры по наиболее актуальным вопросам. При проведении регулярных аэрофотосъемок сельскохозяйственных земель и их постобработке в специализированном программном обеспечении возможно прослеживание динамики изменений в пределах одного и того же поля. И эти данные можно будет точно соотнести с продуктивностью земельных угодий.

СПИСОК ЛИТЕРАТУРЫ

1. Сергеев К. Беспилотники в сельском хозяйстве // Ресурсосберегающее земледелие, 2008. – № 2.
2. Черепанов А.С. Вегетационные индексы // ГеоМатика, 2011. – № 2. – С. 98-102.
3. Бизин И.В. и др. Возможности применения мультироторных беспилотных летательных аппаратов при неуверенном приеме сигналов ГНСС / И.В. Бизин, Ю.В. Василенко, В.В. Власов, А.В. Демидов, Н.В. Канатников, М.В. Смоляков // Информационные системы и технологии, 2014. – № 6(86). – С. 148-153.
4. Bizin I.V. and other. The Use of Multicopters for Traffic Monitoring in Dense Urban Areas / I.V. Bizin, V.V. Vlasov, A.V. Demidov, N.V. Kanatnikov // Conference: Conference on Electronic Governance and Open Society: Challenges in Eurasia, At St. Petersburg, Russian Federation. DOI: 10.1145/2729104.2729124.

Власов Виктор Викторович

ФГБОУ ВПО «Госуниверситет – УНПК», г. Орел
Ассистент кафедры «Информационные системы», программист НОЦ СЦУиСУ
Тел.: 8 953 479 40 79
E-mail: vlasv-viktr@rambler.ru

Власова Наталья Андреевна

Орловский государственный аграрный университет, г. Орел
Аспирант
Тел.: 8 930 861 18 11
E-mail: vlasv90@gmail.com

Демидов Александр Владимирович

ФГБОУ ВПО «Госуниверситет – УНПК», г. Орел
Кандидат технических наук, начальник технического отдела ресурсного центра информатизации образования
Тел.: 8 (4862) 59-46-19
E-mail: a.demidov@ostu.ru

Канатников Никита Владимирович

ФГБОУ ВПО «Госуниверситет – УНПК», г. Орел
Кандидат технических наук, ассистент кафедры «Конструкторско-технологическое обеспечение машиностроительных производств»
Тел.: 8 906 665 00 02
E-mail: nvkanatnikov89@ostu.ru

Смоляков Максим Валерьевич

ФГБОУ ВПО «Госуниверситет – УНПК», г. Орел
Студент кафедры «Информационные системы», программист НОЦ СЦУиСУ
Тел.: 8 909 228 24 24
E-mail: mrdiplodocus@gmail.com

*V.V. VLASOV (Assistant Professor of the Department «Information Systems», Programmer of the Research and Education Center of Special Digital Devices and Control Systems)
State University – ESPC, Orel*

N.V. VLASOVA (*Post-graduate Student*)
Orel State Agrarian University, Orel

A.V. DEMIDOV (*Candidate of Engineering Sciences,*
Head of Technical Department of Resource Centre Informatization of Education)

N.V. KANATNIKOV (*Candidate of Engineering Sciences,*
Assistant Professor of the Department «Design and Technological Support for Engineering Industries»)

M.V. SMOLYAKOV (*Student of the Department «Information Systems», Programmer of the Research and Education Center of Special Digital Devices and Control Systems*)
State University – ESPC, Orel

ON APPLICATION OF USAGE UNMANNED AERIAL VEHICLES IN SPHERE OF PRECISION AGRICULTURE

The article provides a short review and an analysis of possible applications of unmanned aerial vehicles (UAV) in the field of precision agriculture. The article consists an analysis of precision agriculture conception, possibilities of UAV usage in this sphere and analysis of agriculture fields via NDVI-index .

Keywords: *precision agriculture; an unmanned aerial vehicle; monitoring; photogrammetry; NDVI-index.*

BIBLIOGRAPHY (TRANSLITERATED)

1. Sergeev K. *Bespilotniki v sel'skom khozyajstve // Resursosberegayushhee zemledelie, 2008. – № 2.*
2. Cherepanov A.S. *Vegetacionny'e indeksy' // GeoMatika, 2011. – № 2. – S. 98-102.*
3. Bizin I.V. i dr. *Vozmozhnosti primeneniya mul'tirotny'x bespilotny'x letatel'ny'x apparatov pri neuverennom prieme signalov GNSS / I.V. Bizin, Yu.V. Vasilenko, V.V. Vlasov, A.V. Demidov, N.V. Kanatnikov, M.V. Smolyakov // Informacionny'e sistemy' i tehnologii, 2014. – № 6(86). – S. 148-153.*
4. Bizin I.V. and other. *The Use of Multicopters for Traffic Monitoring in Dense Urban Areas / I.V. Bizin, V.V. Vlasov, A.V. Demidov, N.V. Kanatnikov // Conference: Conference on Electronic Governance and Open Society: Challenges in Eurasia, At St. Petersburg, Russian Federation. DOI: 10.1145/2729104.2729124.*

**АЛГОРИТМ ОПТИМИЗАЦИИ
ПРОЦЕССОВ ИНФОРМАЦИОННОГО ОБМЕНА
В МОДУЛЬНЫХ СТРУКТУРАХ СИСТЕМ СБОРА И ОБРАБОТКИ
ДАННЫХ АСУ ГОРОДСКОЙ СЕТИ ГАЗОСНАБЖЕНИЯ**

В статье рассматривается алгоритм оптимизации процессов информационного обмена в модульных структурах систем сбора и обработки данных в АСУ городской сети газоснабжения, базирующийся на использовании замкнутой экспоненциальной модели сети массового обслуживания и отличающаяся возможностью оперирования нечеткими множествами.

Ключевые слова: оптимизация; сети массового обслуживания; нечеткие множества.

Городские сети газоснабжения (ГСГ) представляют собой многообразный комплекс конструкций и сооружений, технических агрегатов и трубопроводов, которые обеспечивают введение и распределение газа среди промышленных, коммунальных и бытовых потребителей с учетом их спроса. Особенности управления городскими сетями газоснабжения связаны со всесторонней деятельностью, а также организацией технологических и финансовых процессов на разных уровнях иерархии, необходимостью предоставления регламентированного количества информационных фондов и важных процессов первичной и вторичной обработки данных газоснабжения.

Функционирование технологических объектов газоснабжения связано со специфическими (по сравнению с традиционными объектами управления) свойствами данных технологических процессов – неоднозначностью реакции технологического объекта на различные управляющие воздействия, многозначностью интерпретации состояний объектов, нестационарностью процессов, обусловленных изменением параметров и переменных состояния технологических объектов во времени, многокритериальностью оценки функционирования объектов газоснабжения. Указанные обстоятельства определяют чрезвычайно сжатый временной интервал реакции для лица, принимающего решение в нештатных ситуациях.

Модульные структуры систем сбора и обработки данных (МСССОД) ГСГ выполняют автоматизированный сбор, обработку физических значений параметров газораспределения в заданных точках контроля. Кроме того, предусматривается не только первичная обработка, но накопление и передача данных должностным лицам ОДУ. В МСССОД газоснабжения используются интеллектуальные датчики, имеющие цифровой интерфейс, память, встроенный микроконтроллер, сетевой адрес и выполняющие автоматическую самодиагностику, калибровку и компенсацию нелинейностей датчика, допускающие настройку пользователем на расстоянии. Данные свойства датчиков уменьшают загруженность управляющих устройств на верхних уровнях и облегчают обслуживание КИП.

Исследование основывается на результатах работ в области теории распределенной обработки данных (М. ван Стеет, В.Г. Хорошевский, Э. Таненбаум), теории надежности процессов обмена данными (С.В. Антонов, И.В. Алексеев, Д.И. Мельников, Ю.С. Злотников), теории алгоритмов (К. Черч, К. Гедель, А. Тьюринг, С. Клини, Э. Пост, А. Колмогоров, А. Марков), теории конечных автоматов (Б.И. Плоткина, В.Б. Кудрявцева, И.Б. Куфарова, А. Гилла), процессов сбора и обработки данных в коммуникационной среде предприятий (И.С. Константинов, А.Д. Иванников, В.Т. Еременко).

Научная новизна данного исследования состоит в получении новых научных результатов:

– математическая модель синтеза, базирующаяся на численных методах многовариантной интеграции, отличающаяся ограничением в виде набора алгоритмических процедур и использованием функции полезности [2, 3];

– методика оптимизации процессов сбора и обработки данных в модульных структурах АСУ ГСГ, базирующаяся на аппарате смешанных сетей массового обслуживания с ограниченным числом классов сообщений и набором дисциплин обслуживания и отличающаяся алгоритмами оптимизации пропускной способности, выбора потоков и их распределения по кратчайшему пути, а также статистического прогнозирования потока данных [14];

– имитационная модель сбора и обработки данных в модульных структурах, основывающаяся на событийном подходе и отличающаяся описанием причинно-следственных связей между существенными событиями [1-9].

МАТЕМАТИЧЕСКАЯ МОДЕЛЬ

В разработанной модели рассматриваются задачи интеграции при разработке и реализации проекта МСССОД газоснабжения, которые выражаются пятью уровнями:

1. Выбор альтернативных вариантов каналов передачи данных газоснабжения.
2. Выбор оборудования сбора данных газоснабжения.
3. Выбор оборудования для обработки данных газоснабжения.
4. Подбор функционального наполнения оборудования среды ГСГ.
5. Формирование набора конечных устройств среды ГСГ.

Каждому уровню в результате интеграции элементов этого уровня соответствует свое множество вариантов [1-9]:

- 1) множество альтернативных вариантов каналов передачи данных газоснабжения

$$\alpha = \{A_1, \dots, A_l, \dots, A_L\} = \{A_l, l = \overline{1, L}\},$$

где $l = \overline{1, L}$ – номер варианта каналов передачи данных газоснабжения 1-го способа создания сети; A_l – вариант создания коммуникационной среды АСУ ГСГ;

- 2) множество вариантов оборудования сбора данных газоснабжения

$$\beta = \{B_1, \dots, B_j, \dots, B_J\} = \{B_j, j = \overline{1, J}\},$$

где $j = \overline{1, J}$ – номер варианта при интеграции компонентов оборудования сбора данных газоснабжения; B_j – вариант интеграции оборудования сбора данных газоснабжения;

- 3) множество вариантов оборудования обработки данных газоснабжения

$$\gamma = \{C_1, \dots, C_t, \dots, C_T\} = \{C_t, t = \overline{1, T}\},$$

где $t = \overline{1, T}$ – индекс комбинации оборудования; C_t – комплект компонентов применяемого оборудования обработки данных газоснабжения;

- 4) множество вариантов функционального наполнения оборудования для p -варианта организации коммуникационной среды АСУ ГСГ

$$\delta = \{D_1, \dots, D_p, \dots, D_P\} = \{D_p, p = \overline{1, P}\},$$

где $p = \overline{1, P}$ – индекс применяемого ПО, D_p – комплект компонентов применяемого ПО;

- 5) множество наборов конечных устройств

$$\varepsilon = \{E_1, \dots, E_q, \dots, E_q\} = \{E_q, q = \overline{1, Q}\},$$

где $q = \overline{1, Q}$ – номер комплекта оконечных устройств коммуникационной среды АСУ ГСГ; E_q – комплект используемого оконечного оборудования коммуникационной среды АСУ ГСГ.

Особенностью решения оптимизационных задач является то, что для численных методов многовариантной интеграции решение всех переменных типов задач происходит в рамках единого алгоритма. При его реализации применяется набор алгоритмических процедур – получение приемлемого решения задачи, доведение до необходимого результата законов распределения альтернативных переменных, получение расчетных оценок оптимального выбора, поиск составляющих векторов безусловных вероятностей, а также получение оценки вариантов по энтропии многовариантной интеграции, ведение учета ограничений общего вида и многокритериальности.

В работе для учета многокритериальности используется одна из разновидностей функции полезности, приводящая к аддитивному критерию оптимизации [1-9]:

$$\Phi(\chi_{mn}, \lambda_{i1}) = \sum_{i_2}^{I_2} \lambda_{i1} (\Psi_{i1}(\chi_{mn}) \rightarrow \max),$$

где λ_{i1} – коэффициенты относительной значимости критериев; $\Psi_{i1}(\chi_{mn})$ – нормированное значение критерия [1-6].

Возможностью вероятностной интерпретации коэффициентов λ_{i1} в рамках схемы ограничения разнообразия множеств обусловлена ориентация на глобальный критерий [1-9].

Рандомизация множества критериев с помощью введения дискретной случайной величины z , номера L критериев Ψ_i , которой являются ее реализациями, лежит в основе процедуры адаптивного многокритериального поиска. Оптимизация по i -ому критерию и получение значений оставшихся критериев происходит на каждом k -ом шаге. Случайный механизм перебора позволяет управлять движением с помощью характеристик случайной величины z . Вероятность p_i может быть использована в качестве одной из таких характеристик [1-6]. Она характеризует возможность на k -ой итерации процесса осуществлять оптимизацию по критерию i . При анализе сформированного вектора $\Psi_i = (\Psi_1, \dots, \Psi_L)$ можно сделать вывод как об окончании процесса и получении Парето-оптимального решения, так и об его продолжении.

Применение математического моделирования позволяет получить МСССОД газоснабжения с минимальными временными и материальными затратами, наилучшими техническими и пользовательскими характеристиками.

МЕТОДИКА ОПТИМИЗАЦИИ

Предлагаемый алгоритм состоит из следующих основных этапов.

Первый этап представлен в виде алгоритма оптимизации пропускной способности в коммуникационной среде модульных структур АСУ ГСГ.

Второй этап представляет собой алгоритм выбора оптимального потока в МСССОД АСУ ГСГ.

Третий этап представляет собой алгоритм оптимизации процесса маршрутизации в МСССОД АСУ ГСГ.

Четвертый этап представлен в виде алгоритма статистического прогнозирования потока данных газоснабжения в модульных структурах их сбора и обработки [14].

Прогноз изменений нагрузки на сеть по времени происходит с применением модели распространения эпидемии. При анализе колебаний трафика в определенный сезон

и изменения трафика в интервале суток существует возможность применения регрессионных моделей и математического аппарата, который предназначен для анализа временных рядов.

Выбор оптимальных потоков и определение оптимальных маршрутов в коммуникационной среде ГСГ реализован на основании критерия средней задержки.

Оптимизация подбора оборудования состоит в следующем: максимальная производительность сети может быть достигнута при выполнении ограничений на стоимость либо минимальная стоимость сети может быть достигнута при выполнении ограничений на производительность [14].

ИМИТАЦИОННАЯ МОДЕЛЬ

В разработанной имитационной модели запрос рассматривается как множество, состоящее из трех элементов – важность запроса, обработчик, тип запроса. В моделируемой МСССОД газоснабжения при формировании запроса определяется степень его важности. При выборе обработчика значения переменной лежат в диапазоне от автоматического выбора, замешательства администратора, ситуационного анализа экспертной системы, поддержки обновлений до выработки управляющих воздействий системы устранения неполадок. Переменная «тип запроса» имеет следующие значения – автоматический выбор, вопрос, обновление программного обеспечения, неполадки программного обеспечения, неполадки оборудования.

Запросы с помощью позиции администратора передаются переходу «поиск», который моделирует поиск и подготовку ответа для МСССОД газоснабжения.

Остальные типы обработчиков функционируют одинаково. В случае поступления запроса в буфер системы устранения неполадок обработчик с помощью интеллектуальной базы ищет соответствие текущего состояния системы эталонному. При обнаружении соответствия текущего состояния описанному инциденту в базе запрос может быть передан на обработку и исполнение. В случае отсутствия описания в интеллектуальной базе данных запрос передается администратору. После этого экспертная система обеспечивает с помощью интерфейса оповещение должностного лица о состоянии запроса, при этом обучение экспертной системы обеспечивает расширение базы инцидентов, накопление соответствующей экспертной информации в базе данных.

Временная модель МСССОД газоснабжения отличается от первичной модели добавленными в нее временными характеристиками переходов. Все представленные переходы, а именно регистрация запросов на входе либо передача запросов между подсистемами, имеют длительность по времени меньше или равную секунде. Наиболее интересны переходы, моделирующие поиск решения инцидента каким-либо обработчиком в системе.

Построенная имитационная модель позволяет провести сравнительный анализ обработчиков по критерию «среднее время обработки запроса» (рис. 1).

Показано, что коэффициент качества обработки запросов всей системой и среднее время обработки запроса увеличивается с увеличением базы знаний экспертной подсистемы, подсистемы устранения неполадок, подсистемы поддержки обновлений.

При оценке качества работы системы определено количество повторно обрабатываемых запросов.

МСССОД газоснабжения рассматривается как множество масштабируемых систем массового обслуживания [10-13]. Над данной системой проведена оценка эффективности ее функционирования с возрастанием числа таких же каналов обработки.

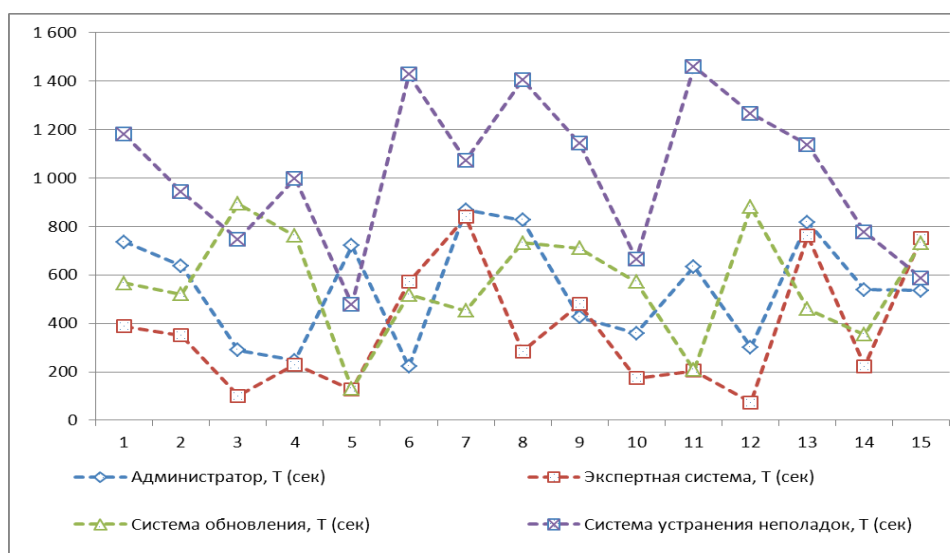


Рисунок 1 – Среднее время обработки запроса каждым обработчиком

Сбор данных модульной системы с исполнительных устройств моделируется в виде одноканальной экспоненциальной СМО [10-13]. В этом случае запросы окажутся соизмеримы обращениям к агрегату системы для завершения операции ввода/вывода; интенсивность таких обращений описывается средним временем выполнения требуемой операции. Предложенная имитационная модель позволила получить функциональные зависимости (рис. 2) – количество запросов в очереди и времени ожидания в очереди от общего количества запросов, поступивших за сутки. Кроме того, представляет определенный интерес функциональная зависимость длины очереди от количества поступивших заявок за сутки.

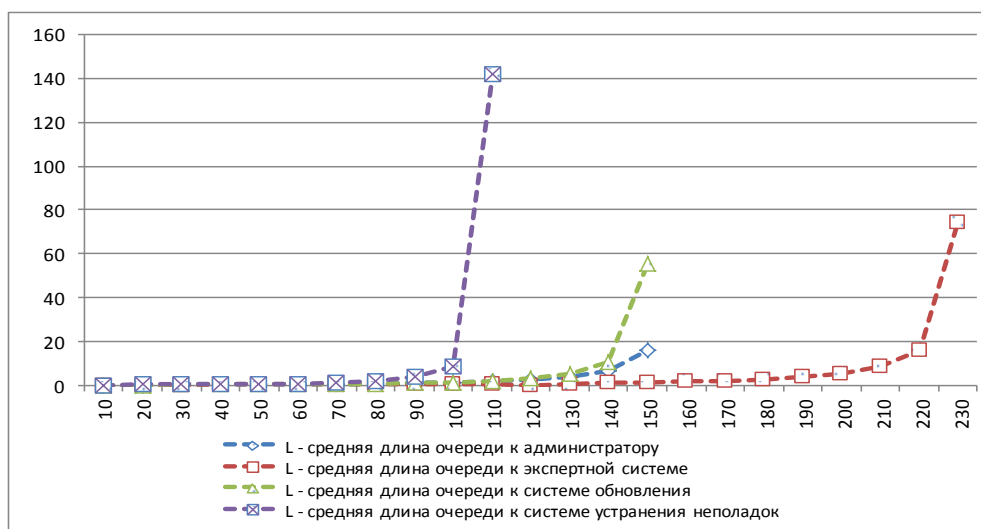


Рисунок 2 – Зависимость очереди запросов от суммарного количества запросов за день

Анализ результатов моделирования показал, что среднее время, затраченное на обработку одного запроса разрабатываемой системы поддержки, сопоставимо с аналогичным параметром у системы с одним обработчиком. Коэффициент качества обработки по сравнению с системами, имеющими 2 и 3 обработчика, несколько ниже. Показано, что среднее время, затраченное на обработку одного запроса разрабатываемой системы поддержки, будет сокращаться с ростом базы данных МСССОД газоснабжения. Определено, что коэффициент качества обработки будет увеличиваться при росте объема базы данных. В

то же время с точки зрения эффективности по количеству успешно обработанных запросов за рабочий день представленная модель показывает существенный выигрыш.

Разработанный алгоритм оптимизации процессов информационного обмена в модульных структурах систем сбора и обработки данных АСУ ГСГ на основе оценки наиболее важных показателей, влияющих на процесс газоснабжения и выбранного критерия оптимизации маршрута доставки данных, позволяет прогнозировать нагрузку на исследуемую модульную структуру при сезонных колебаниях процессов газораспределения.

СПИСОК ЛИТЕРАТУРЫ

1. Еременко В.Т. Оптимизация ресурсов и управление процессами информационного обмена в сетях АСУТП на основе полевых шин / В.Т. Еременко, С.И. Афонин, С.А. Максаков, А.И. Куленич // Вестник компьютерных и информационных технологий, 2011. – № 9. – С. 46-49.
2. Еременко В.Т. Математическая модель оценки производительности беспроводной вычислительной сети АСУ предприятия / В.Т. Еременко, С.И. Афонин, Д.А. Краснов и др. // Информационные системы и технологии, 2011. – № 5. – С. 11-20.
3. Афонин С.И. Моделирование информационных потоков в сетях передачи данных интегрированных АСУ / С.И. Афонин, В.Т. Еременко, Т.М. Парамохина, Л.В. Кузьмина, Д.А. Плащенков // Информационные системы и технологии, 2011. – № 6. – С. 35-42.
4. Еременко В.Т. Методы решения задач распределения информационных потоков в сетях передачи данных предприятия на основе резервирования ресурсов / В.Т. Еременко, С.И. Афонин, Л.В. Кузьмина и др. // Информационные системы и технологии, 2012. – № 1. – С. 78-84.
5. Еременко В.Т., Афонин С.И. Создание теоретических основ автоматизации и построения технологической составляющей АСУ территориально распределенных предприятий // Информационные системы и технологии, 2012. – № 2. – С. 99-105.
6. Еременко В.Т. Рекурсивный алгоритм оценки минимальной величины канального ресурса в сети передачи данных / В.Т. Еременко, Л.В. Кузьмина, Д.А. Плащенков, Д.А. Краснов // Информационные системы и технологии, 2012. – № 4. – С. 97-102.
7. Еременко В.Т. Метод проектирования сетей передачи данных, совместимых с неблокируемой маршрутизацией / В.Т. Еременко, А.И. Офицеров, С.А. Черепков // Вестник компьютерных и информационных технологий, 2012. – № 4. – С. 38-46.
8. Еременко В.Т. Решение задач управления сетевыми ресурсами в условиях динамического изменения конфигурации беспроводной сети АСУП / В.Т. Еременко, Д.В. Анисимов, Д.А. Плащенков, Д.А. Краснов, С.А. Черепков, А.Е. Георгиевский // Информационные системы и технологии, 2012. – № 6. – С. 114-119.
9. Еременко В.Т. Анализ моделей управления трафиком в сетях АСУП на основе технологии MPLS / В.Т. Еременко, С.В. Еременко, Д.В. Анисимов, С.А. Черепков, А.А. Лякишев // Информационные системы и технологии, 2013. – № 1. – С. 106-112.
10. Еременко В.Т. Моделирование пропускной способности сегмента беспроводной сети АСУП на базе стандарта 802.11 / В.Т. Еременко, Д.В. Анисимов, С.А. Черепков, А.А. Лякишев, П.А. Чулахин // Информационные системы и технологии, 2013. – № 2. – С. 82-86.
11. Еременко В.Т. Математическое моделирование беспроводного сегмента вычислительной сети АСУ ПП / В.Т. Еременко, Д.В. Анисимов, Т.М. Парамохина, А.А. Лякишев // Информационные системы и технологии, 2013. – № 3. – С. 67-72.
12. Вишневский В.М., Федотов Е.В. Анализ методов маршрутизации при проектировании сетей пакетной коммутации // 3-rd I.S. Teletraffic Theory and Computing Modeling. – София, 1992.
13. Рындин А.А. Автоматизация проектирования сетей передачи данных распределенных информационно-телекоммуникационных систем / А.А. Рындин, С.В. Сапегин, А.В. Хаустович. – Воронеж: ВГТУ, 2001.

14. Лякишев А.А. Автоматизация процесса информационного обмена в модульных структурах систем сбора и обработки данных АСУ // Сборник материалов II международной научно-практической конференции «Актуальные проблемы естественных и математических наук в России и за рубежом». – Инновационный центр развития образования и науки, Новосибирск, 2015 – С. 112-115.

Орешина Марина Николаевна

ФГБОУ «Московский государственный университет пищевых производств», г. Москва
Доктор технических наук, профессор кафедры «Информатика и управление»
Тел.: +7 (499) 750 01 11, доб. 7242
E-mail: mar-ore@yandex.ru

Лякишев Александр Александрович

ФГБОУ ВПО «Госуниверситет – УНПК», г. Орел
Аспирант кафедры «Электроника, вычислительная техника и информационная безопасность»
Тел.: 8 919 777 06 64
E-mail: vpksearch@yandex.ru

M.N. OREShINA (*Doctor of Engineering Sciences,
Professor of the Department «Informatics and Management»
Moscow State University of Food Production, Moscow*)

A.A. LYaKIShEV (*Post-graduate Student of the Department
«Electronics, Computer Science and Information Security»
State University – ESPC, Orel*)

**ALGORITHM OF OPTIMIZATION OF THE PROCESS INFORMATION
EXCHANGE IN THE MODULAR STRUCTURE OF DATA COLLECTION AND PROCESSING
SYSTEMS IN THE ACS OF A CITY GAS SUPPLY NETWORK**

The article describes the algorithm of optimization of information exchange in the modular structure of data collection and processing systems in the ACS of a city gas supply network based on the using of the exponential model of a closed queuing network and differed by the ability to operate on fuzzy sets.

Keywords: optimization; queuing networks; fuzzy sets.

BIBLIOGRAPHY (TRANSLITERATED)

1. Eremenko V.T. Optimizaciya resursov i upravlenie processami informacionnogo obmena v setyax ASUTP na osnove polevy'x shin / V.T. Eremenko, S.I. Afonin, S.A. Maksakov, A.I. Kulenich // Vestnik komp'yuterny'x i informacionny'x tehnologij, 2011. – № 9. – S. 46-49.
2. Eremenko V.T. Matematicheskaya model' ocenki proizvoditel'nosti besprovodnoj vy'chislitel'noj seti ASU predpriyatiya / V.T. Eremenko, S.I. Afonin, D.A. Krasnov i dr. // Informacionny'e sistemy' iologii, 2011. – № 5. – S. 11-20.
3. Afonin S.I. Modelirovanie informacionny'x potokov v setyax peredachi danny'x integrirovanny'x ASU / S.I. Afonin, V.T. Eremenko, T.M. Paramoxina, L.V. Kuz'mina, D.A. Plashhenkov // Informacionny'e sistemy' iologii, 2011. – № 6. – S. 35-42.
4. Eremenko V.T. Metody' resheniya zadach raspredeleniya informacionny'x potokov v setyax peredachi danny'x predpriyatiya na osnove rezervirovaniya resursov / V.T. Eremenko, S.I. Afonin, L.V. Kuz'mina i dr. // Informacionny'e sistemy' iologii, 2012. – № 1. – S. 78-84.
5. Eremenko V.T., Afonin S.I. Sozdanie teoreticheskix osnov avtomatizacii i postroeniya texnologicheskoy sostavlyayushhej ASU territorial'no raspredelenny'x predpriyatij // Informacionny'e sistemy' iologii, 2012. – № 2. – S. 99-105.
6. Eremenko V.T. Rekursivnyj algoritm ocenki minimal'noj velichiny' kanal'nogo resursa v seti peredachi danny'x / V.T. Eremenko, L.V. Kuz'mina, D.A. Plashhenkov, D.A. Krasnov // Informacionny'e sistemy' iologii, 2012. – № 4. – S. 97-102.

7. Eremenko V.T. Metod proektirovaniya setej peredachi danny'x, sovmestimy'x s neblokiruemoj marshrutizaciej / V.T. Eremenko, A.I. Oficerov, S.A. Cherepkov // Vestnik komp'yuterny'x i informacionny'x texnologij, 2012. – № 4. – S. 38-46.
8. Eremenko V.T. Reshenie zadach upravleniya setevy'mi resursami v usloviyax dinamicheskogo izmeneniya konfiguracii besprovodnoj seti ASUP / V.T. Eremenko, D.V. Anisimov, D.A. Plashhenkov, D.A. Krasnov, S.A. Cherepkov, A.E. Georgievskij // Informacionny'e sistemy' i texnologii, 2012. – № 6. – S. 114-119.
9. Eremenko V.T. Analiz modelej upravleniya trafikom v setyax ASUP na osnove texnologii MPLS / V.T. Eremenko, S.V. Eremenko, D.V. Anisimov, S.A. Cherepkov, A.A. Lyakishev // Informacionny'e sistemy' i texnologii, 2013. – № 1. – S. 106-112.
10. Eremenko V.T. Modelirovanie propusknoj sposobnosti segmenta besprovodnoj seti ASUP na baze standarta 802.11 / V.T. Eremenko, D.V. Anisimov, S.A. Cherepkov, A.A. Lyakishev, P.A. Chupaxin // Informacionny'e sistemy' i texnologii, 2013. – № 2. – S. 82-86.
11. Eremenko V.T. Matematicheskoe modelirovanie besprovodnogo segmenta vy'chislitel'noj seti ASU PP / V.T. Eremenko, D.V. Anisimov, T.M. Paramoxina, A.A. Lyakishev // Informacionny'e sistemy' i texnologii, 2013. – № 3. – S. 67-72.
12. Vishnevskij V.M., Fedotov E.V. Analiz metodov marshrutizacii pri proektirovanii setej paketnoj kommutacii // 3-rd I.S. Teletraffic Theory and Computing Modeling. – Sofiya, 1992.
13. Ryndin A.A. Avtomatizaciya proektirovaniya setej peredachi danny'x raspredelenny'x informacionno-telekommunikacionny'x sistem / A.A. Ry'ndin, S.V. Sapegin, A.V. Xaustovich. – Voronezh: VGTU, 2001.
14. Lyakishev A.A. Avtomatizaciya processa informacionnogo obmena v modul'ny'x strukturax sistem sbora i obrabotki danny'x ASU // Sbornik materialov II mezhdunarodnoj nauchno-prakticheskoj konferencii «Aktual'ny'e problemy' estestvenny'x i matematicheskix nauk v Rossii i za rubezhom». – Innovacionny'j centr razvitiya obrazovaniya i nauki, Novosibirsk, 2015 – S. 112-115.

УДК 004.651

А.Н. ГРИГОРЬЕВА

АЛГОРИТМ ЭФФЕКТИВНОГО ПОСТРОЕНИЯ МНОЖЕСТВА КЛЮЧЕЙ ИНДЕКСА НА ОСНОВЕ МУЛЬТИГРАММ

В данной статье рассмотрены методы индексирования строк, обоснован выбор индекса на основе мультиграмм для ускорения поиска по фрагментам и регулярным выражениям. Для эффективного построения индекса разработана модификация алгоритма построения множества ключей индекса.

Ключевые слова: информационный поиск; индексирование; мультиграмма; бор.

ВВЕДЕНИЕ

В связи с активным развитием информационных технологий потребности в быстром, удобном и качественном поиске постоянно возрастают. Некоторые виды поиска (например, поиск по словам) хорошо проработаны и используются во многих поисковых системах. Однако существует ряд задач, требующих других видов поиска, таких, как поиск по подстроке, по регулярным выражениям или более простым шаблонам и другие.

В качестве практического примера рассмотрим поиск повторяющихся фрагментов. Множество инженерных задач, таких, как рефакторинг, оценка качества кода, поиск и исправление ошибок, требуют выделения синтаксически или семантически схожих фрагментов исходного кода, обычно называемых клонами. Многочисленные исследования показывают, что в больших программных продуктах значительная часть кода (от 5% до 50%) дублируется.

В современных информационных системах объемы данных достаточно велики и требуют эффективных алгоритмов поиска информации. Для обеспечения эффективного поиска используются специальные вспомогательные структуры данных – индексы. Целью данной работы является создание индекса с использованием подстрок в качестве индексированных элементов, который позволяет ускорять поиск по подстрокам и более сложным шаблонам (вплоть до регулярных выражений), а также по сходству. Для этого был решен ряд задач:

- выполнен анализ существующих методов индексирования строк, обоснован выбор мультиграммного индекса в качестве основы разработки;
- выполнена модификация алгоритма построения индекса, позволяющая существенно повысить скорость его работы;
- получены экспериментальные результаты по времени работы и требованию к памяти.

АНАЛИЗ МЕТОДОВ ИНДЕКСИРОВАНИЯ СТРОК

Для индексирования строк разработан целый ряд структур данных – суффиксные деревья, суффиксные массивы, String-B-деревья, индексы на базе подстрок фиксированной и переменной длины (k-gram и multigram) и др. К сожалению, многие структуры хорошо работают лишь в оперативной памяти, но возможность их применения существенно снижается при росте объемов входных данных. Классические же индексы для внешней памяти (разновидности B-деревьев и хеш-таблицы) напрямую для рассмотренных видов

поиска не подходят. Таким образом, разработка индекса для строк и последовательностей во внешней памяти является достаточно актуальной задачей.

Выполним краткий аналитический обзор наиболее известных способов индексирования строк.

СУФФИКСНЫЕ ДЕРЕВЬЯ

Суффиксное дерево – сжатый бор, содержащий только все суффиксы входной строки. Суффиксное дерево позволяет выяснять, входит ли строка s в исходную строку t , за время $O(m)$, где m – длина строки s .

Для построения СД используются разнообразные алгоритмы – Вайнера, Мак-Крейга, Укконена, с различными требованиями к памяти и линейным временем работы. Возможно построение обобщенного суффиксного дерева над несколькими строками.

Суффиксные деревья позволяют эффективно решать весьма широкий ряд. Основным их недостатком являются существенные требования к оперативной памяти (размер дерева даже в лучших программных реализациях превышает размер входных данных более чем в 10 раз). Кроме того, для индексации динамических (изменяющихся) данных недостаточно хранить только само суффиксное дерево, также необходимы вспомогательные структуры, обеспечивающие возможность вносить изменения в индексируемые данные, не перестраивая индекс «с нуля» [1]. Использование СД во внешней памяти весьма затруднительно, так как при поиске требуется произвольное перемещение по дереву, что приводит к большому числу обращений к дисковым страницам.

СУФФИКСНЫЕ МАССИВЫ

Альтернативным вариантом уменьшения затрат памяти является использование структуры данных «суффиксный массив». Суффиксный массив – это набор лексикографически отсортированных суффиксов строк.

Суффиксный массив требует значительно меньше памяти по сравнению с суффиксным деревом. Однако он позволяет решать лишь часть задач обработки строк, которые можно решать с помощью суффиксного дерева.

Существует несколько способов построить массив номеров суффиксов:

- достаточно простой и понятный алгоритм построения со сложностью $O(n \cdot \log n)$;
- алгоритм Каркайнена-Сандерса со сложностью $O(n)$;
- построение из суффиксного дерева $O(n)$.

Поиск подстроки в суффиксном массиве выполняется за $O(m + \log n)$, где n – длина входных данных, m – длина подстроки.

Использование данной структуры во внешней памяти затруднительно как с точки зрения скорости построения, так и поиска, так как обе операции требуют большого числа обращений к дисковым страницам.

СТРОКОВЫЕ В-ДЕРЕВЬЯ

Строковое В-дерево (String-B-Tree) представляет собой структуру для хранения текстовых данных во внешней памяти: это комбинация В-дерева и бора Патриции для индексации внутренних узлов, что позволяет эффективно выполнять поиск подстрок.

Основные достоинства:

- во всех случаях полезное использование пространства вторичной памяти составляет свыше 50%;
- произвольный доступ к записи реализуется посредством малого количества обращений к дисковым страницам;

– в среднем достаточно эффективно реализуются операции включения и удаления записей; при этом сохраняется естественный порядок ключей с целью последовательной обработки, а также соответствующий баланс дерева для обеспечения быстрой произвольной выборки.

К-ГРАММНЫЙ И МУЛЬТИГРАММНЫЙ ИНДЕКСЫ

Индекс на основе k -грамм состоит в том, чтобы включать в индекс каждую последовательность из k символов в тексте (k -грамму). Наиболее часто используемым вариантом является триграммный индекс ($k = 3$).

Однако триграммный индекс позволяет ускорять далеко не все запросы поиска. При увеличении же длины k подход становится слишком дорогостоящим как по затратам времени, так и по объему памяти.

В работе [2] было предложено включать в такой индекс граммы различной длины, но лишь те, которые будут наиболее полезны для поиска. Полученная структура получила название мультиграммного индекса. Она позволяет добиться весьма низких требований к памяти при высокой скорости поиска и подходит для индексирования данных во внешней памяти.

Основоположниками мультиграммного индекса были Junghoo Cho и Sridhar Rajagopalan. Они разработали его с целью ускорения запросов поиска по регулярным выражениям.

Индекс состоит из множества ключей и списков вхождений, которые для определенного ключа ссылаются на документы, в которых этот ключ встречается. Все ключи представляют собой подстроки с низкой селективностью (то есть встречающиеся в небольшой доле документов).

Селективность граммы x равна

$$sel(x) = M(x) / N, \quad (1)$$

где N – количество входных блоков данных (документов); $M(x)$ – число блоков, содержащих x .

Грамма x является c -полезной, если

$$sel(x) \leq c \quad (2),$$

где

$$0 \leq c \leq 1 \quad (3).$$

Константа c выбирается заранее, исходя из конкретной задачи. Будем далее называть c -полезные граммы просто полезными.

В своем подходе J. Cho и S. Rajagopalan предложили хранить в индексе только полезные граммы. Кроме того, ни одна граммa не должна являться суффиксом или префиксом другой, такое множество грамм они назвали пресуффиксно-свободным. Такой подход позволяет сделать множество ключей индекса довольно компактным – оно помещается в оперативную память даже при входных данных объемом десятки и сотни Гб. Этот факт является основной причиной выбора нами данного метода индексирования.

Однако имеются некоторые замечания к описанному в статье [2] алгоритму построения индекса:

– во-первых, в ней не указывается, каким образом осуществляется построение списков вхождений;

– во-вторых, в алгоритме не уточняется, каким именно образом нужно хранить в памяти множество грамм. От выбора же структуры данных может существенно зависеть производительность;

– в-третьих, описанный алгоритм требует выполнения достаточно большого

количества проходов по входным данным – оно равно длине наибольшей граммы в результирующем множестве. Это приводит к существенным временным затратам.

Способ эффективного построения списков вхождений при помощи алгоритма Ахо-Корасик был описан в статье [3]. В данной же статье предложена модификация алгоритма построения множества ключей индекса, позволяющая существенно повысить его скорость.

БЫСТРОЕ ПОСТРОЕНИЕ МНОЖЕСТВА КЛЮЧЕЙ ИНДЕКСА

В таблице 1 в третьем столбце показано экспериментальное время построения мультиграммного индекса с использованием оригинального алгоритма. Для экспериментов были использованы реальные данные разного размера, взятые из двух источников: программный код решений студентов ВоГУ (г. Вологда) из системы «Дистанционный практикум по программированию» [4] и подмножество набора данных «Web-страницы коллекции РОМИП» [5].

Таблица 1 – Экспериментальное время построения мультиграммного индекса с использованием оригинального алгоритма

Входные данные для экспериментов	Размер входных данных	Время построения ядра индекса, сек
Решения студентов ВоГУ (г. Вологда) из системы «Дистанционный практикум по программированию»	20000 файлов 16Мб	63
Набор данных «Web-страницы коллекции РОМИП»	6942 файлов 77Мб	402
Набор данных «Web-страницы коллекции РОМИП»	110998 файлов 924Мб	3408

Для хранения множества ключей индекса и их селективности при построении использовался класс `std::map<std::string, int>` из стандартной библиотеки языка C++. Как можно заметить, уже при размере входных данных порядка одного гигабайта время построения индекса становится достаточно существенным. Таким образом, повышение эффективности операции построения индекса является достаточно актуальной задачей.

В статье [6] для хранения грамм индекса в оперативной памяти предложено использовать структуру данных «луч» (другое название – «бор»). Данная структура позволяет эффективно выполнять поиск подстрок в заданном множестве – время поиска подстроки пропорционально ее длине и не зависит от размера множества, в котором осуществляется поиск.

Основной проблемой при использовании структуры данных «луч» является организация быстрого поиска нужного ребра, выходящего из узла, в случае использования алфавита большого размера. Обычно с этой целью луч комбинируют с хеш-таблицей или деревом поиска, что не только усложняет программную реализацию, но и замедляет операции над лучом.

В данной статье предложено использовать подход, описанный в [7]. Каждый символ исходного алфавита представляется в виде строки из двухбитных символов, то есть мы переходим к новому алфавиту размера 4 (выбор константы 4 обоснован в [7]). Например, для обычной 8-битной кодовой таблицы каждый символ входных данных преобразуется в строку из четырех двухбитных символов. Хотя данный подход изначально был предложен для разреженных суффиксных деревьев, но он вполне подходит и в этом случае.

В результате из каждого узла луча может выходить максимум 4 ребра. Для их представления достаточно в каждом узле хранить массив из четырех указателей. Извлечение

из символа необходимой группы битов эффективно выполняется битовыми машинными операциями. Результаты данной модификации приведены в таблице 2.

Таблица 2 – Результаты применения модифицированного алгоритма

Входные данные для экспериментов	Размер входных данных	Время построения ядра индекса, сек (оригинальный алгоритм)	Время построения ядра индекса, сек (хранение ключей в луче)	Время построения ядра индекса, сек (сокращение числа проходов)	Количество мультиграмм, их максимальная длина
Решения студентов ВоГУ (г. Вологда) из системы «Дистанционный практикум по программированию»	20000 файлов 16Мб	63	27	26	7306,8
Набор данных «Web-страницы коллекции РОМИП»	6942 файлов 77Мб	402	134	91	22344,10
Набор данных «Web-страницы коллекции РОМИП»	110998 файлов 924Мб	3408	3247	1576	37082,7

При сравнении третьего и четвертого столбцов таблицы 2, видно, что описанный подход существенно ускоряет построение индекса над небольшими данными. Однако с ростом объема входных данных разница уменьшается, поскольку время работы начинает в основном зависеть от количества проходов по входным данным, хранящимся во внешней памяти. Следовательно, для дальнейшего повышения производительности необходимо добиться снижения количества проходов по данным. В данной статье предлагается уменьшить вдвое количество проходов по входным данным с помощью слияния пар соседних проходов в один. Рассмотрим изменения в алгоритме, которые необходимо для этого сделать. Оригинальный алгоритм построения множества ключей индекса выглядит следующим образом [2]:

```

k=1, expand={.} // . – пустая строка
while (expand не пусто)
  k-grams := все k-граммы в исходных данных,
             где (k-1)-префикс ∈ expand
  expand := {}
  For each gram x in k-grams
    If sel(x) ≤ c Then
      insert(x, index)
    Else
      expand := expand ∪ {x}
  k:=k+1
    
```

На каждом проходе строится множество k-grams – все граммы длины k из исходных данных, у которых префикс длины (k-1) принадлежит множеству expand. В данной статье предложено добавить в данный алгоритм построение еще одного множества – (k+1)-grams – все граммы длины (k+1) из исходных данных, у которых префикс длины (k-1) принадлежит множеству expand. При этом необходимо выполнение следующего условия: еслиграмма во

множестве k -grams является полезной, то следует удалить все граммы из множества $(k+1)$ -grams, для которых она является префиксом.

Ниже приведен модифицированный алгоритм построения индекса.

```

k=1, expand={.} // . – пустая строка
while (expand не пусто)
  k-grams := все k-граммы в исходных данных,
             где (k-1)-префикс ∈ expand
  (k+1)-grams := все (k+1)-граммы в исходных данных,
                 где (k-1)-префикс ∈ expand
  expand := {}

  For each gram x in k-grams
    If sel(x) ≤ c Then
      insert(x, index)
      Удалить из (k+1)-grams все граммы с префиксом x

  For each gram x in (k+1)-grams
    If sel(x) ≤ c Then
      insert(x, index)
    Else
      expand := expand ∪ {x}

k:=k+2

```

В таблице 2 в пятом столбце приведено время работы программной реализации данного алгоритма. Сравнивая со значениями в предыдущем столбце, можно сделать вывод, что существенный прирост производительности наблюдается, начиная лишь с некоторого размера обрабатываемых данных. Это вполне объяснимо: данные небольшого объема кэшируются в оперативной памяти, при этом на скорость работы влияет общее количество операций, а не число проходов. Однако при увеличении размера входных данных время работы начинает определяться преимущественно количеством операций обращения к внешней памяти, которое сократилось примерно вдвое.

Рассматривая полученный алгоритм, может возникнуть следующий вопрос: можно ли получить еще больший прирост производительности путем сливания не пар, а троек (и большего количества) проходов по входным данным, В принципе, это возможно. Однако проблемой является ограниченный объем оперативной памяти для хранения временных данных. Мощность множества $(k+1)$ -grams примерно в $|\Sigma|$ раз больше, чем мощность множества k -grams, где $|\Sigma|$ – размер входного алфавита. Аналогично мощность множества $(k+2)$ -grams будет во столько же раз больше мощности множества $(k+1)$ -grams и т.д. Такой экспоненциальный рост ограничивает количество сливаемых проходов весьма небольшим значением и для большинства практических задач 2 будет разумным выбором. Однако в конкретном практическом случае используется алфавит малого размера и количество сливаемых проходов выгодно увеличить. Более того, целесообразно реализовать алгоритм, который автоматически выбирает данный параметр, исходя из размера алфавита и доступного объема оперативной памяти, что является перспективой для дальнейшей работы.

ЗАКЛЮЧЕНИЕ

Индекс на основе подстрок переменной длины с низкой селективностью – мультиграмм – является перспективной структурой для работы с данными большого объема во внешней памяти.

Для повышения эффективности алгоритма построения индекса предложено

использовать модифицированный бор для хранения множества ключей, а также уменьшение числа проходов по входным данным за счет слияния пар соседних проходов в один.

Корректность и эффективность модифицированного алгоритма подтверждена проведенными вычислительными экспериментами на различных видах и объемах входных данных.

СПИСОК ЛИТЕРАТУРЫ

1. Чернов А.Ф. Анализ и разработка индекса для поиска последовательностей элементов произвольного типа по их фрагментам в реляционных базах данных (диссертация на соискание ученой степени кандидата технических наук) // ФГБОУ ВПО «Санкт-Петербургский государственный электротехнический университет «ЛЭТИ» им. В.И. Ульянова (Ленина)», 2014
2. Cho J., Rajagopalan S. A Fast Regular Expression Indexing Engine // In Proc. of the 18th Intl. Conf. on Data Engineering (ICDE'02), 2002.
3. Андрианов И.А., Григорьева А.Н. Алгоритм быстрого построения списков вхождений для индекса на основе мультиграмм // Системы управления и информационные технологии (научно-технический журнал); Москва-Воронеж. – Издательство «Научная книга», 2014. – № 4(58). – С. 29-34.
4. Система «Дистанционный практикум по программированию» [Электронный ресурс]. – URL: http://atpp.vstu.edu.ru/cgi-bin/arh_problems.pl.
5. Web-страницы коллекции РОМИП [Электронный ресурс]. – URL: <http://www.romip.ru>.
6. Андрианов И.А., Минин В.Н. Применение декомпозиции на основе подстрок при построении индексов для поиска по сходству // Сборник работ участников конкурса «Интернет-математика 2007».
7. Андрианов, И.А., Чернов А.Ф. Индексирование и поиск в последовательностях для больших баз данных; монография. – Вологда, ВоГУ, 2013. – 165 с.

Григорьева Анна Николаевна

ФГБОУ ВПО «Вологодский государственный университет», г. Вологда

Аспирант

Тел.: 8 921 232 27 10

E-mail: gran80@mail.ru

*A.N. GRIGOR'EVA (Post-graduate Student)
Vologda State University, Vologda*

THE ALGORITHM FOR EFFECTIVE KEYSSET CONSTRUCTION FOR A MULTIGRAM-BASED INDEX

This article examines the methods of indexing of strings. The index based on the multigrams is selected, since it is suitable for acceleration of the search for fragments and regular expressions. In order to effectively build an index, the modification of the algorithm for constructing the set of index keys is developed.

Keywords: *information retrieval; indexing; multigram; trie.*

BIBLIOGRAPHY (TRANSLITERATED)

1. Chernov A.F. Analiz i razrabotka indeksa dlya poiska posledovatel'nostej e'lementov proizvol'nogo tipa po ix fragmentam v relyacionny'x bazax danny'x (dissertaciya na soiskanie uchenoj stepeni kandidata texnicheskix nauk) // FGBOU VPO «Sankt-Peterburgskij gosudarstvennyj e'lektrotexnicheskij universitet «LETI» im. V.I. Ul'yanova (Lenina)», 2014
2. Cho J., Rajagopalan S. A Fast Regular Expression Indexing Engine // In Proc. of the 18th Intl. Conf. on Data Engineering (ICDE'02), 2002.

3. Andrianov I.A., Grigor'eva A.N. Algoritm by'strogo postroeniya spiskov vvozhdenij dlya indeksa na osnove mul'tigramm // *Sistemy' upravleniya i informacionny'e tehnologii (nauchno-texnicheskij zhurnal)*; Moskva-Voronezh. – Izdatel'stvo «Nauchnaya kniga», 2014. – № 4(58). – S. 29-34.
4. Sistema «Distancionny'j praktikum po programmirovaniyu» [E'lektronny'j resurs]. – URL: http://atpp.vstu.edu.ru/cgi-bin/arh_problems.pl.
5. Web-stranicy' kollekcii ROMIP [E'lektronny'j resurs]. – URL: <http://www.romip.ru>.
6. Andrianov I.A., Minin V.N. Primenenie dekompozicii na osnove podstrok pri postroenii indeksov dlya poiska po sxodstvu // *Sbornik rabot uchastnikov konkursa «Internet-matematika 2007»*.
7. Andrianov, I.A., Chernov A.F. Indeksirovanie i poisk v posledovatel'nostyax dlya bol'shix baz danny'x; monografiya. – Vologda, VoGU, 2013. – 165 s.

Н.А. ОРЕШИН, С.А. ЧЕРЕПКОВ, В.С. ШУМИЛИН, А.Т. ТАШМАТОВ

ОДНОПРОДУКТОВАЯ ДВУХПОЛЮСНАЯ ПОТОКОВАЯ МОДЕЛЬ ТРАНСПОРТНОЙ СЕТИ

В статье рассматриваются математические аспекты оценки пропускной способности групп направлений транспортной сети на основе использования однопродуктовой двухполюсной потоковой модели.

Ключевые слова: пропускная способность; потоковая модель; транспортная сеть.

Эффективность использования существующей транспортной сети оценивается, с одной стороны, суммарной задействованной канальной емкостью, необходимой для обеспечения требований направлений в каналах связи, с другой стороны – степенью практической реализации пропускной способности сети. При этом чем больше в анализируемой сети реализуется ее пропускная способность, тем выше эффективность ее использования.

При решении задачи оценки пропускной способности будем использовать однопродуктовую двухполюсную модель, представленную на рисунке 1.

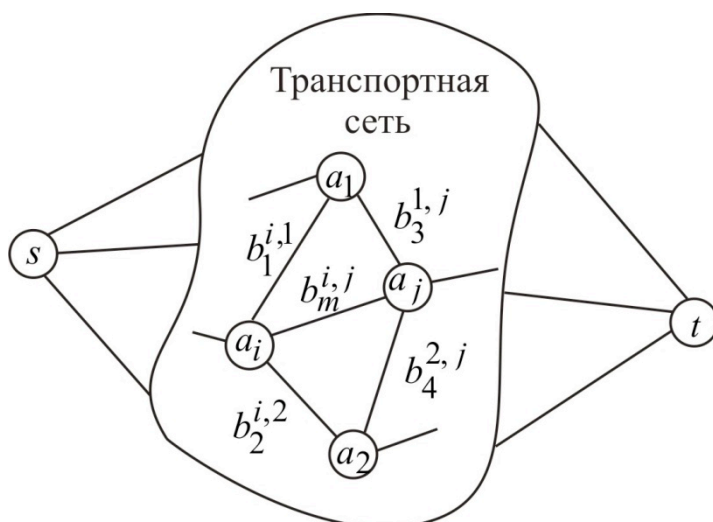


Рисунок 1 – Однопродуктовая двухполюсная модель транспортной сети

Важным компонентом этой модели транспортной сети является поток, который математически может быть определен в формах узлы-дуги и дуги-цепи.

Потоком, протекающим по m -ому ребру $b_m^{i,j}$ от узла a_i к узлу a_j в форме узлы-дуги, назовем целочисленную функцию $f_m^{b_m^{i,j}}$, определенную на множестве дуг $\vec{b}_m^{i,j} \in \vec{B}$, $\vec{B} = \{ \vec{b}_m^{i,j} \mid m = \overline{1, M}; i, j = \overline{1, N} \}$ орграфа $\vec{G} = (A, \vec{B})$, если:

$$\sum_{b_m^{i,j} \in B_1^i} f_m^{b_m^{i,j}} - \sum_{b_m^{i,j} \in B_2^j} f_m^{b_m^{i,j}} = 0, \forall a_j \in A;$$

$$0 \leq f_m^{b^{i,j}} \leq c^{i,j}, \forall \bar{b}_m^{i,j} \in \bar{B}, \quad (1)$$

где B_1^i – множество дуг, по которым поток вытекает из a_i -ого узла; B_2^i – множество дуг, по которым поток втекает в a_i -ый узел; $f_m^{b^{i,j}}$ – поток, протекающий по дуге $\bar{b}_m^{i,j}$ от вершины a_i к вершине a_j (однопродуктовый поток по дуге); $c^{i,j}$ – емкость (пропускная способность) дуги $\bar{b}_m^{i,j}$; $C = \left\{ c_m^{i,j} \mid m = \overline{1, M}; i, j = \overline{1, N} \right\}$ – множество пропускных способностей дуг $\bar{b}_m^{i,j} \in \bar{B}$, соединяющих вершину a_i с вершиной a_j ; $\bar{B} = \left\{ \bar{b}_m^{i,j} \mid m = \overline{1, M}; i, j = \overline{1, N} \right\}$ – множество дуг орграфа; N – количество вершин графа; M – количество ребер графа.

Потоком в форме дуги-цепи назовем целочисленную функцию $f^{\pi_p^{s,t}}$, определенную на множестве дуг $\bar{b}_m^{i,j} \in \bar{B}$, $\bar{B} = \left\{ \bar{b}_m^{i,j} \mid m = \overline{1, M}; i, j = \overline{1, N} \right\}$ орграфа $\bar{G} = (A, \bar{B})$, если для каждого пути этого орграфа $\pi_p^{s,t} = (b^{s,l}, b^{i,j}, \dots, b^{m,t})$, $p = \overline{1, P}$ выполняются условия:

$$f^{\pi_p^{s,t}} \leq \min \left[c^{i,j} \in C / b^{i,j} \in \pi_p^{s,t} \right]; \quad (2)$$

$$f_m^{b^{i,j}} = y_p^{i,j} \cdot f^{\pi_p^{s,t}}; \quad (3)$$

$$y_p^{i,j} = \begin{cases} 1, & \text{если } \bar{b}_m^{i,j} \in \pi_p^{s,t}; \\ 0, & \text{если } \bar{b}_m^{i,j} \notin \pi_p^{s,t}; \\ -1, & \text{если } \bar{b}_m^{j,i} \in \pi_p^{s,t}; \end{cases} \quad (4)$$

$$f_m^{b^{i,j}} = -f_m^{b^{j,i}}, \quad (5)$$

где $f^{\pi_p^{s,t}}$ – поток, протекающий по p -ому пути $\pi_p^{s,t} = (b^{s,i}, b^{i,j}, \dots, b^{m,t})$ от вершины s к вершине t (поток по пути); $f_m^{b^{i,j}}$ – поток, протекающий по дуге $\bar{b}_m^{i,j}$, входящей в p -ый путь, от вершины s к вершине t (поток по дуге); $\min \left[c^{i,j} \in C / b^{i,j} \in \pi_p^{s,t} \right]$ – минимальное значение пропускных способностей ребер, входящих в p -ый путь $\pi_p^{s,t} = (b^{s,i}, b^{i,j}, \dots, b^{m,t})$.

Величину, равную минимальной пропускной способности ребра из множества ребер, образующих путь, по которому протекает поток величиной $\max f^{\pi_p^{s,t}}$, назовем пропускной способностью пути и обозначим символом $c^{\pi_p^{s,t}}$, то есть $c^{\pi_p^{s,t}} = \max f^{\pi_p^{s,t}}$.

Для решения задачи оценки пропускной способности групп направлений транспортной сети ее целесообразно представить в виде однопродуктовой двухполюсной потоковой модели

$$G = \left(A, B, s, t, C, v, \Pi^{s,t}, f^{s,t}(\chi) \right),$$

где $A = \{a_n / n = \overline{1, N}\}$ – множество вершин (узлов) модели; N – количество вершин (узлов) модели; $B = \{b_m^{i,j} / m = \overline{1, M}; i, j = \overline{1, N}\}$ – множество ветвей модели; M – количество ветвей модели; s – исток модели; t – сток модели; $C = \{c_m^{i,j} / m = \overline{1, M}; i, j = \overline{1, N}\}$ – множество пропускных способностей ребер $b_m^{i,j} \in B$, соединяющих вершину a_i с вершиной a_j ; v – требование в каналах направления связи, образованного между истоком s и стоком t ; $\Pi^{s,t} = \{\pi_p^{s,t} / p = \overline{1, P}\}$ – множество путей, организуемых в потоковой модели; P – мощность множества путей; $f^{s,t}(\chi)$ – однопродуктовый двухполюсный поток, являющейся целочисленной функцией от структурно-метрического разложения $\chi(A, B, C, s, t, v)$, удовлетворяющий следующим условиям:

1. Поток $f^{\pi_p^{s,t}}$, протекающий по p -тому пути $\pi_p^{s,t} \in \Pi^{s,t}$, не превышает минимальной пропускной способностью ветви $b_m^{i,j}$, входящей в этот путь:

$$\forall \pi_p^{s,t} \in \Pi^{s,t} \\ f^{\pi_p^{s,t}} \leq \min [c_m^{i,j} \in C / b_m^{i,j} \in \pi_p^{s,t}]. \quad (6)$$

2. Направление протекания потока в $b_m^{i,j}$ ветви по пути $\pi_p^{s,t}$ удовлетворяет следующим соотношениям:

$$\forall b_m^{i,j} \in \pi_p^{s,t} \\ f^{b_m^{i,j}} = y_p^{i,j} \cdot f^{\pi_p^{s,t}}; \quad (7)$$

$$y_p^{i,j} = \begin{cases} 1, & \text{если } \vec{b}_m^{i,j} \in \pi_p^{s,t}; \\ 0, & \text{если } \vec{b}_m^{i,j} \notin \pi_p^{s,t}; \\ -1, & \text{если } \vec{b}_m^{j,i} \in \pi_p^{s,t}. \end{cases} \quad (8)$$

3. Величина потока $f^{s,t}(\chi)$ равна сумме величин потоков $f^{b_m^{i,j}}$ по всем путям $\pi_p^{s,t} \in \Pi^{s,t}$, организуемых в потоковой модели

$$f^{s,t}(\chi) = \sum_{p=1}^P f^{b_m^{i,j}}. \quad (9)$$

4. Величина потока $f_m^{i,j}$ в ветви $b_m^{i,j}$ равна сумме величин потоков $f_m^{i,j}$ по всем путям, включающим в себя эту ветвь

$$\forall b_m^{i,j} \in B$$

$$f_m^{i,j} = \sum_{\pi_p^{s,t} \in \Pi^{i,j}} f \pi_p^{s,t}, \quad (10)$$

где $\Pi^{i,j}$ – множество путей, содержащих ребро $b_m^{i,j}$.

5. Величина потока в ветви не превышает величины пропускной способности этой ветви

$$\forall b_m^{i,j} \in B$$

$$f_m^{i,j} \leq c_m^{i,j}. \quad (11)$$

ВЫВОДЫ

Данная однопродуктовая двухполюсная потоковая модель транспортной сети обладает следующими достоинствами:

- сравнительная простота моделирования;
- разработанность математического аппарата для ее исследования.

СПИСОК ЛИТЕРАТУРЫ

1. Еременко В.Т. Основы построения информационно-телекоммуникационных систем: учебное пособие. – Орел: ОрЮИ МВД России, 1999. – Ч. 2. – 129 с.
2. Еременко В.Т. Теория информации и информационных процессов: монография / В.Т. Еременко, И.С. Константинов, А.В. Коськин, В.А. Лобанова и др.; под ред. д.т.н. В.Т. Еременко, д.т.н. А.П. Фисуна. – Орел: ОГУ, ОрелГТУ, 2008. – 478 с. – 500 экз.
3. Филипс Д., Гарсиа-Диас А. Методы анализа сетей: пер. с англ. под ред. Сушкова Б.Г. – М.: Мир, 1984. – 496 с.
4. Форд А., Фалкерсон Д. Потоки в сетях: пер. с англ. Вайнштейна И.А. – М.: Мир, 1966. – 276 с.
5. Берж К. Теория графов и ее применение. – М.: Иностранная литература: пер. с англ., 1962. – 320 с.
6. Басакер Р., Саати Т. Конечные графы и сети: пер. с англ. под ред. Теймана А.И. – М.: Наука, 1973. – 368 с.
7. Фрэнк Г., Фриш И. Сеть связи и потоки: под ред. Поспелова. – М.: Связь, 1978. – 448 с.

Николай Алексеевич Орешин

Академия ФСО России, г. Орел
Кандидат технических наук, профессор
Тел.: 8 (4862) 54-96-91

Сергей Анатольевич Черепков

Академия ФСО России, г. Орел
Сотрудник
Тел.: 8 (4862) 54-94-78

Вячеслав Сергеевич Шумилин

Академия ФСО России, г. Орел
Сотрудник
Тел.: 8 (4862) 54-96-91

Артур Тахирович Ташматов
Академия ФСО России, г. Орел
Сотрудник

N.A. OREShIN (*Candidate of Engineering Sciences, Professor*)

S.A. ChEREPKOV (*Employee*)

V.S. ShUMILIN (*Employee*)

A.T. TASHMATOV (*Employee*)

Academy of Federal Agency of Protection of the Russian Federation, Orel

SINGLE-PRODUCT DOUBLE-POLE STREAMING MODEL OF THE TRANSPORT NETWORK

The article deals with the mathematical aspects of assessing the capacity of groups of areas of the transport network, based on the use of single-product double-pole streaming model.

Keywords: *bandwidth; streaming model; transport network.*

BIBLIOGRAPHY (TRANSLITERATED)

1. Eremenko V.T. Osnovy' postroeniya informacionno-telekommunikacionny'x sistem: uchebnoe posobie. – Orel: OrYuI MVD Rossii, 1999. – Ch. 2. – 129 s.
2. Eremenko V.T. Teoriya informacii i informacionny'x processov: monografiya / V.T. Eremenko, I.S. Konstantinov, A.V. Kos'kin, V.A. Lobanova i dr.; pod red. d.t.n. V.T. Eremenko, d.t.n. A.P. Fisuna. – Orel: OGU, OrelGTU, 2008. – 478 s. – 500 e'kz.
3. Fillips D., Garsia-Dias A. Metody' analiza setej: per. s angl. pod red. Sushkova B.G. – M.: Mir, 1984. – 496 s.
4. Ford A., Falkerson D. Potoki v setyax: per. s angl. Vajnshtejna I.A. – M.: Mir, 1966. – 276 s.
5. Berzh K. Teoriya grafov i ee primenenie. – M.: Inostrannaya literatura: per. s angl., 1962. – 320 s.
6. Basaker R., Saati T. Konechnye grafy i seti: per. s angl. pod red. Tejmana A.I. – M.: Nauka, 1973. – 368 s.
7. Fre'nk G., Frish I. Set' svyazi i potoki: pod red. Pospelova. – M.: Svyaz', 1978. – 448 s.

MISD-МЕТОД И АППАРАТНЫЙ СЕМАФОР ГОТОВНОСТИ РЕЗУЛЬТАТА

Статья содержит описание метода и специализированной схемы обнаружения окончания вычислений. Длительность вычислений зависит от структуры данных, что позволяет в среднем уменьшить это время. Специализированная схема на основе регистра сдвига имеет однородную структуру. Однородная структура позволяет использовать эту схему для различных разрядностей данных во многих типовых операциях.

Ключевые слова: параллельные вычисления; специализированный сдвиговый регистр; стартовые точки.

АКТУАЛЬНОСТЬ РАБОТЫ

По мнению ведущих отечественных и зарубежных ученых [1], количественный рост сложности задач (объем данных, скорость обработки, иерархичность и нерегулярность информационных структур) привел к необходимости поиска структурно-схемотехнических методов и решений упреждающего обнаружения конечного результата. Данный подход при фиксации таковой частоты работы параллельных устройств позволит в среднем сократить общее время вычислений. Традиционные методы проектирования устройств вычислительной техники на схемотехническом и системном уровнях опираются на следующие принципы:

- общая синхронизация всех блоков и модулей устройств, имеющих элементы памяти для исключения гонок;
- фиксированное время выполнения операций, зависящее от организации вычислительного процесса и не зависящее от структуры обрабатываемых данных.

Вместе с тем однородные вычислительные устройства, использующие потенциальный параллелизм данных и предназначенные для обработки массивов разнородного данных (особенно переменной длины), допускают множественные условия фиксации конечного результата. При массовой обработке потоков данных переменное время завершения операций обеспечивает в среднем сокращенное время вычислений. Таким образом, актуальность исследования определяется применением интеллектуальных методов процесса поиска, извлечения и обработки дополнительной информации из структуры исходных данных для досрочного принятия решения о формировании конечного результата и разработкой нестандартных архитектурных решений субсистемного уровня организации вычислений [2].

MISD-ПОДХОД К ОРГАНИЗАЦИИ ВЫЧИСЛЕНИЙ

Предлагается подход организации вычислений, относящийся по классификации Флинна к классу MISD. Сущностью такой организации вычислений является разбиение общего вычислительного процесса по формату данных на самостоятельные локальные процессы, имеющие собственные стартовые точки, что порождает параллелизм вычислений на схемотехническом уровне проектирования. MISD-подход характеризуется рядом выигранных принципов в организации:

- количество и расположение стартовых точек локальных процессов динамически определяется структурой входных данных;
- наличие множества параллельных вычислительных процессов приводит к переменной длительности выполнения и обязательное введение схемы контроля готовности конечного результата;
- однородность состава функциональных узлов операционной части для автоматического распараллеливания вычислений.

На основе MISD- подхода определен метод проектирования устройств ВТ, состоящий в создании динамично формируемых однородных по составу вычислительных модулей, распределенных по длине обрабатываемых данных и имеющих в общем случае различную длительность процессов, тем самым обеспечивая совмещение выполняемых операций или отдельных этапов операций.

Таким образом, актуальность MISD-подхода состоит в создании параллельных архитектур однородных устройств, сочетающих лучшие технические решения вычислительных систем (параллелизм и конвейеризация) с досрочным завершением вычислительных процессов.

МЕТОД ПРОЕКТИРОВАНИЯ АППАРАТНОГО СЕМАФОРА ГОТОВНОСТИ РЕЗУЛЬТАТА

Традиционно аппаратная реализация большинства операций ориентирована на организацию в пределах одного устройства единого вычислительного процесса над входными данными. Вместе с тем известно, что для ряда операций (сортировка методом вставок, операция подстановки, операция сжатия, логические операции и т.д.) [3] существует возможность создания в пределах устройства нескольких пространственно распределенных локальных процессов с собственными стартовыми точками в зависимости от структуры входных данных. Распределение локальных процессов по длине обрабатываемых данных приводит к независимости выполнения локальных вычислительных процессов между собой. Независимость и, как следствие, потенциальный параллелизм локальных фрагментов данных достигаются за счет одновременного вычисления стартовых точек по структуре входным данным. Следствием такого метода организации вычислений является переменная длительность времени выполнения операции, определяемая длительностью наибольшего локального вычислительного процесса. Таким образом, отличительная особенность MISD-метода вычислений по создаваемым параллельным локальным процессам заключается в зависимости времени выполнения операции от структуры обрабатываемых данных [3].

ПРОЕКТИРОВАНИЕ АППАРАТНОГО СЕМАФОРА ГОТОВНОСТИ РЕЗУЛЬТАТА

В теории параллельного программирования для синхронизации параллельно выполняющихся процессов используется семафор – защищенная переменная, осуществляющая разграничение доступа к общему ресурсу со стороны нескольких выполняющихся и/или взаимодействующих процессов [4]. Для повышения эффективности и надежности параллельного программирования элементарные операции установки флага завершения и проверки состояния в рамках семафора получили аппаратную реализацию в виде специализированных схем планирования и контроля выполнения параллельных процессов, которые целесообразно применять в нестандартных архитектурах вычислительных систем и устройств.

Для контроля готовности результата от нескольких динамично создаваемых локальных вычислительных процессов в специализированное устройство, имеющее единый тактовый генератор, вводится дополнительная схема, которая отслеживает текущее состояние параллельно выполняемых локальных процессов и устанавливает флаги их готовности. Данную схему условимся называть аппаратным семафором готовности результатов. Его основное функциональное назначение – обнаруживать момент завершения вычислительного процесса и операции соответственно. Аппаратный семафор строится на основе типового сдвигового регистра.

В известных сдвиговых регистрах [4, 5] единый процесс сдвига над триггерами осуществляется по входу С синхронизации и имеет одну стартовую точку S (рис. 1).

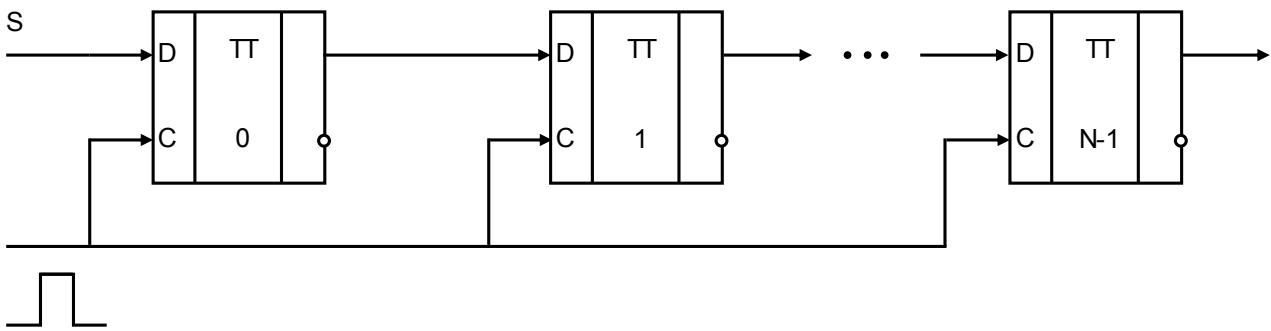


Рисунок 1 – Схема сдвигового регистра

Для организации контроля готовности локальных процессов в состав n -разрядного сдвигового регистра вводятся мультиплексыры « $2 \rightarrow 1$ ». На информационные входы триггеров подаются внешние стартовые точки $S_1 - S_{n-1}$ или предыдущее значение от соседнего триггера (рис. 2). Таким образом, в составе единого сдвигового регистра формируются несколько пространственно распределенных сдвиговых процессов.

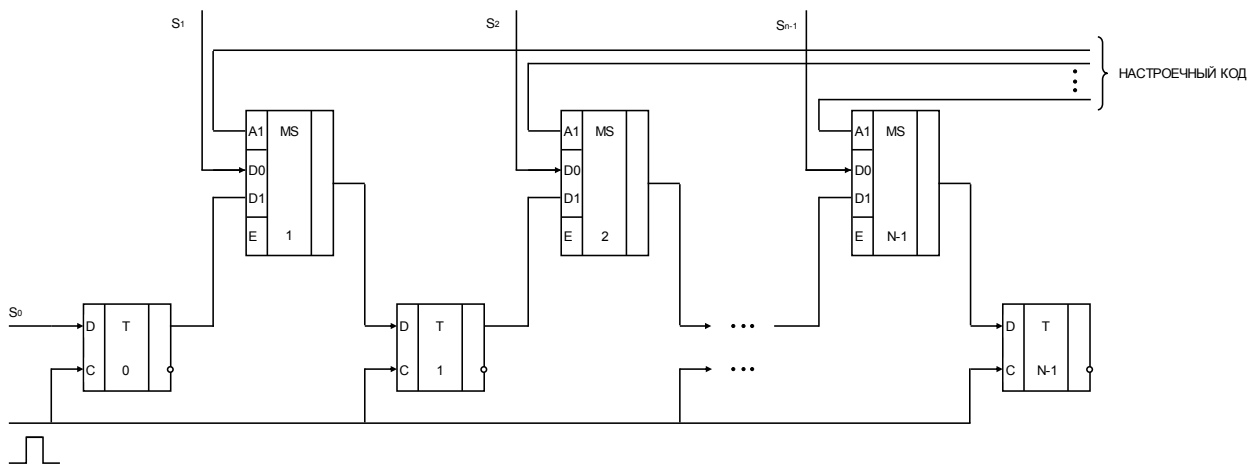


Рисунок 2 – Схема контролирующего сдвигового регистра

На адресные входы мультиплексыры « $2 \rightarrow 1$ » подается настроечный код. Настроечный код – это невесовой код-идентификатор позиции локального процесса. Если j -й бит настроечного кода равен 1, то показывает, что с j -й позиции создается новый локальный вычислительный процесс и соответствующий ему контролирующий сдвиговый процесс.

На рисунке 3 приведен пример формирования трех контролирующих процессов сдвига в аппаратном семафоре готовности результата с настроечным кодом 0101.

Настроечный код вычисляется в устройстве автоматически по обрабатываемым данным в зависимости от вида операции. Входы $S_1 \div S_4$ являются самостоятельными входными данными для сдвиговых процессов.

Созданные пространственно распределенные локальные процессы стартуют одновременно с контролирующими сдвиговыми процессами в семафоре готовности результата. Корректная работа специализированной сдвиговой схемы контроля возможна при условии, что квант обработки единицы данных в вычислительном устройстве гарантированно не меньше задержки на 1 такт специализированной сдвиговой схемы. В противном случае возникает ситуация потери обработанных единиц данных и некорректно формируется момент готовности результата. Функциональная схема ядра аппаратного семафора готовности результата представлена на рисунке 4.

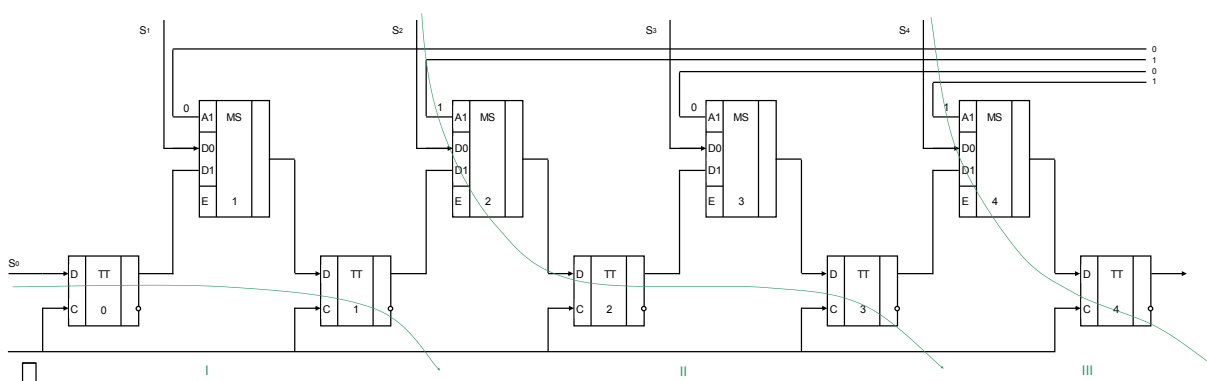


Рисунок 3 – Пример работы аппаратного семафора готовности результата

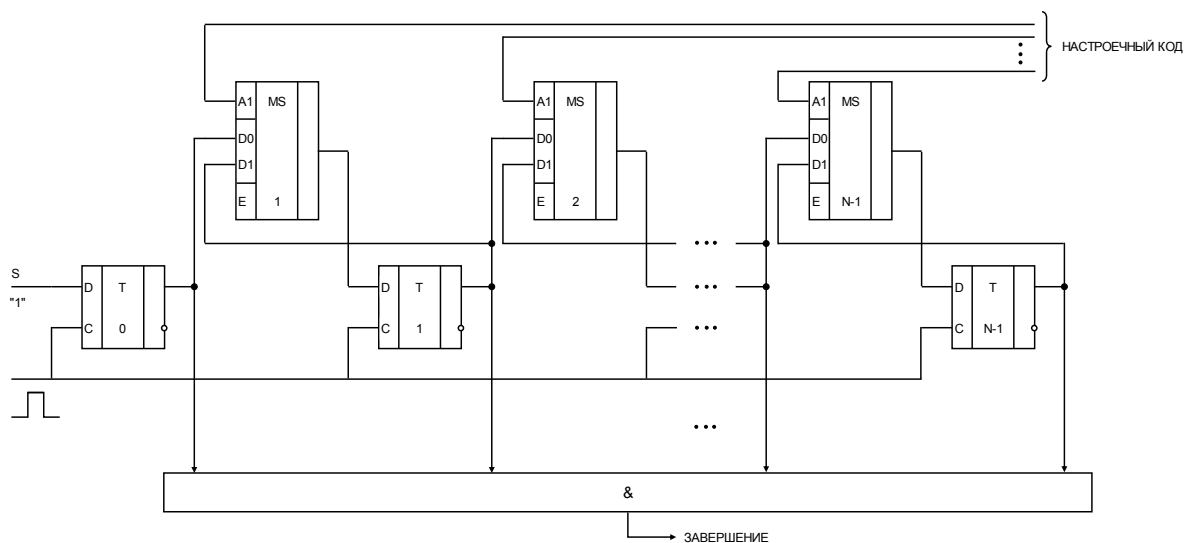


Рисунок 4 – Функциональная схема ядра аппаратного семафора готовности результата

Вычисление стартовых точек S имеет индивидуальную схему, определяемую структурой и форматами исходных операндов, а также выполняемой операцией. Вычисленный невесовой настроечный код загружается в схему контролирующего сдвигового регистра и по тактовым сигналам C выполняется распространение 1 от стартовых точек до конечной позиции сдвигового процесса.

Практическая ценность данной схемы контролирующего сдвигового регистра заключается в том, что традиционная схема сдвига (рис.1) является частным случаем при значении настроечного кода, равного $\underbrace{00\dots0}_{n-1}$. При значении настроечного кода, равного $\underbrace{11\dots1}_{n-1}$

в схеме контролирующего сдвигового регистра формируется $n-1$ контролирующий процесс.

Временной выигрыш устройств с переменной длительностью завершения работы связан со смещенным влево от максимальной длительности (n тактов) количеством тактов на самый длительный локальный процесс. В частности, для сумматоров с динамически создаваемыми локальными процессами суммирования для операндов разрядностью $n=16$ бит среднее время суммирования составило $T_{сред}=4,34$ такта, для $n=32$ среднее время суммирования составило $T_{сред}=5,33$). [3]

ВЫВОД

Разработана схема аппаратного семафора готовности результата, которая позволяет контролировать динамически создаваемые распределенные локальные вычислительные

процессы и сигнализировать о готовности результата в зависимости от структуры обрабатываемых данных. Аппаратный семафор имеет однородную структуру, что позволяет его масштабировать под изменяемые разрядности операндов. Аппаратный семафор может найти применение в архитектурах машин потоков данных [1], специализированных вычислителях, осуществляющих параллельную обработку, в том числе символьных данных [6].

СПИСОК ЛИТЕРАТУРЫ

1. Титенко Е.А. Организация реконфигурируемого мультипроцессора потока данных // Вестник компьютерных и информационных технологий, 2012. – № 2. – С. 24-29.
2. Титенко Е.А., Семенихин Е.А., Петрик Е.А. Структурно-функциональная организация арбитра параллельной обработки запросов // Информационно-измерительные и управляющие системы. – М.: Радиотехника, 2010. – № 11. – Т. 8. – С. 30-35.
3. Титенко Е.А., Скорняков К.С., Бусыгин К.Н. Методы и сумматоры с параллельными групповыми процессами // Известия Юго-Западного государственного университета. Серия «Управление, вычислительная техника, информатика». – Медицинское приборостроение, 2013. – № 2. – Ч. 1. – С. 161-165.
4. Зотов И.В. и др. Аппаратная реализация барьерной синхронизации в матричных мультикомпьютерах // Информационно-измерительные и управляющие системы, 2013. – № 8. – С. 41-45.
5. Титенко Е.А., Евсюков В.С., Семенихин Е.А. Аппаратный семафор для асинхронных устройств // Теоретические и прикладные вопросы современных информационных технологий: материалы Всероссийской научно-технической конференции. – Улан-Удэ: Изд-во ВСГТУ, 2006. – Ч. I. – С. 196-200.
6. Титенко Е.А., Емельянов С.Г., Зерин И.С. Однородные вычислительные структуры для параллельных символьных вычислений // Известия Юго-Западного государственного университета. – № 6(39). – Ч. 2. – С. 77-82.

Титенко Евгений Анатольевич

ФГБОУ ВПО «Юго-Западный государственный университет», г. Курск

Кандидат технических наук, доцент, доцент кафедры информационных систем и технологий

E-mail: johntit@mail.ru

E.A. TITENKO (*Candidate of Engineering Sciences, Associate Professor,
Associate Professor of the Department of Information Systems and Technologies*)
Southwest State University, Kursk

MISD-METHOD AND HARDWARE SEMAPHORE OF FINAL RESULT

The article is show of the high-end method and specialized scheme of the computing. The computation time depends on the data structure that allows to reduce the average time. Specialized scheme based on the shift register has a homogeneous structure. Homogeneous structure allows the use of this scheme for a different bit length in many standard operations.

Keywords: *parallel computing; specialized shift register; start points.*

BIBLIOGRAPHY (TRANSLITERATED)

1. Titenko E.A. Organizaciya rekonfiguriruemogo mul'tiprocessora potoka danny'x // Vestnik komp'yuterny'x i informacionny'x texnologij, 2012. – № 2. – S. 24-29.
2. Titenko E.A., Semenixin E.A., Petrik E.A. Strukturno-funkcional'naya organizaciya arbitra parallel'noj obrabotki zaprosov // Informacionno-izmeritel'ny'e i upravlyayushhie sistemy'. – M.: Radiotexnika, 2010. – № 11. – Т. 8. – S. 30-35.

3. Titenko E.A., Skornyakov K.S., Busy'gin K.N. Metody' i summatory' s parallel'ny'mi gruppovy'mi processami // Izvestiya Yugo-Zapadnogo gosudarstvennogo universiteta. Seriya «Upravlenie, vy'chislitel'naya texnika, informatika». – Medicinskoe priborostroenie, 2013. – № 2. – Ch. 1. – S. 161-165.
4. Zotov I.V. i dr. Apparatnaya realizaciya bar'ernoj sinxronizacii v matrichny'x mul'tikomp'yuterax // Informacionno-izmeritel'ny'e i upravlyayushhie sistemy', 2013. – № 8. – S. 41-45.
5. Titenko E.A., Evsyukov V.S., Semenixin E.A. Apparatny'j semafor dlya asinxronny'x ustrojstv // Teoreticheskie i prikladny'e voprosy' sovremenny'x informacionny'x tehnologij: materialy' Vserossijskoj nauchno-texniceskoj konferencii. – Ulan-Ude': Izd-vo VSGTU, 2006. – Ch. I. – S. 196-200.
6. Titenko E.A., Emel'yanov S.G., Zerin I.S. Odnorodny'e vy'chislitel'ny'e struktury' dlya parallel'ny'x simvol'ny'x vy'chislenij // Izvestiya Yugo-Zapadnogo gosudarstvennogo universiteta. – № 6(39). – Ch. 2. – S. 77-82.

УДК 004.931

Д.А. ТОРОПЧИН, К.В. ЛЫСЫХ, Н.И. КОРСУНОВ

**РАСПОЗНАВАНИЕ ИЗОБРАЖЕНИЙ
НА ОСНОВЕ АППРОКСИМАЦИИ ФУНКЦИЙ
И КОМБИНИРОВАННЫХ НЕЙРОННЫХ СЕТЕЙ**

В данной работе предлагается метод распознавания изображений на основе комбинированной нейронной сети.

Ключевые слова: ИНС; КНС; распознавание изображений.

Известно, что искусственные нейронные сети (ИНС) широко применяются при решении задач распознавания образов [1], аппроксимации функции и ряде других задач. При этом архитектура и обучение ИНС ориентирована на класс решаемых задач. Архитектура ИНС имеет некоторое сходство с естественными нейронными сетями. ИНС, предназначенные для решения различных задач, могут существенно различаться алгоритмами функционирования.

ИНС состоит из элементов, называемых формальными нейронами, которые сами по себе очень просты и связаны с другими нейронами. Каждый нейрон преобразует набор сигналов, поступающих к нему на вход в выходной сигнал. Именно связи между нейронами, кодируемые весами, играют ключевую роль. Одно из преимуществ ИНС (а также недостаток при реализации их на последовательной архитектуре) – это то, что все элементы могут функционировать параллельно, тем самым существенно повышая эффективность решения задачи, особенно в обработке изображений. Кроме того, что ИНС позволяют эффективно решать многие задачи, они предоставляют мощные гибкие и универсальные механизмы обучения, что является их главным преимуществом перед другими методами (вероятностные методы, линейные разделители, решающие деревья и т.п.). Обучение избавляет от необходимости выбирать ключевые признаки, их значимость и отношения между признаками. Но тем не менее выбор исходного представления входных данных (вектор в n -мерном пространстве, частотные характеристики, вейвлеты и т.п.) существенно влияет на качество решения и является отдельной темой. ИНС обладают хорошей обобщающей способностью, т.е. они могут успешно распространять опыт, полученный на конечном обучающем наборе, на все множество образов.

Решение задач аппроксимации используют многослойные ИНС прямого распространения, обучение которых ведется методом обратного распространения, а для решения задач распознавания образов более эффективно использование ИНС с обратными связями, например, сетей Гроссберга [2].

Однако любой образ представляет некоторую функцию и его распознавание тесно связано с аппроксимацией функций [3]. Представляется целесообразным использование ИНС одинаковой архитектуры как для распознавания образов, так и для аппроксимации функции.

Основанием для построения такой сети является представление функций

$f(x, y) = \cup f_i(x_i, y_j), i = 1, n, j = 1, m$, где x_i и y_j определяют границы подобласти, разбиения $n \times m$ области

(x, y) , f_i функция аппроксимирующая $f(x, y)$ в данной подобласти.

И если известно изображение $f(x,y)$ и задано разбиение (x,y) на квадраты (прямоугольники) i, j , то распознавание изображения $f(x,y)$ сводится к определению $f_i(x,y_j)$ и последовательности из этих функций.

Для решения данной задачи предлагается использовать комбинированные нейронные сети (КНС) [3].

На рисунке 1 приведена структура отдельной ИНС с пятью выходами, распознающая соответствующее число классов объектов. В модели, основанной на комбинированной структуре, для аппроксимации используются ИНС с такой же структурой, но с одним выходным нейроном.

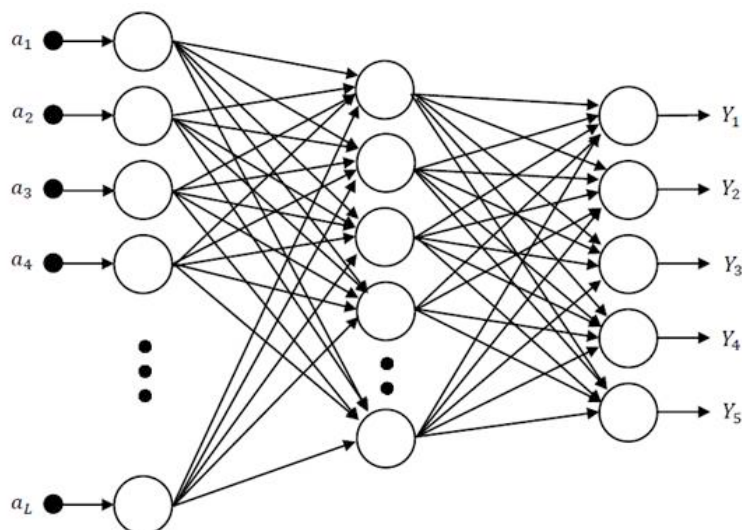


Рисунок 1 – Структура отдельной ИНС

Разработанная модель на основе комбинированной структуры для классификации образов показана на рисунке 2. На каждом отрезке используются i -е ИНС для определения аппроксимации функции на этом отрезке из заданного ограниченного множества базисных функций. По анализу последовательности функций на выходе y_i определяют класс распознаваемого образа. Класс определяется заданной последовательностью $f_1, f_2, f_3, \dots, f_n$, где f_i – одна из базисных функций.

ИНС осуществляют преобразование входных данных, представляющих образ, которое можно представить как

$$f_i(\varphi(a)) = y_i,$$

где $\varphi(a)$ – нелинейное преобразование входных данных ИНС определенной архитектуры;

i – номер ИНС в комбинированной структуре.

Множество ИНС, задействованных в комбинированной структуре, можно представить как

$$F \ni \{f_1(\varphi(a)), f_2(\varphi(a)), \dots, f_k(\varphi(a))\},$$

где k – число используемых ИНС.

В зависимости от решаемых задач используемые в комбинированной структуре ИНС могут иметь различные архитектуры. Таким образом, в общем случае комбинированная структура имеет вид:

$$F \ni \{f_i(\varphi_j(a))\}, i \in K, j \in J,$$

где K – ИНС; J – число различных архитектур.

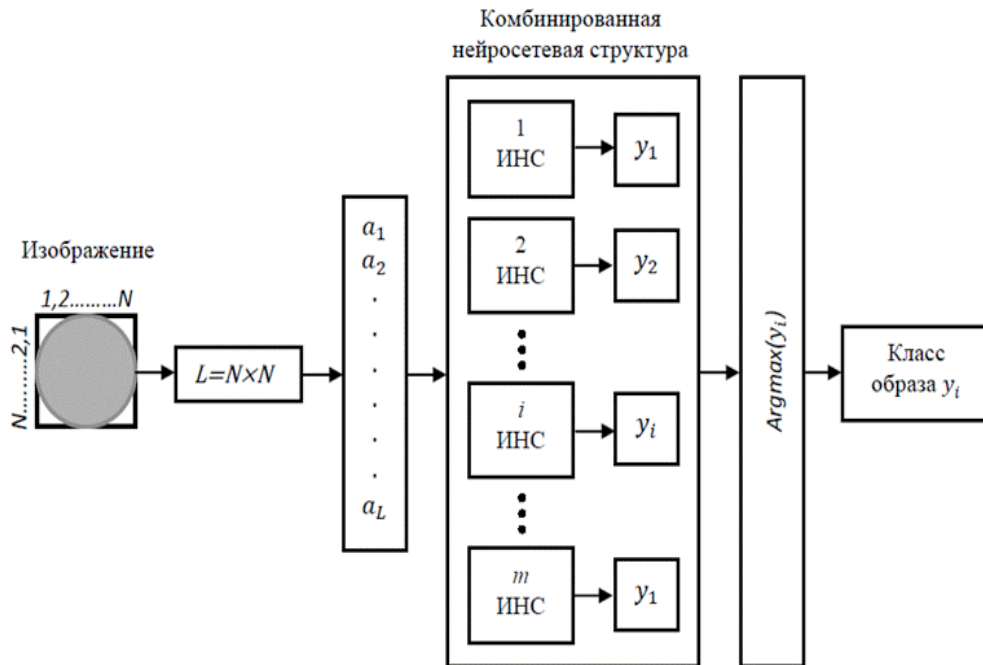


Рисунок 2 – Комбинированная нейронная сеть

Рассмотрим использование КНС для задач классификации и кластеризации. Так как предложенная структура КНС совмещает в себе свойства пластичности и стабильности, а также не требует априорного знания о фиксированном количестве необходимых классов, то предлагается использовать комбинированную модель, совмещающую в себе достоинства многослойного персептрона и модели адаптивно резонансной теории АРТ-2 (сеть адаптивно резонансной теории с дискретными функциями активации нейронов) [4].

Обучение нейронной сети классификации изображений основано на алгоритме обучения без учителя, включающего следующие шаги:

1. Формируются базисные функции.
2. Область существования изображения разбивается на прямоугольники (квадраты) подобласти по методу Кулаковского [5].
3. Архитектура i -ой сети представляется множеством сетей АРТ-2, количество этих сетей соответствует количеству прямоугольников (квадратов), разбиения области.
4. Число нейронов в каждой сети определяется количеством эталонных (базисных) функций.
5. На каждом отрезке (подобласти) определяется наиболее близкое изображение к одной из базисных функций. Определение ведется в соответствии с алгоритмом функционирования ИНС АРТ.
6. Предъявляется изображение. Обучение происходит одновременно для всех подобластей.
7. На выходе ИНС получим последовательность, которая максимально близко соответствует заранее определенному классу изображения.

Для определения базисной функции на отрезке аппроксимации (в квадрате) используется мера близости сети АРТ-2:

$$\rho = \frac{|x|}{|s|},$$

Последовательность существования базисных функций соответствует возрастающей последовательности, номеров квадратов разбиения области существования.

Введем меру близости формируемой последовательности эталонным последовательностям. Так как базисные функции задаются идентификатором в виде целого числа, критерий может быть представлен двоичным вектором. Поэтому близость двоичных векторов будем определять расстоянием Хэмминга. Тогда отнесение изображения к соответствующему классу определяется формулой:

$$d_{diff} = d_i - d_j,$$

где d_i – эталонная последовательность; d_j – полученная последовательность.

Задается мера близости t данной последовательности к i -й эталонной последовательности $i = 1, k$ и изображение относится к i -ому, если:

$$d_{diffmin} \leq t + 1. \quad (1)$$

Если классифицируемый вектор имеет d_{diff} , соответствующее (1) не для единственного i , то для отнесения его к тому или другому классу вводится дополнительное условие:

$$d_{diff} \leq 2t + 1,$$

позволяющее выделить класс по заданной подобласти существования искажений.

Таким образом, введение комбинированной нейронной сети, а также аппроксимации функций позволяет классифицировать изображения.

СПИСОК ЛИТЕРАТУРЫ

1. Васильев В.И. Проблема обучения распознаванию образов. Принципы, алгоритмы, реализация. – Киев, «Выща школа», 1989. – 64 с.
2. Сосулин Ю.Г., Фам Чунг Зунг. Нейросетевое распознавание двумерных изображений // Радиотехника и электроника, 2003. – Т. 48. – № 8. – С. 969-978.
3. Маркин М.И. Синтез нейроаппроксиматоров с неоднородной архитектурой. Материал технической информации // Программирование, 2003. – № 4. – С. 53-65.
4. Хайкин, Саймон. Нейронные сети: полный курс. – 2-е издание: пер. с англ. – М.: издательский дом «Вильямс», 2006. – 1104 с.: ил. – парал. тит. англ.
5. Кулаковский А.И. Об одном методе построения функционального преобразователя с несколькими входами. – Автоматика и телемеханика, 1966. – Выпуск 11. – С. 146-153.

Торопчин Дмитрий Анатольевич

Белгородский государственный национальный исследовательский университет, г. Белгород
Аспирант кафедры прикладной математики и информатики
E-mail: re1aps@rocketmail.com

Лысых Константин Витальевич

Белгородский государственный национальный исследовательский университет, г. Белгород
Аспирант кафедры математического и программного обеспечения информационных систем
E-mail: lysykh_k@bsu.edu.ru

Корсунов Николай Иванович

Белгородский государственный национальный исследовательский университет, г. Белгород

Доктор технических наук, профессор кафедры математического и программного обеспечения информационных систем
E-mail: korsunov@bsu.edu.ru

D.A. TOROPChIN (*Post-graduate Student of the Department of Applied Mathematics and Informatics*)

K.V. LY'SY'X (*Post-graduate Student of the Department of Mathematical and Software Information Systems*)

N.I. KORSUNOV (*Doctor of Engineering Sciences,
Professor of the Department of Mathematical and Software Information Systems
Belgorod National Research University, Belgorod*)

**IMAGE RECOGNITION BASED ON APPROXIMATION OF FUNCTIONS
AND NEURAL NETWORKS COMBINED**

In this article offered method of image recognition based on neural networks combined.

Keywords: *Artificial Neural Networks; Combined Neural Networks; image recognition.*

BIBLIOGRAPHY (TRANSLITERATED)

1. Vasil'ev V.I. Problema obucheniya raspoznavaniyu obrazov. Principy', algoritmy, realizaciya. – Kiev, «Vy'shha shkola», 1989. – 64 s.
2. Sosulin Yu.G., Fam Chung Zung. Nejrosetevoe raspoznavanie dvumerny'x izobrazhenij // Radiotekhnika i e'lektronika, 2003. – Т. 48. – № 8. – S. 969-978.
3. Markin M.I. Sintez nejroapproximirovannykh arhitekturoj. Material texnicheskoj informacii // Programmirovanie, 2003. – № 4. – S. 53-65.
4. Xajkin, Sajmon. Nejronny'e seti: polny'j kurs. – 2-e izdanie: per. s angl. – M.: izdatel'skij dom «Vil'yams», 2006. – 1104 s.: il. – paral. tit. angl.
5. Kulakovskij A.I. Ob odnom metode postroeniya funkcional'nogo preobrazovatelya s neskol'kimi vxodami. – Avtomatika i telexanika, 1966. – Vy'pusk 11. – S. 146-153.

УДК 004.045

Г.В. БАБИН, В.И. ХАЛИМОН

ПРОГРАММНЫЙ КОМПЛЕКС ДЛЯ ПОЛУЧЕНИЯ И АНАЛИЗА ДАННЫХ В ИЗМЕРИТЕЛЬНЫХ СИСТЕМАХ С ИСПОЛЬЗОВАНИЕМ ИСТОЧНИКА МЕЧЕННЫХ НЕЙТРОНОВ

Приведено описание программного комплекса, написанного на языке C++, для обработки непрерывного потока информации с различных внешних измерительных систем, которые используют в качестве канала передачи данных RS-482, RS-232, Ethernet UDP/IP. Представлена последовательность действий для инициализации устройств и их настройки, а также вывод их в рабочий режим с поддержкой активных команд. В статье описаны ключевые модули, которые служат для приема, отсылки и анализа данных. Представлены различные методы организации взаимодействия потоков для каждого модуля (общая память, Futures&Promises, STM (Software Transactional Memory), асинхронные сообщения между потоками).

Ключевые слова: аналого-цифровые преобразователи; внешние измерительные устройства; многопоточность; активные команды; спектральные данные.

ВВЕДЕНИЕ

Создание гибких программных модулей для работы с внешними неравноправными устройствами или интерфейсами зачастую является сложной задачей. В основном проблема заключается в обработке активных команд. Под активной командой будем понимать самопроизвольную отсылку посылки в порт от внешнего устройства без предварительного запроса. Необходимость использования данных команд обусловлена ситуациями, при которых обработка поступившей от внешнего устройства информации имеет наивысший приоритет, например, оповещение управляющего блока о внезапном изменении состояния устройства.

Целью данной статьи является описание реализации взаимодействия между потоками в программном комплексе Control-GTII для работы с неравноправными внешними устройствами на примере установки для досмотра УВП-1103 в ОАО НТЦ РАТЭК. В автономный измерительный комплекс входит 7 измерительных плат для сцинтилляционных детекторов и одна плата для работы с альфа-детекторами. Также в комплекс входят дополнительные внешние устройства, логика работы с которыми существенно отличается от работы с измерительными блоками детекторов. В УВП-1103 такими устройствами являются: связующий микроконтроллер, блок охлаждения, блок управления генератором. Такое сочетание нескольких различных по своей архитектуре внешних устройств приводит к серьезному усложнению программы и ее логики и в конечном итоге сводится к реализации многопоточной системы обработки информации.

ОПИСАНИЕ ПОДКЛЮЧЕНИЯ ВНЕШНИХ УСТРОЙСТВ СИСТЕМЫ ДОСМОТРА И РЕАЛИЗАЦИИ ПРОГРАММНОГО КОМПЛЕКСА

Работа с контроллерами измерительных плат детекторов происходит по UDP/IP-протоколу. Данные блоки имеют набор от 34 до 100 команд типа «запрос-ответ» и от 6 до 20 активных команд (без запроса), которые могут отправляться в любое время, указывая на внештатную ситуацию. Все остальные внешние интерфейсы используют в качестве канала передачи и приема информации последовательный порт. Принципиальная схема подключения внешних интерфейсов представлена на рисунке 1.

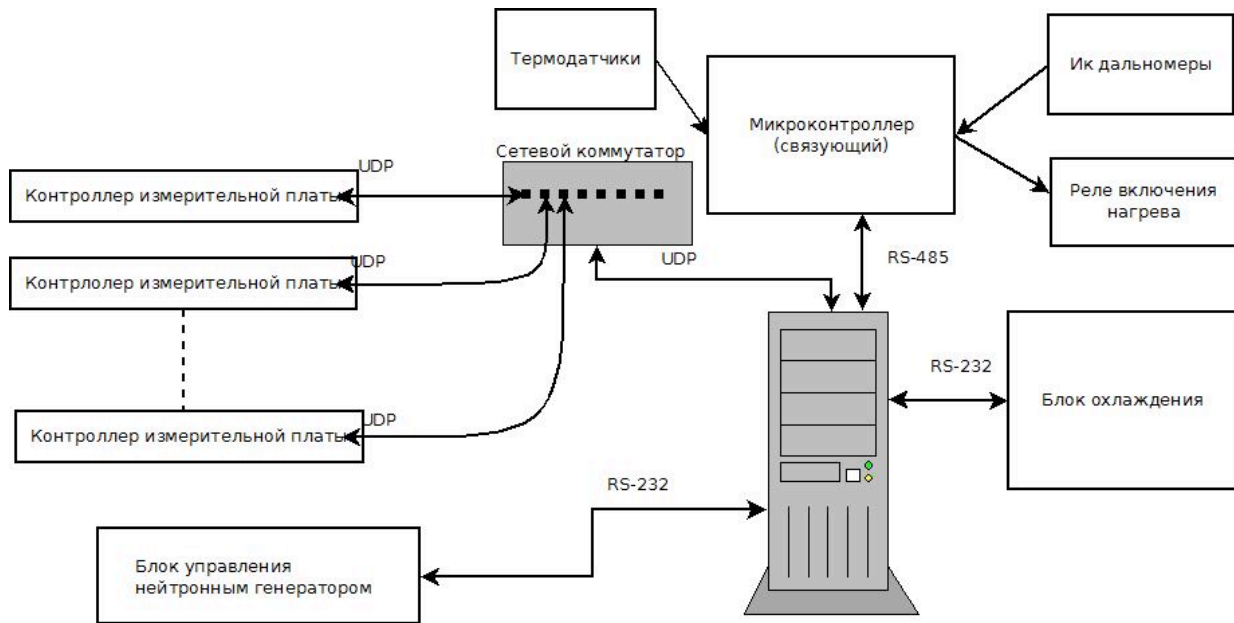


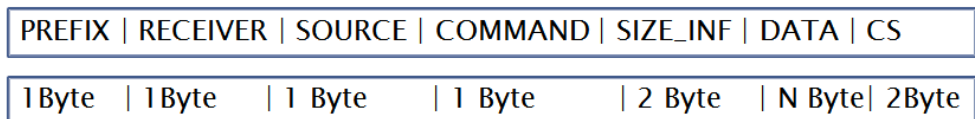
Рисунок 1 – Принципиальная схема подключения внешних интерфейсов

Работа с измерительными блоками происходит по UDP-протоколу и поэтому необходимо свести вероятность потери информации при передаче данных к минимуму. Одним из методов разрешения проблемы потери данных может служить создание отдельного потока для взаимодействия с каждым из устройств. Для этого был создан в программном комплексе отдельный класс, в котором в отдельном потоке был запущен обработчик событий поступления данных в порт. Необходимо отметить, что непосредственное считывание и запись данных в порт происходит в асинхронном режиме, так как одновременное использование устройства для двух различных функций физически невозможно, поэтому взаимодействие с основным потоком реализовано через STM. В результате одновременное обращение в порт становится запрещенной операцией. Определение активной команды во время взаимодействия происходит либо при некорректном ответе на посылку с запросом, либо при поступлении данных в порт при отсутствии запроса.

После поступления данных в порт и их анализа происходит установка сигнала потоку, отвечающего за разбор посылки. В нем происходит анализ посылки на уровне протокола, т.е. в нем реализована процедура поиска полезной команды, префикса, имени устройства и подсчет контрольной суммы. Затем информативная часть посылки копируется в промежуточный буфер для дальнейшей обработки. Пример структуры посылки представлен на рисунке 2, а схема модулей работы с внешними интерфейсами – на рисунке 3.

В зависимости от поступившей команды поток, отвечающий за обработку полезных данных, может осуществлять взаимодействие с предыдущим потоком с помощью общей памяти, либо по архитектуре Futures&Promises. Такой способ взаимодействий строится на промежуточном классе флагов и реализации условного ожидания. В классе флагов могут быть выставлены два вида событий: пассивный (управляющий поток его проверит по выполнению всех внутренних команд) и активный (поток начнет выполнение незамедлительно после окончания последней процедуры). Необходимость условно активного флага связана с возможным поступлением команды в порт, действия по которой носят относительно незамедлительный характер выполнения, например, получение активной команды.

Пример посылки входящей



Пример посылки исходящей

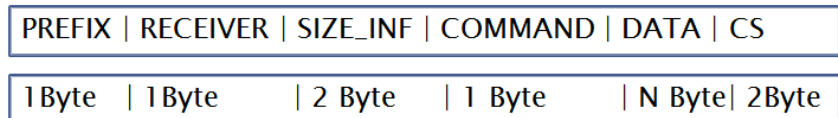


Рисунок 2 – Пример протокола посылок:

PREFIX – префикс посылки, *RECEIVER* – имя получателя, *SOURCE* – отправитель команды, *SIZE_INF* – размер посылки без префикса, *COMMAND* – команда, *CS* – контрольная сумма, *DATA* – данные

После получения команды от потока обработчика посылок данные сохраняются в буфере. В зависимости от посылки строится различная логика работы последующих потоков обработки данных – без оповещения управляющего потока, обработка по таймеру, активная обработка.

1. Режим работы без оповещения управляющего потока.

Данный режим реализован таким образом, что для определенных видов команд последовательность действий выполняется без оповещения управляющего потока. В программном комплексе обработки данных с установки типа УВП-1103, предназначенной для обнаружения опасных веществ в ОАО «НТЦ «РАТЭК», был реализован режим работы без оповещения управляющего потока. Режим предусматривает последовательный опрос состояния измерительных блоков (энергетическое смещение, коэффициенты усиления, временная задержка, флаг режектирования, высокое напряжение) и вывод нейтронного генератора в рабочий режим.

2. Обработка по таймеру.

Обработка по таймеру часто является рабочим стационарным режимом, в котором идет опрос рабочих параметров через определенный промежуток времени τ , временные интервалы устанавливаются через инициализационный файл или пользователем.

3. Активная обработка.

Активная обработка команд – это незамедлительное выполнение определенных операций при получении активной команды от внешнего устройства. Довольно часто активная обработка посылок происходит при выполнении действий, носящих опасный характер в том смысле, что данные операции могут вывести всю систему из рабочего состояния (резкий перегрев, переохлаждение, непредвиденное движение, выход из строя датчиков).

В программном комплексе Control-GТII дополнительно была реализована процедура параллельного анализа данных, в частности, для кластеризации информации для каждой пары измерительных блоков.

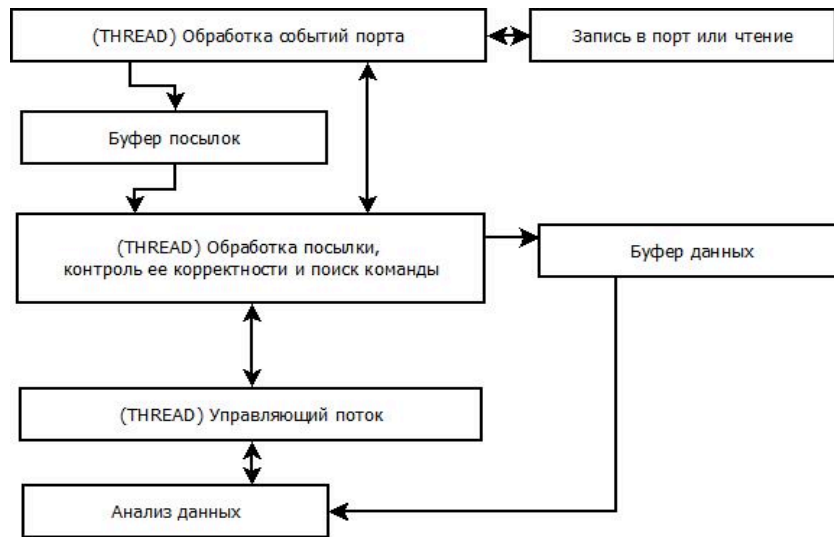


Рисунок 3 – Унифицированная схема модулей работы с внешними интерфейсами (THREAD означает поток)

ВЗАИМОДЕЙСТВИЕ УПРАВЛЯЮЩИХ ПОТОКОВ С ОСНОВНЫМ

При инициализации любого из потоков посредством флагов устанавливается внутренний приоритет на обработку событий. В реализованном программном обеспечении для УВП-1103 приоритетность сигналов установлена в следующем порядке:

1. Обработка данных и управление с контроллера движения и ИК-датчиками. Обработка данных и управление с контроллера для поддержки рабочей температуры.
2. Обработка данных и управление с блока управления генератора.
3. Обработка данных и управление с измерительных блоков гамма и альфа детекторов.

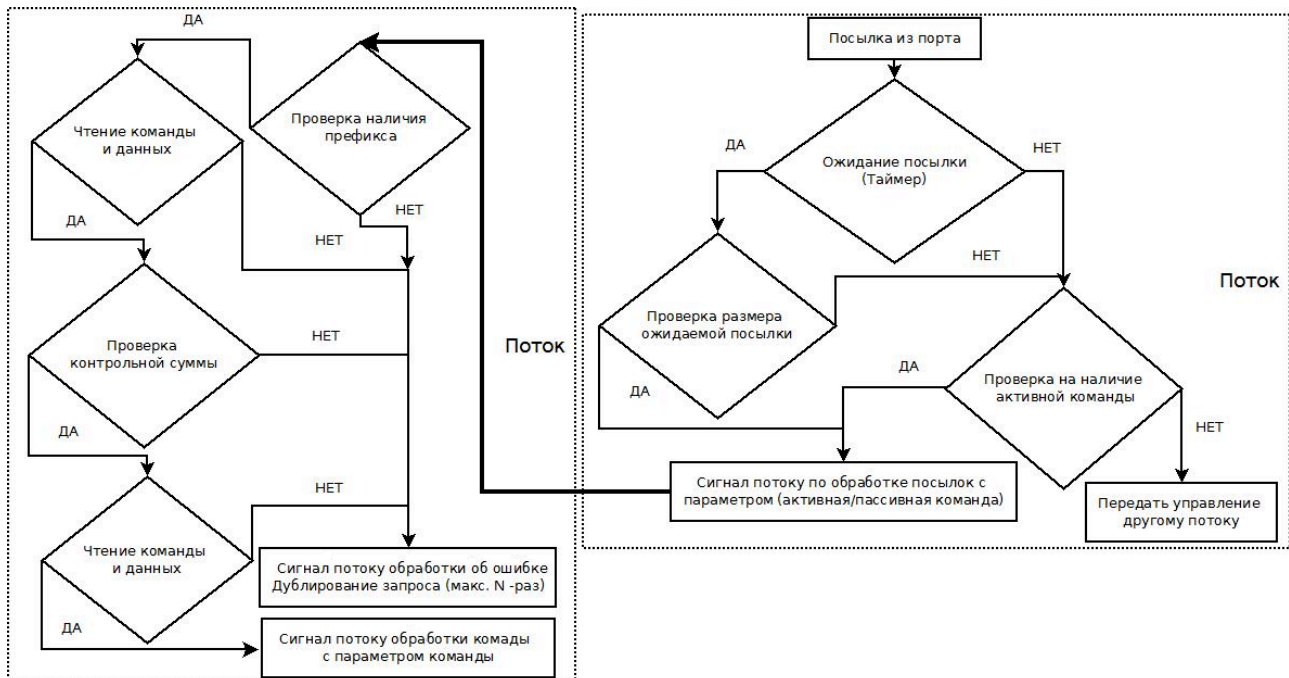


Рисунок 4 – Схема модуля по обработке команд

Если в порт поступила посылка, содержащая активную команду, то действия, связанные с ней, имеют наивысший приоритет. Структурная схема взаимодействия модулей

с основным потоком показана на рисунке 5. Как видно из данной схемы, некоторые потоки, такие, как поток управления блоком генератора, платой с контроллером ИК-датчиков и транспортера и системы охлаждения, могут взаимодействовать в обход основного потока. Причина такой реализации связана с необходимостью выполнения определенных действий без промедления, другими словами, без обращения к основному потоку. В частности, при выходе из строя блока управления генератором необходимо остановить движение установки и выключить измерение раньше, чем оповестить оператора через интерфейс. Пунктирными линиями на рисунке 5 показано, что каждый управляющий модуль может функционировать автономно без основного потока. Это позволяет повысить выживаемость системы в случае выхода из строя одного из устройств. В качестве примера приведем логику работы с измерительными платами.

Первоначально строится список условных состояний устройства в управляющем модуле. В случае измерительной платы детекторов таких состояний три:

- инициализация;
- установка/запрос параметров;
- рабочий режим.

Для каждого состояния строится структура, характеризующаяся следующими параметрами:

- глобальный счетчик ошибок;
- максимальное количество глобальных ошибок;
- структуры команд.

Каждой команде присваивается уникальный номер, номер следующей команды, номер предыдущей команды, счетчик ошибок и максимальное их количество. Логика работы построена так, что, после удачного выполнения команды происходит переход к следующей команде, со сбросом счетчика ошибок данной команды для данного состояния. В случае превышения допустимого количества ошибок выполняется переход к предыдущей команде. Любая ошибка увеличивает счетчик глобальных ошибок. Первая команда не имеет номера предыдущей команды, а последняя – следующей, так как при выполнении (или ошибке соответствующей операции) производится переход в другое условное состояние устройства. Структура для активных команд имеет упрощенный вид, есть только уникальный номер состояния, в который необходимо перевести систему, и номер команды, с которой необходимо начать следующие действия. Условно логика работы представлена на рисунке 6.

При таком построении логики работы достигается автономность управляющего модуля или, другими словами, взаимодействие и обработка данных для одного устройства становится независимыми от других.

ЗАКЛЮЧЕНИЕ

Представленная на примере программного комплекса для установки по обнаружению опасных веществ организация многопоточности является гибкой для модификации, так как все потоки разделены на определенные блоки и каждый из них выполняет свою задачу, что, в свою очередь, позволяет достаточно легко модифицировать программный код, используя модульный принцип программирования.

Предложенный подход по распределению приоритетов модулей по анализу данных для различных устройств позволяет достоверно определить последовательность работ в нештатной ситуации, что зачастую является довольно проблематичной задачей при создании и использовании многопоточного приложения.

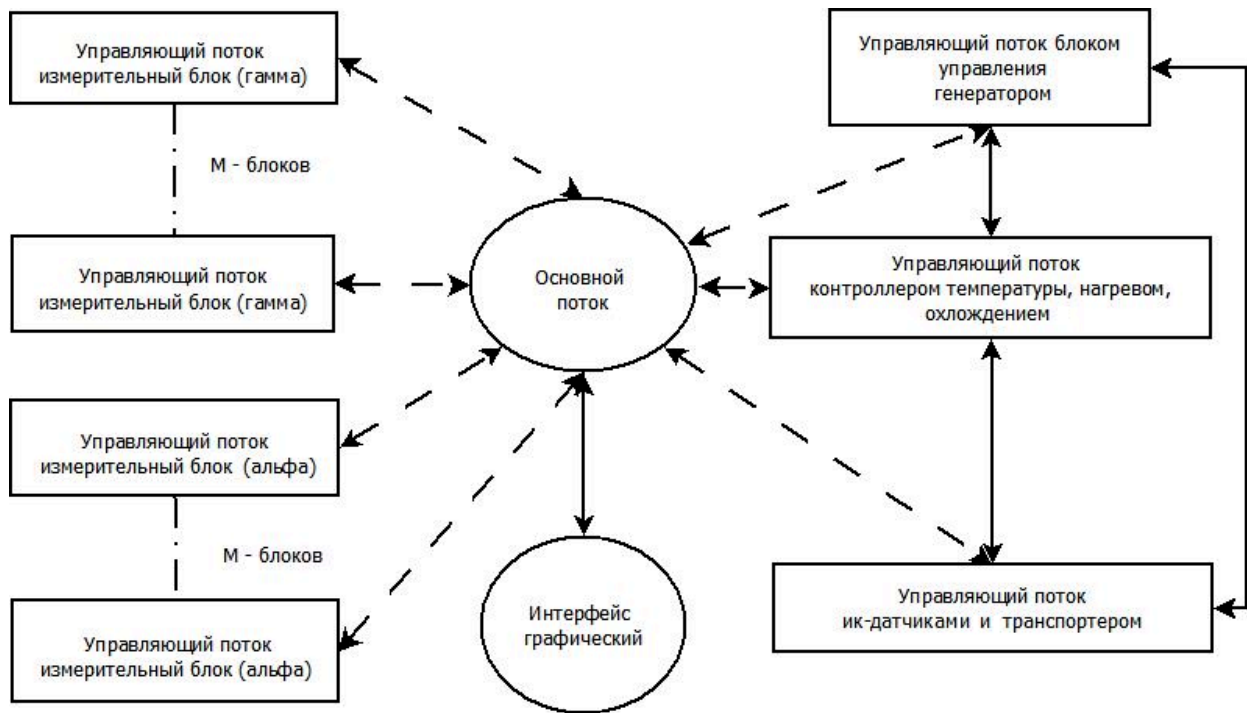


Рисунок 5 – Схема организации взаимодействия между потоками

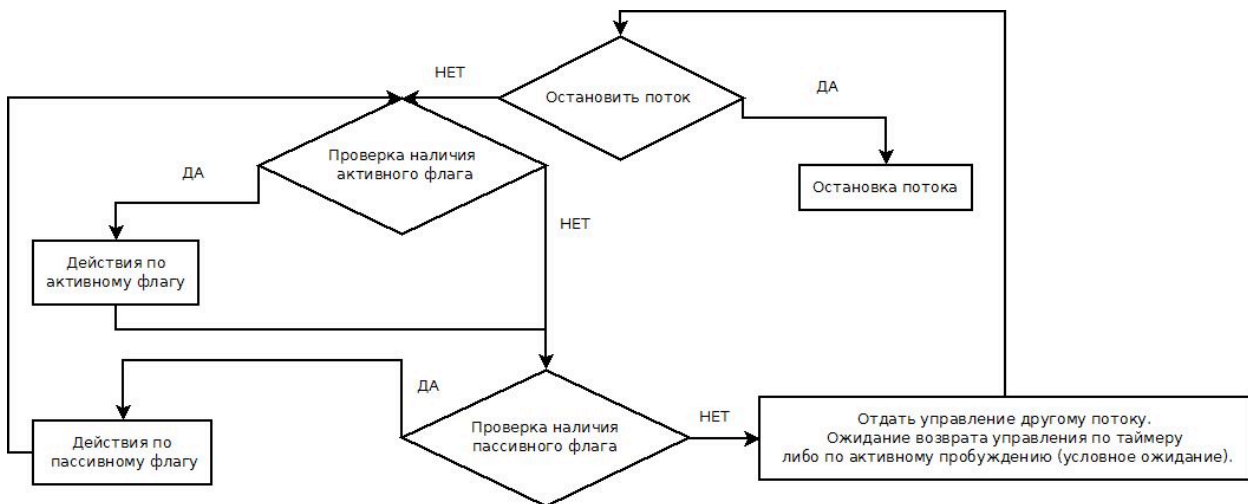


Рисунок 6 – Условная схема работы управляющего потока

СПИСОК ЛИТЕРАТУРЫ

1. Bisseling H. Parallel Scientific Computation a Structured Approach Using BSP and MPI. Oxford University Press, 2004.
2. Воеводин Вл.В., Воеводин В.В. Параллельные вычисления. – СПб: БХВ-Петербург, 2002. – 608 с.
3. Воеводин Вл.В. Теория и практика исследования параллелизма последовательных программ // Программирование, 1992. – № 3. – С. 38-53.
4. Жуматий С.А., Савольев С.И. Оценка загруженности компьютера в различных UNIX-системах // Численные методы, параллельные вычисления и информационные технологии // Издательство Московского университета, 2008. – 143 с.
5. Johnsson S.L. Communication Efficient Basic Linear Algebra Computations on Hypercube Architectures // Journal of Parallel and Distributed Computing, 1987. – № 4. – 133 p.

- Hill M.D., Hower D., Moore K.E. A Case for Deconstructing Hardware Transactional Memory Systems // Univ. of Wisconsin Computer Sciences Technical Report CS-TR-2007-1594, June 2007.

Бабин Герман Владимирович

Санкт-Петербургский государственный технологический институт (технический университет), г. Санкт-Петербург

Аспирант кафедры системного анализа

Тел.: 8 911 750 38 66

E-mail: germanbabin@gmail.com

Халимон Виктория Ивановна

Санкт-Петербургский государственный технологический институт (технический университет), г. Санкт-Петербург

Доктор технических наук, профессор

G.V. BABIN (*Post-graduate Student of the Department of System Analysis*)

V.I. XALIMON (*Doctor of Engineering Sciences, Professor*)

St. Petersburg State Technological Institute (Technical University), Saint Petersburg

**SOFTWARE FOR OBTAINING AND ANALYSING OF DATA
IN MEASURING SYSTEMS USING TAGGED NEUTRON SOURCE**

This article is described the architecture of the software package which are written in C++ for processing a continuous stream of information from a variety of external measuring systems. This system was used data link RS-482, RS-232, Ethernet UDP/IP. In this article was shown the sequence of actions for initialization and configuration of device, and also perpetrating to the working mode with active commands. The key modules are described to receive, send and data analysis. Different methods to organize communication streams for various operations (Shared Memory, Futures&Promises, STM (Software Transactional Memory), asynchronous messages between threads) are presented.

Keywords: *analog-to-digital converters; external measuring devices; multi-threading, active command; spectral data.*

BIBLIOGRAPHY (TRANSLITERATED)

- Bisseling H. Parallel Scientific Computation a Structured Approach Using BSP and MPI. Oxford University Press, 2004.
- Voevodin V.I., Voevodin V.V. Parallel'ny'e vy'chisleniya. – SPb: BXV-Peterburg, 2002. – 608 s.
- Voevodin V.I. Teoriya i praktika issledovaniya parallelizma posledovatel'ny'x programm // Programirovanie, 1992. – № 3. – S. 38-53.
- Zhumatij S.A., Sovol'ev S.I. Ocenka zagruzhennosti komp'yutera v razlichny'x UNIX-sistemax // Chislenny'e metody', parallel'ny'e vy'chisleniya i informacionny'e tehnologii // Izdatel'stvo Moskovskogo universiteta, 2008. – 143 s.
- Johnsson S.L. Communication Efficient Basic Linear Algebra Computations on Hypercube Architectures // Journal of Parallel and Distributed Computing, 1987. – № 4. – 133 p.
- Hill M.D., Hower D., Moore K.E. A Case for Deconstructing Hardware Transactional Memory Systems // Univ. of Wisconsin Computer Sciences Technical Report CS-TR-2007-1594, June 2007.

РЕАЛИЗАЦИЯ МУЛЬТИПРОТОКОЛЬНОЙ КОММУТАЦИИ ПО МЕТКАМ НА БАЗЕ ОПЕРАЦИОННЫХ СИСТЕМ С ОТКРЫТЫМ ИСХОДНЫМ КОДОМ

В статье предложена методика построения программного маршрутизатора с поддержкой технологии MPLS на базе операционной системы с открытым исходным кодом. Рассматривается функциональная структура программного маршрутизатора, а также предлагаются программные модули, дополняющие функциональные возможности операционной системы для использования в качестве коммуникационного оборудования.

Ключевые слова: мультипротокольная коммутация; MPLS; маршрутизация; свободное программное обеспечение; открытый код.

Применение операционных систем с открытым исходным кодом в качестве платформы для построения коммуникационного оборудования широко известно и получило повсеместное распространение в IP-сетях [1]. Причиной такого успеха можно считать высокопроизводительную и надежную сетевую подсистему, организующую на уровне ядра ОС разбор, обработку и пересылку пакетов [2]. С точки зрения архитектуры операционной системы это позволяет проектировать высокоуровневые сетевые приложения, упрощая при этом ядро до минимального набора системных вызовов, необходимых для функционирования в качестве сетевого устройства. В частности, такой подход может быть использован для построения роутера с открытым исходным кодом [2], правила пересылки пакетов в котором определяются на уровне приложения. Производительность подобных решений является достаточной для практического использования и сопоставимой с производительностью аппаратных решений от мировых лидеров в данной отрасли [3].

Дальнейшим развитием исследований является расширение функциональных возможностей роутера с открытым исходным кодом до возможностей аппаратных аналогов. Современные вычислительные сети имеют тенденцию к расширению спектра предоставляемых услуг [4]. В связи с этим на коммуникационном оборудовании реализуется поддержка множества протоколов и технологий, реализующих высокоуровневые сервисы конечных пользователей. С другой стороны, поддержка данных функций в распределенной сети ведет к отказу от традиционных способов пересылки пакетов и развитию методов коммутации и маршрутизации, позволяющих наиболее эффективным образом использовать имеющиеся вычислительные ресурсы и пропускную способность. Данное обстоятельство существенно усложняет конфигурирование сетевых устройств, а, следовательно, и внедрение новых услуг поверх существующих физических сетей. На сегодняшний день единой методики решения обозначенных проблем не существует, однако активно развивается концепция программно-коммутируемых сетей (SDN) и технология мультипротокольной коммутации по меткам (MPLS), которые направлены на более эффективное управление информационными потоками. Поскольку обе технологии изначально были заявлены как свободные с точки зрения лицензионных ограничений, то целесообразным является их реализация в рамках маршрутизатора на базе операционной системы с открытым исходным кодом. Целью данной работы является предложение методики по созданию подобного устройства, ориентированного на применение в действующих сетях связи и обладающего широкими возможностями по переносимости и виртуализации.

Технология MPLS предполагает отказ от стандартной методики маршрутизации пакетов. Вместо нее применяется маршрутизация потоков, выделяемых из общего трафика

согласно некоторым критериям [5]. Данный подход позволяет существенно снизить вычислительные затраты на обработку одного пакета данных, поскольку, во-первых, решение о перенаправлении выполняется на канальном, а не на сетевом уровне модели OSI, во-вторых, решение о перенаправлении пакета известно заранее. Еще одним преимуществом является возможность гибкого управления потоками данных, которая достигается за счет перераспределения меток на промежуточных устройствах в зависимости от меняющихся требований к качеству обслуживания, надежности, пропускной способности и т.п. Реализация же технологии MPLS в рамках открытой ОС позволяет внедрить высокоуровневую систему управления, распределяющую метки согласно централизованно разработанным правилам. При этом в отличие от концепции SDN маршрутизатор остается полностью функциональным и может работать автономно при потере связи с центральным устройством управления.

Применение технологии MPLS вносит изменения в структуру передаваемых сегментов данных. Помимо адресов канального и сетевого уровня кадр дополняется специальной меткой, которая определяет маршрут следования инкапсулированных данных [6]. Следовательно, сетевая подсистема должна корректно обрабатывать такой заголовок, присваивать следующую метку или же передавать пакет без меток на граничном устройстве MPLS-домена. Назначение меток соответствующим информационным потокам осуществляется на основе имеющейся информации о структуре сети. В наиболее общем случае эта информация может быть получена из таблиц маршрутизации.

Таким образом, в роутере на базе операционной системы с открытым исходным кодом должны быть реализованы:

- протоколы динамической маршрутизации;
- механизмы распространения меток согласно имеющейся маршрутной информации;
- средства распознавания кадров канального уровня, содержащих метки;
- механизмы продвижения пакетов на основе таблицы меток.

С учетом архитектуры современных операционных систем с монолитным ядром реализация взаимодействия вышеописанных функциональных блоков возможна только при поддержке технологии MPLS на уровне ядра. На сегодняшний день поддержка функций MPLS штатно заложена в ядре ОС OpenBSD, а также в специальном проекте для ОС Linux [7]. Поскольку обе системы являются совместимыми на уровне исходных кодов, то в дальнейшем именно они будут рассматриваться в качестве базы для решения поставленной задачи.

Проанализируем возможности реализации обозначенных функциональных блоков. Поддержка статической маршрутизации является неотъемлемой частью операционной системы и реализуется стандартными библиотечными модулями. Эти же модули отвечают непосредственно за процесс пересылки пакетов на основе адресов сетевого уровня. Таким образом, для поддержки протоколов динамической маршрутизации необходимо на основе входящего трафика заполнить таблицу маршрутизации. Данная функция реализуется с использованием специализированного программного обеспечения, в качестве которого предлагается использовать проект Quagga. Основными критериями выбора данного проекта являются поддержка распространенных протоколов в соответствии с RFC (OSPF, IS-IS, BGP), высокая стабильность и производительность [3].

Механизм распространения меток является связующим звеном при переходе от маршрутизации на третьем уровне модели OSI к коммутации по меткам. Существующие подходы представлены статическим назначением меток, а также стандартными протоколами LDP и RSVP-TE [8]. Поскольку различия в производительности между двумя протоколами минимальны и ими можно пренебречь, то основополагающим критерием выбора является поддержка со стороны выбранных операционных систем. Наибольшей степенью интеграции

с выбранным программным обеспечением обладает протокол LDP, реализуемый как дополнение к проекту Quagga. Менеджер Zebra в указанном проекте фактически играет роль связующего звена между протоколами маршрутизации, протоколом распространения меток и сетевой подсистемой ядра ОС. Такая архитектура позволяет организовать единый программный интерфейс между всеми компонентами MPLS-модуля.

Наиболее значимой является реализация механизмов распознавания и продвижения MPLS-пакетов. Каждая MPLS-метка добавляется между заголовками канального и сетевого уровней, следовательно, и функции ее обработки распределяются между ними. Первичный разбор кадра осуществляется на аппаратном уровне с последующей передачей инкапсулированных данных Ethernet-драйверу. Драйвер, в свою очередь, принимает решение о передаче данных либо IP-модулю, либо MPLS-модулю в зависимости от содержимого кадра. Последним принимается решение о соответствующей обработке метки: замене, добавлении, снятии. Основанием для принятия такого решения является обозначенная ранее таблица меток.

Процесс продвижения пакетов также вносит некоторые изменения в процесс взаимодействия модулей сетевой подсистемы. Вычисление аппаратного адреса получателя должно производиться не по его IP-адресу, а по добавленной метке. Таким образом, запрос от Ethernet-драйвера должен обрабатываться не стандартными модулями стека TCP/IP, а MPLS-модулем ядра. Таким образом, на пограничных устройствах будут задействованы обе подсистемы, а для промежуточных устройств MPLS-домена продвижение пакетов будет выполняться исключительно средствами разрабатываемого модуля. Можно говорить о возможности выделения MPLS-модуля и выносе его на соответствующую аппаратную платформу, однако данная возможность выходит за рамки рассматриваемой тематики и является темой для дальнейших исследований.

Стоит отметить, что предложенная методика обладает высокой степенью переносимости и масштабируемости. В частности, маршрутизатор на базе операционной системы с открытым исходным кодом может быть как на базе реального оборудования, так и на базе различных сред виртуализации. Как показано в работе [9], данные подходы отличаются производительностью, однако их применение может быть обусловлено местом развертывания MPLS-домена, а также спецификой решаемых задач передачи данных.

Таким образом, можно представить следующую функциональную структуру роутера на базе ОС Linux или OpenBSD, реализующую мультипротокольную коммутацию по меткам (рис. 1). С учетом архитектуры ядра выбранных операционных систем функции по обработке MPLS-пакетов распределяются между сетевым адаптером, драйвером, программным обеспечением в пользовательской области памяти и ядром. Такая реализация полностью соответствует концепции MPLS, так как отделяет уровень управления передачей данных от непосредственно передачи данных. А реализация уровня управления в пользовательской области позволяет разрабатывать высокоуровневые приложения по управлению всей сетью, а не только отдельными устройствами.

Как видно из приведенных данных, эффективность использования имеющейся пропускной способности находится в пределах 66-70%. Данный факт объясняется обработкой MPLS-пакетов на уровне ядра операционной системы, а не на уровне аппаратуры. При этом загрузка центрального процессора не выходила за пределы 25%. Величина джиттера во всех экспериментах говорит о возможности использовать предложенный маршрутизатор с открытым исходным кодом для передачи данных, чувствительных к временным задержкам.

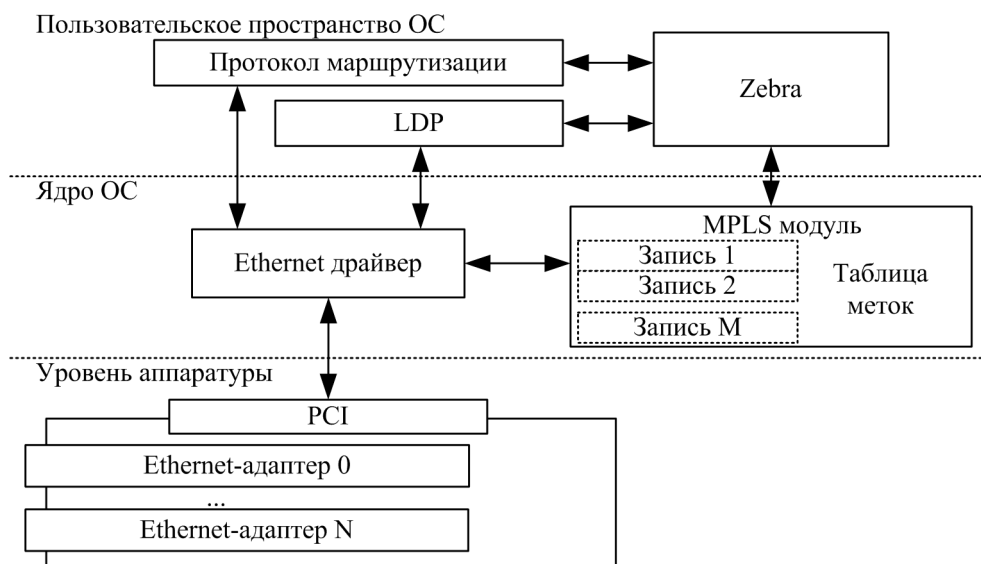


Рисунок 1 – Схема подсистемы поддержки MPLS на базе операционной системы

Таблица 1 – Результаты серии экспериментов

Полоса пропускания, Mbits/sec	Джиттер, ms
666	0.031
686	0.028
697	0.023
689	0.015
689	0.024
687	0.022
696	0.017
688	0.018

Основываясь на полученных данных, можно сделать следующие выводы. Существующими открытыми программными средствами возможно реализовать маршрутизатор с поддержкой мультипротокольной коммутации по меткам на базе операционной системы с открытым исходным кодом, причем эффективность такого устройства будет сопоставима с аппаратными решениями. В то же время неоспоримым преимуществом такого программного маршрутизатора является разделение уровней управления и обработки данных, что практически недостижимо для существующих аппаратных устройств. Немаловажным оказывается также и снижение стоимости на один-два порядка по сравнению с существующими аппаратными аналогами.

Дальнейшее развитие предложенной методики заключается в увеличении производительности разработанной подсистемы, что достигается выносом всех описанных функций на аппаратный уровень. Также важными аспектами функционирования, не затронутыми в данной работе, являются безопасность и надежность. Поскольку на практике возможно сочетание различных типов оборудования, поддерживающих различные алгоритмы распространения меток и резервирования ресурсов, то их включение в соответствующие модули операционной системы является первоочередной задачей для внедрения предложенной технологии на практике.

СПИСОК ЛИТЕРАТУРЫ

1. Bianco A. and other. Open-source PC-Based Software Routers: A Viable Approach to High-Performance Packet Switching / A. Bianco, J.M. Finochietto, G. Galante, M. Mellia, F. Neri // In Proc. Of the 3rd International Workshop on QoS in Multiservice IP Networks, QoSIP 2005, Catania, Italy, Feb. 2-4, 2005. – P. 353-366.
2. Rami Rosen Linux Kernel Networking: Implementation and Theory, Apress, 2013, ISBN 9781430261964. – P. 648
3. Ruambo F., Gebre E. Performance Analysis of a Bifrost Open Source Router. Master's Thesis in Computer Network Engineering School of Information Science, Computer and Electrical Engineering Halmstad University, 2012.
4. Олифер В.Г., Олифер Н.А. Компьютерные сети: принципы, технологии, протоколы; 4-е издание. – СПб.: Питер, 2010. – 886 с.: ил.
5. Олифер В.Г., Олифер Н.А. Виртуальные частные сети на основе MPLS // Журнал сетевых решений LAN, 2002. – № 3. – С. 54-58.
6. Rosen E., Viswanathan A., Callon R. Multiprotocol Label Switching Architecture, RFC 3031, Internet Engineering Task Force, January, 2001.
7. Maravic I., Smiljanic A. MPLS Implementation for the Linux Kernel // High Performance Switching and Routing (HPSR), 2012. IEEE 13th International Conference, Belgrade. – P. 23-28.
8. Adami, D. Signalling Protocols in DiffServ-aware MPLS Networks: Design and Implementation RSVP-TE Network Simulator / D. Adamiand, C. Callegari, S. Giordano, F. Mustacchio, M. Pagano, F. Vitucci // IEEE Global Telecommunications Conference, 5.
9. Bredel M., Bozakov Z. Yuming Jiang Analyzing Router Performance Using Network Calculus with External Measurements // Quality of Service (IWQoS), 2010. 18th International Workshop, Beijing. – P. 1-9.

Долгов Евгений Петрович
 Академия ФСО России, г. Орел
 Кандидат технических наук
 E-mail: dolevp@list.ru

E.P. DOLGOV (*Candidate of Engineering Sciences*)
Academy of Federal Agency of Protection of the Russian Federation, Orel

IMPLEMENTATION OF MULTIPROTOCOL LABEL SWITCHING BASED ON OPEN SOURCE OPERATING SYSTEMS

The paper proposes a method of constructing a software router with support for MPLS-based operating system with open source. The functional structure of the routing software as well as software modules are available to complement the functionality of the operating system to be used as communication equipment.

Keywords: *multiprotocol switching; MPLS; routing; free software; open source.*

BIBLIOGRAPHY (TRANSLITERATED)

1. Bianco A. and other. Open-source PC-Based Software Routers: A Viable Approach to High-Performance Packet Switching / A. Bianco, J.M. Finochietto, G. Galante, M. Mellia, F. Neri // In Proc. Of the 3rd International Workshop on QoS in Multiservice IP Networks, QoSIP 2005, Catania, Italy, Feb. 2-4, 2005. – P. 353-366.
2. Rami Rosen Linux Kernel Networking: Implementation and Theory, Apress, 2013, ISBN 9781430261964. – P. 648
3. Ruambo F., Gebre E. Performance Analysis of a Bifrost Open Source Router. Master's Thesis in Computer Network Engineering School of Information Science, Computer and Electrical Engineering Halmstad University, 2012.

4. Olifer V.G., Olifer N.A. Komp'yuterny'e seti: principy', texnologii, protokoly'; 4-e izdanie. – SPb.: Piter, 2010. – 886 s.: il.
5. Olifer V.G., Olifer N.A. Virtual'ny'e chastny'e seti na osnove MPLS // Zhurnal setevy'x reshenij LAN, 2002. – № 3. – S. 54-58.
6. Rosen E., Viswanathan A., Callon R. Multiprotocol Label Switching Architecture, RFC 3031, Internet Engineering Task Force, January, 2001.
7. Maravic I., Smiljanic A. MPLS Implementation for the Linux Kernel // High Performance Switching and Routing (HPSR), 2012. IEEE 13th International Conference, Belgrade. – P. 23-28.
8. Adami, D. Signalling Protocols in DiffServ-aware MPLS Networks: Design and Implementation RSVP-TE Network Simulator / D. Adamiand, C. Callegari, S. Giordano, F. Mustacchio, M. Pagano, F. Vitucci // IEEE Global Telecommunications Conference, 5.
9. Bredel M., Bozakov Z. Yuming Jiang Analyzing Router Performance Using Network Calculus with External Measurements // Quality of Service (IWQoS), 2010. 18th International Workshop, Beijing. – P. 1-9.

УДК 621.396

А.Г. ДУБРОВИН, А.И. ОФИЦЕРОВ,
Н.Н. ЕРОФЕЕВ, И.С. ОВЕЧКИН

СПОСОБ ПОВЫШЕНИЯ ЭФФЕКТИВНОСТИ ИСПОЛЬЗОВАНИЯ ВЫДЕЛЕННОГО РЕСУРСА В МУЛЬТИСЕРВИСНЫХ СЕТЯХ СПУТНИКОВОЙ СВЯЗИ

Разработан способ определения длительности суперфрейма обратного канала спутниковой сети стандарта DVB-RCS, обеспечивающий эффективное использование выделенной сети частотной полосы. Основное отличие разработанного способа от известных ранее заключается в определении длительности суперфрейма, позволяющее сократить время на процедуру формирования частотно-временного плана и нахождения пакетных данных в буфере спутникового терминала.

Ключевые слова: интерактивная сеть спутниковой связи; обратный канал; длительность суперфрейма; частотно-временной ресурс.

ВВЕДЕНИЕ

Проведенный анализ направлений развития спутниковых систем связи в нашей стране и за рубежом позволяет утверждать, что на современном этапе производители оборудования концентрируют свои усилия на расширении количества видов услуг и повышении их качества (*QoS*). Происходит активное совершенствование технических характеристик и массогабаритных показателей спутниковых терминалов (СТ), технологическую основу которых составляют земные станции с малой апертурой антенных систем – *VSAT*-терминалы. Переход на пакетные методы передачи информации обуславливает совершенствование технологий передачи по единому каналу связи различных видов услуг: цифровая речь, видеотрафик, передача данных, доступ в сеть Интернет [4].

Вместе с тем высокая стоимость аренды частотно-энергетического ресурса (ЧЭР) транспондеров спутников связи, увеличение объемов передаваемой информации приводят к необходимости интеграции технических решений по передаче мультимедийного трафика с мобильных узлов и стационарных объектов связи на основе методов пакетной коммутации.

Предложенный в стандарте [1] метод многочастотного доступа с временным разделением *MF-TDMA* (*Multi Frequency – Time Division Multiple Access*) в обратных спутниковых каналах позволяет динамически изменять величину занимаемой спутниковой сетью полосы частот в зависимости от объема передаваемого активными спутниковыми терминалами трафика данных. Однако ориентация технологии *MF-TDMA* на коммерческое применение вызывает сложности при необходимости гарантированного обеспечения требований к качеству обслуживания категоризированных абонентов. Кроме того, отсутствие в стандарте математически обоснованных решений по определению необходимой для работы сети доли частотно-временного ресурса транспондера спутника связи привело к использованию производителями оборудования эвристических способов распределения ресурса. Данные особенности применения коммерческих технологий в силовых структурах и ведомствах обуславливают необходимость постановки задачи по разработке способа повышения эффективности использования ресурса, выделенного сети спутниковой связи.

ПОСТАНОВКА ЗАДАЧИ ИССЛЕДОВАНИЯ

В сети спутниковой связи с технологией *MF-TDMA* доступный ресурс $(0, t) \times (f_{sf})$ (мкс×МГц) представляется набором несущих частот (частотной полосой), на которых реализован многостанционный доступ с временным разделением. Во временной

области единицей измерения является суперфрейм, который разделен на фреймы, каждый из которых, в свою очередь, разделен на таймслоты (рис. 1).

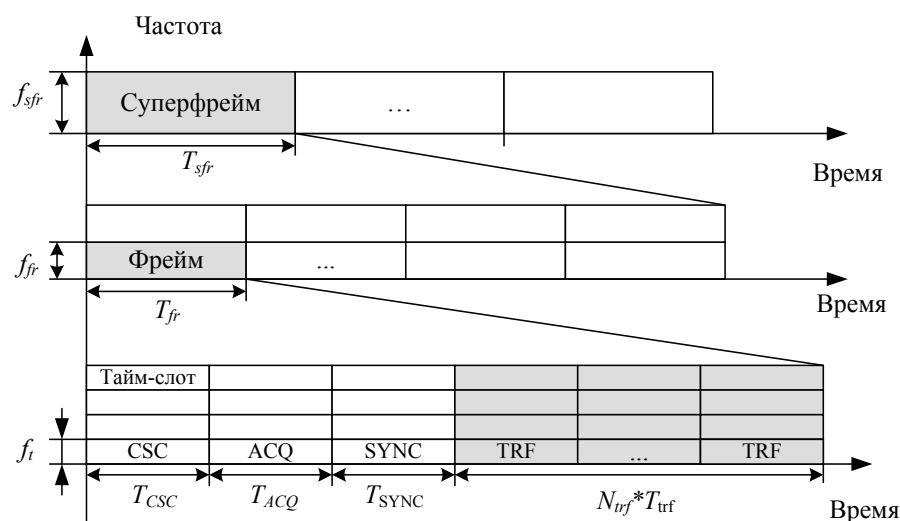


Рисунок 1 – Частотно-временная структура суперфрейма

Передача пакетного трафика спутниковыми терминалами осуществляется по одному или последовательно по нескольким частотным каналам на временных позициях трафиковых таймслотов. Последовательность таймслотов в пределах фрейма имеет следующий порядок: таймслот *CSC* общеканальной сигнализации, таймслот *ACQ* частотной коррекции, таймслоты *SYNC* синхронизации и трафиковые таймслоты.

В соответствии со стандартом [1] скорость передачи информации на несущей зависит от условий распространения радиосигнала, определяемых через отношение сигнал/шум в прямом или обратном канале, и изменяется от наименьшей (для худших условий распространения) до максимальной символьных скоростей для наилучших условий функционирования.

В свою очередь, информационная скорость несущей также может варьироваться путем изменения позиционности модуляции и скорости кодирования и для одной и той же символьной скорости R_s может различаться в несколько раз. Например, для четырех символьных скоростей спутникового модема, изменяющихся от 256 до 2048 ксимб/с, информационная скорость одного временного слота будет изменяться более чем в 50 раз (табл. 1).

Таблица 1 – Количество информационных бит в слоте в зависимости от конфигурации спутникового модема

	$R_{s1}=256$ ксимб/с	$R_{s2}=512$ ксимб/с	$R_{s3}=1024$ ксимб/с	$R_{s4}=2048$ ксимб/с
<i>BPSK</i> , $R=1/2$	512	1024	2048	4096
<i>QPSK</i> , $R=1/2$	1024	2048	4096	8192
<i>QPSK</i> , $R=3/4$	1536	3072	6144	12288
<i>8PSK</i> , $R=2/3$	2048	4096	8192	16384
<i>16-QAM</i> , $R=3/4$	3072	6144	12288	24576
<i>16-QAM</i> , $R=7/8$	3584	7168	14336	28672

Анализ ряда научных работ показал, что повышение эффективности использования частотно-временного ресурса достигается либо за счет сокращения избыточности в

формируемом СТ запросе динамического резервирования пропускной способности [2, 4], либо за счет формирования центральной станцией (ЦС) сети структуры суперфрейма, обеспечивающего сокращение частотной полосы, необходимой для передачи трафика [7, 8] и, как следствие, сокращения средств на аренду ресурса.

Однако в условиях, когда частотная полоса ретранслятора связи выделена спутниковой сети на условиях долгосрочной аренды, варьируемым параметром для представленной на рисунке 1 структуры суперфрейма является его длительность T_{sfr} .

Длительность суперфрейма должна изменяться достаточно гибко, чтобы эффективно реагировать на запросы различных категорий. Стандартом [1] определено четыре категории запросов на выделение ресурса – *CRA*, *RBDC*, *VBDC* и *FCA*, – различающиеся между собой требованиями по временной задержке доставки пакетов, джиттеру задержки и вероятности потери данных.

Слишком большая длительность СФ приводит к существенному увеличению среднего времени ожидания процедуры формирования ЦС частотно-временного плана передачи и, как следствие, выделения СТ ресурса пропускной способности, что может привести к отказу в обслуживании некоторых видов сервиса, поэтому СФ с небольшой длительностью будет более адаптивным к реагированию на скорость формирования СФ, но при этом в нем возрастает доля служебных кадров, что также приводит к снижению общей пропускной способности спутникового канала и при той же самой входной нагрузке может привести к наступлению режима перегрузки [5, 6].

Таким образом, решением сформулированной задачи является разработка правила определения оптимальной длительности суперфрейма, обеспечивающего сокращение времени нахождения пакетных данных в буфере СТ, а также снижения вероятности наступления перегрузки, возникающей при отсутствии возможности передачи трафика.

СПОСОБ РЕШЕНИЯ И РЕЗУЛЬТАТЫ МОДЕЛИРОВАНИЯ

Пусть D (ксимв/с) – суммарная пропускная способность сети, B_{Σ} (ксимв/с) – сумма скоростей запросов на выделение пропускной способности. Пропускная способность сети D зависит от конфигурации суперфрейма – количества несущих частот в СФ и количества поднесущих во фрейме.

Пусть N_s (ксимв) – количество служебных таймслотов, необходимых для передачи сигнальных таймслотов в суперфрейме, которое не зависит от длительности СФ и определяется только количеством активных СТ.

На первоначальном этапе длительность СФ T_{sfr} определяется произвольно, исходя из накопленных во время предыдущих сеансов статистических данных. В этом случае средняя задержка в формировании следующего ЧВП составит $T_{sfr}/2$. Введем переменную, $Y(T_{sfr})$, которая будет характеризовать степень изменения средней задержки от длительности суперфрейма, тогда

$$Y(T_{sfr}) = \begin{cases} e^{\gamma \left(\frac{T_{sfr}}{2} \right)}, & \text{если } T_{sfr} < T_{sfr \max} \\ \infty, & \text{если } T_{sfr} > T_{sfr \max} \end{cases}, \quad (1)$$

где $\gamma (> 0)$ – нормировочный коэффициент для средней задержки, характеризующий вид трафика. Значение $Y(T_{sfr})$ будет возрастать при увеличении T_{sfr} до значения, не

превышающего $T_{sfr\ max}$, где $T_{sfr\ max}$ – инженерный параметр, соответствующий максимально допустимой задержки в формировании ЧВП. Выражение (1) ограничивает максимальное значение средней задержки значением $T_{sfr\ max}/2$.

Очевидно, что чем меньше T_{sfr} , тем меньше будет значение $Y(T_{sfr\ max})$, однако слишком малая величина длительности суперфрейма может привести к наступлению режима перегрузки, возникающего при условии

$$N_s + B_{\Sigma}T_{sfr} > DT_{sfr}, \quad (2)$$

что соответствует превышению количества поступивших запросов над пропускной способностью сети $B_{\Sigma} \geq D$. В этом случае сеть не может адаптироваться ко всем поступившим запросам вне зависимости от длительности суперфрейма T_{sfr} . С другой стороны, перегрузку можно избежать, устанавливая T_{sfr} таким, что

$$T_{sfr} \geq \frac{N_s}{D - B_{\Sigma}} \quad (3)$$

Предполагая, что потери в использовании ресурса при перегрузке пропорциональны степени перегрузки, равной отношению $(N_s + B_{\Sigma}T_{sfr})/DT_{sfr}$, то очевидно, что потери при перегрузке могут быть описаны выражением:

$$U(T_{sfr}) = \begin{cases} e^s \left(\frac{N_s + B_{\Sigma}T_{sfr}}{DT_{sfr}} \right), & \text{если } 0 < T_{sfr} < \frac{N_s}{D - B_{\Sigma}}, \\ 0, & \text{если } T_{sfr} \geq \frac{N_s}{D - B_{\Sigma}} \end{cases}, \quad (4)$$

где $s (>0)$ – коэффициент, характеризующий степень перегрузки для различных видов трафика.

Окончательно потери в использовании ресурса будут определяться суммой потерь от увеличения средней задержки и от перегрузки (выражения 1 и 4):

$$P(T_{sfr}) = Y(T_{sfr}) + U(T_{sfr}). \quad (5)$$

Подставляя в данное выражение соответствующие значения для функции потерь при задержке и перегрузке, получим:

$$P(T_{sfr}) = \begin{cases} e^{\gamma} \left(\frac{T_{sfr}}{2} \right) + e^s \left(\frac{N_s + B_{\Sigma}T_{sfr}}{DT_{sfr}} \right), & \text{если } T_{sfr} < \min \left\{ T_{sfr\ max}, \left(\frac{N_s}{D - B_{\Sigma}} \right) \right\} \\ e^{\gamma} \left(\frac{T_{sfr}}{2} \right), & \text{если } \frac{N_s}{D - B_{\Sigma}} \leq T_{sfr} \leq T_{sfr\ max} \\ \infty, & \text{во всех остальных случаях} \end{cases}. \quad (6)$$

Оптимальное значение длительности суперфрейма T_{sfr} , при котором будут обеспечены требования по средней задержке и отсутствие перегрузки, может быть определено как:

$$T_{sfr}^* = \begin{cases} \frac{N_s}{D - B_{\Sigma}}, & \text{если } \frac{N_s}{D - B_{\Sigma}} \leq T_{sfr \max} \\ \min \left\{ \sqrt{\frac{e^S}{e^{\gamma}} \frac{2N_s}{D}}, T_{sfr \max} \right\}, & \text{если } \frac{N_s}{D - B_{\Sigma}} > T_{sfr \max} \end{cases} \quad (7)$$

Если в выражении (7) e^S будет существенно больше e^{γ} , то данное выражение можно преобразовать к виду:

$$T_{sfr}^* = \min \left\{ \frac{N_s}{D - B_{\Sigma}}, T_{sfr \max} \right\}. \quad (8)$$

В соответствии с разработанным методом значение $T_{sfr \min}$ в выражении (8) является оптимальной длительностью суперфрейма, для которого выражение $\frac{N_s}{D - B_{\Sigma}}$

определяет минимальную длительность, позволяющую избежать перегрузки (выражение 3).

Кроме того, для предотвращения перегрузки значение e^S в соответствии с выражением (6) должно быть как можно больше, что позволяет установить не только малую длительность СФ, но и повысить гибкость системы при распределении ресурса в обратных каналах.

ЗАКЛЮЧЕНИЕ

В научной статье проведен анализ основных направлений развития мультисервисных сетей спутниковой связи, обосновано направление совершенствования способов распределения частотно-временного ресурса ретрансляторов связи. Сформулирована и решена научная задача по определению длительности суперфрейма, обеспечивающая сокращение задержки передачи данных в спутниковом канале и снижение вероятности наступления перегрузки в сети. Применение обоснованных математических подходов к процедуре формирования структуры суперфрейма позволяет повысить эффективность использования частотно-временного ресурса сети при обеспечении требований к качеству сервиса различных классов пакетных данных.

Проверка эффективности разработанного метода осуществлялась методом имитационного моделирования для следующих исходных данных: количество активных СТ равно 20, длительность между двумя независимыми запросами пропускной способности имеет геометрическое распределение и составляет 10 суперфреймов для *CRA* трафика и 5 суперфреймов для *RBDC* трафика, соответственно. Выигрыш в коэффициенте использования пропускной способности составил от 5 до 18% по сравнению с суперфреймом фиксированной длительности [3].

СПИСОК ЛИТЕРАТУРЫ

1. ETSI EN 301 790 v.1.3.1 (03.2003) European Standard Digital Video Broadcasting (DVB);

- Interaction Channel for Satellite Distribution Systems.
2. Ki-Dong Lee, Yong Hoon Cho, Ho Jin Lee. Apparatus and Method for Dynamic Resource Allocation in Interactive Satellite Multimedia System. Патент на изобретение US 7346069 B2, H04L 12/28, от 18.03.2008.
 3. Дубровин А.Г. Способ динамического резервирования пропускной способности обратных каналов в сети спутниковой связи интерактивного доступа / А.Г. Дубровин, А.А. Илюхин, В.В. Попов, Б.Г. Катыгин. Патент на изобретение RU 2012126525 от 25.07.13.
 4. Дубровин А.Г., Илюхин А.А. Способ формирования частотно-временного плана передачи интерактивного трафика в мультисервисных спутниковых сетях // Телекоммуникации. – М.: «Наука и технологии», 2012. – № 11. – С. 12-19.
 5. Еременко В.Т. Анализ моделей управления трафиком в сетях АСУП на основе технологии MPLS / В.Т. Еременко, С.В. Еременко, Д.В. Анисимов, С.А. Черепков, А.А. Лякишев // Информационные системы и технологии, 2013. – № 1. – С. 106-112.
 6. Еременко В.Т. Синтез локально-оптимальной структуры классификатора информационных ресурсов по критерию минимума средней длины процедуры поиска / В.Т. Еременко, А.А. Батенков, И.С. Полянский, К.А. Батенков, М.А. Сазонов // Вестник компьютерных и информационных технологий, 2013. – № 7. – С. 3-8.
 7. Офицеров А.И., Еременко В.Т. Методы управления информационными потоками в сетях передачи данных на основе резервирования ресурсов // Методы и устройства передачи и обработки информации. Межвузовский сборник научных трудов. Выпуск 11. – М.: «Радиотехника», 2009. – С. 340-346.
 8. Офицеров А.И., Еременко В.Т., Черепков С.А. Метод проектирования сетей передачи данных, совместимых с неблокируемой маршрутизацией // Вестник компьютерных и информационных технологий, 2012. – № 4. – С. 38-46.

Дубровин Александр Георгиевич

Академия ФСО России, г. Орел
Кандидат технических наук, сотрудник
Тел.: 8 (4862) 54-94-59
E-mail: dubrovin1973@mail.ru

Офицеров Александр Иванович

Академия ФСО России, г. Орел
Кандидат технических наук, сотрудник
Тел.: 8 (4862) 54-99-13
E-mail: oficerow@mail.ru

Ерофеев Николай Николаевич

Академия ФСО России, г. Орел
Сотрудник
Тел.: 8 (4862) 54-94-59
E-mail: erofeev.96@mail.ru

Овечкин Иван Сергеевич

Академия ФСО России, г. Орел
Сотрудник
Тел.: 8 (4862) 54-99-13
E-mail: ivanovechkin@mail.ru

A.G. DUBROVIN (*Candidate of Engineering Sciences, Employee*)

A.I. OFICEROV (*Candidate of Engineering Sciences, Employee*)

N.N. EROFEEV (*Employee*)

I.S. OVEChKIN (*Employee*)

Academy of Federal Agency of Protection of the Russian Federation, Orel

METHOD OF INCREASE OF EFFICIENCY OF USAGE OF THE SELECTED RESOURCE ON MULTISERVICE COMMUNICATION SATELLITE NETWORKS

In this article developed a method for determining superframe length of a return channel of a satellite communication network of DVB-RCS standard. The method provides high resource utilization of network bandwidth. A main difference between this method and are known earlier is determining superframe length to getting reduce a burst time plan allocation and waiting for traffic data standing in a satellite terminal buffer.

Keywords: *an interactive satellite communication network; a return channel; superframe length; time and bandwidth resource.*

BIBLIOGRAPHY (TRANSLITERATED)

1. ETSI EN 301 790 v.1.3.1 (03.2003) European Standard Digital Video Broadcasting (DVB); Interaction Channel for Satellite Distribution Systems.
2. Ki-Dong Lee, Yong Hoon Cho, Ho Jin Lee. Apparatus and Method for Dynamic Resource Allocation in Interactive Satellite Multimedia System. Patent na izobrenenie US 7346069 B2, H04L 12/28, ot 18.03.2008.
3. Dubrovin A.G. Sposob dinamicheskogo rezervirovaniya propusknoj sposobnosti obratny'x kanalov v seti sputnikovoj svyazi interaktivnogo dostupa / A.G. Dubrovin, A.A. Ilyuxin, V.V. Popov, B.G. Katy'gin. Patent na izobrenenie RU 2012126525 ot 25.07.13.
4. Dubrovin A.G., Ilyuxin A.A. Sposob formirovaniya chastotno-vremennogo plana peredachi interaktivnogo trafika v mul'tiservisny'x sputnikovy'x setyax // Telekommunikacii. – M.: «Nauka i texnologii», 2012. – № 11. – S. 12-19.
5. Eremenko V.T. Analiz modelej upravleniya trafikom v setyax ASUP na osnove texnologii MPLS / V.T. Eremenko, S.V. Eremenko, D.V. Anisimov, S.A. Cherepkov, A.A. Lyakishev // Informacionny'e sistemy' i texnologii, 2013. – № 1. – S. 106-112.
6. Eremenko V.T. Sintez lokal'no-optimal'noj struktury' klassifikatora informacionny'x resursov po kriteriyu minimuma srednej dliny' procedury' poiska / V.T. Eremenko, A.A. Batenkov, I.S. Polyanskij, K.A. Batenkov, M.A. Sazonov // Vestnik komp'yuterny'x i informacionny'x texnologij, 2013. – № 7. – S. 3-8.
7. Oficerov A.I., Eremenko V.T. Metody' upravleniya informacionny'mi potokami v setyax peredachi danny'x na osnove rezervirovaniya resursov // Metody' i ustrojstva peredachi i obrabotki informacii. Mezhvuzovskij sbornik nauchny'x trudov. Vy'pusk 11. – M.: «Radiotexnika», 2009. – S. 340-346.
8. Oficerov A.I., Eremenko V.T., Cherepkov S.A. Metod proektirovaniya setej peredachi danny'x, sovместimy'x s neblokiruemoj marshrutizaciej // Vestnik komp'yuterny'x i informacionny'x texnologij, 2012. – № 4. – S. 38-46.

ПРОГРАММНЫЙ КОМПЛЕКС ДЛЯ ИССЛЕДОВАНИЯ КАЧЕСТВА АЛГОРИТМОВ ОБНАРУЖЕНИЯ СЕТЕВЫХ АНОМАЛИЙ

В статье предложен подход к оценке качества алгоритмов обнаружения аномалий, возникающих в информационных системах и технологических сетях организаций. Выделены причины возникновения сетевых аномалий и алгоритмы, позволяющие их выявлять. Разработана архитектура программного комплекса и описаны алгоритмы работы клиентской и серверной частей. Исследованы алгоритмы на основе дискретного вейвлет-преобразования с применением статистических критериев и обнаружения аномалий Бродского-Дарховского. В результате проведения экспериментов и оценки качества был найден алгоритм, имеющий наилучшее значение F-меры и чувствительности при классификации аномалий в сетевом трафике.

Ключевые слова: атака; информационная система; злоумышленник; ошибки первого рода; ошибки второго рода; размер окна; точность.

ВВЕДЕНИЕ

В настоящее время актуальна проблема обеспечения безопасности информационных и технологических систем предприятий. Поскольку сетевые атаки могут привести к утечке, искажению или недоступности данных, администраторам необходимо следить за поведением компьютерных сетей и подключенных к ним информационных ресурсов. По данным отчета аналитического центра компании InfoWatch по сравнению с 2014 годом число утечек информации в мире выросло на 22%, в России – на 73%. Поэтому для предотвращения утечки информации по сети или минимизации ее последствий важно обнаружить сетевую атаку на ранней стадии ее воздействия. Любое отступление или уклонение от правильного поведения системы может указывать на возможную атаку или вторжение. Таким образом, от качества алгоритма обнаружения сетевых аномалий может зависеть, насколько точно программное средство, реализующее данный алгоритм, способно выявить аномалию или возможную атаку в сети предприятия.

ЗАДАЧИ ПРОГРАММНОГО КОМПЛЕКСА ДЛЯ ИССЛЕДОВАНИЯ КАЧЕСТВА АЛГОРИТМОВ ОБНАРУЖЕНИЯ СЕТЕВЫХ АНОМАЛИЙ

Существует множество алгоритмов для обнаружения аномалий в сети. В данной работе авторами для исследования качества и точности были выбраны два алгоритма:

- алгоритм на основе дискретного вейвлет-преобразования с применением статистических критериев (ДВП);
- алгоритм обнаружения аномалий Бродского-Дарховского (БД).

Выбранные алгоритмы в соответствии с [1, 2] являются одними из наиболее распространенных и простых в реализации. Для каждого алгоритма характерно свое количество ошибок первого (FN) и второго рода (FP) при обнаружении аномалий в сетевом трафике. Чем меньше число данных ошибок, тем более точным будет алгоритм (табл. 1).

Таблица 1 – Возможные результаты классификации аномалий

Класс A=anomaly		Принадлежность к классу	
		ДА	НЕТ
Результат	Правильный	True positive (TP)	False positive (FP)

классификации		$anomaly \rightarrow anomaly$	$normal \rightarrow anomaly$ (Ошибка второго рода)
	Неправильный	False negative (FN) $anomaly \rightarrow normal$ (Ошибка первого рода)	True negative (TN) $normal \rightarrow normal$

Для оценки качества классификации чаще всего используются следующие метрики – чувствительность и специфичность.

Чувствительность (Se) – доля истинно положительных случаев при классификации аномалий алгоритмом:

$$Se = \frac{TP}{TP + FN} 100\% . \quad (1)$$

Специфичность (Sp) – доля истинно отрицательных случаев, которые были правильно идентифицированы алгоритмом:

$$Sp = \frac{TN}{TN + FP} 100\% . \quad (2)$$

Алгоритм с высокой чувствительностью при классификации сетевого трафика часто дает истинный результат при наличии положительного исхода (обнаруживает положительные случаи). А алгоритм с высокой специфичностью чаще дает истинный результат при наличии отрицательного исхода (обнаруживает отрицательные случаи). «Идеальный» алгоритм обладает 100% чувствительностью и специфичностью. Однако при решении задач классификации аномалий и оценки качества алгоритма классификации необходимо учитывать, что максимальная специфичность и чувствительность на практике одновременно недостижимы, следовательно, необходимо искать наиболее рациональное соотношение между ними. Поэтому в качестве метрики оценки качества алгоритмов обнаружения сетевых аномалий в данной работе авторами предлагается использовать F-меру с установкой различного веса для специфичности и чувствительности алгоритма. Это связано с тем, что по данным [3] при классификации аномалий и атак наибольшее значение имеет именно минимизация ошибок первого рода, которые связаны с пропуском алгоритмом аномалии или атаки при ее наличии (чувствительность алгоритма).

$$F = (\beta^2 + 1) \frac{SpSe}{\beta^2 Sp + Se} , \quad (3)$$

где β – вес критерия, $\beta > 1$, поскольку в данном случае при оценке качества алгоритмов обнаружения сетевых аномалий приоритет отдается чувствительности алгоритма.

Для проведения исследований качества алгоритмов был разработан программный комплекс, выполняющий задачи, отраженные в таблице 2.

Таблица 2 – Задачи программного комплекса

Функции	Задачи
Выбор алгоритма	1. Обнаружение сетевых аномалий с помощью алгоритма Бродского-Дарховского. 2. Обнаружение сетевых аномалий с помощью алгоритма дискретного вейвлет-преобразования с применением статистических критериев.
Анализ аномалий	Проведение анализа интервалов на наличие аномальных пакетов, применяя выбранный алгоритм.
Оценка результатов по критериям	Оценка качества работы алгоритмов по следующим критериям: – наличие ошибки первого рода; – наличие ошибок второго рода;

	<ul style="list-style-type: none"> - чувствительность алгоритма; - специфичность алгоритма; - F-мера качества алгоритма.
--	---

АРХИТЕКТУРА ПРОГРАММНОГО КОМПЛЕКСА

Для решения поставленных перед программным комплексом задач и реализации его функциональных возможностей была разработана клиент-серверная архитектура, представленная на рисунке 1.

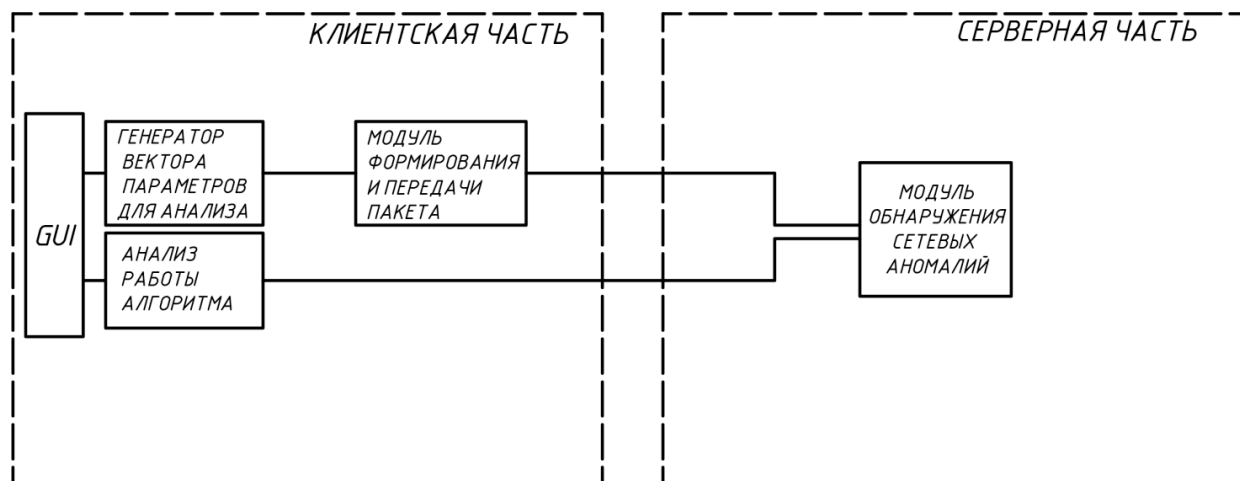


Рисунок 1 – Клиент-серверная архитектура программного комплекса

Клиентская часть состоит из графического пользовательского интерфейса (GUI), «Генератора вектора параметров для анализа», модуля «Анализ работы алгоритма» и модуля «Формирование и передача пакетов».

Графический интерфейс пользователя предназначен для обеспечения удобства взаимодействия пользователя с программным комплексом. Состоит из вкладки «Анализ» и вкладки «Параметры».

Вкладка «Анализ» (рис. 2) реализует возможность установки пользователем параметров модели сетевого трафика (число интервалов, число аномалий – зона 3) и параметров алгоритма поиска аномалий (тип алгоритма – зона 2, размер окон – зона 4). Позволяет выводить графическое представление сгенерированной модели сетевого трафика (зона 5), а также данные о результатах работы алгоритмов и параметрах, характеризующих его качество, представленных в виде таблицы (зона 7). Таблица содержит колонки: дата, название алгоритма, размер окна W1, размер окна W2, количество правильно обнаруженных аномалий, ошибки первого рода, ошибки второго рода.

Вкладка «Параметры» (рис. 3) осуществляет управление настройками подключения к серверу (хост и порт серверной части – зона 8).

«Генератор вектора параметров для анализа» – модуль, реализующий следующие операции:

- создание модели по заданным пользователем параметрам для дальнейшего преобразования в трафик;
- передача модели в модуль GUI для графического представления;
- передача модели в модуль анализа работы алгоритма для сравнения исходной модели аномального трафика и данных об обнаружении аномалий с сервера.

Модуль «Формирование и передача пакета» реализует создание трафика по модели, полученной с модуля «Генератора вектора параметров для анализа», а также для соединения клиентской части программного комплекса с серверной и передачи сгенерированного

трафика на серверную часть. Помимо этого, модуль передает данные о выбранном типе алгоритма анализа и его параметрах.

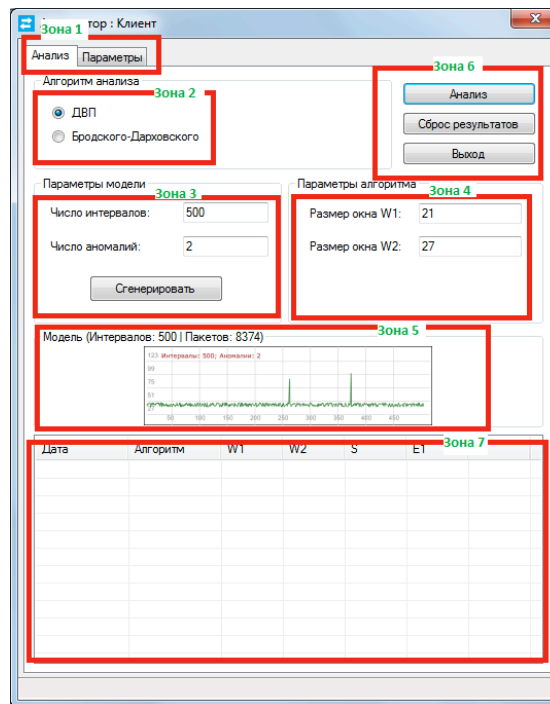


Рисунок 2 – Пользовательский интерфейс клиента. Вкладка «Анализ» (экранный снимок)

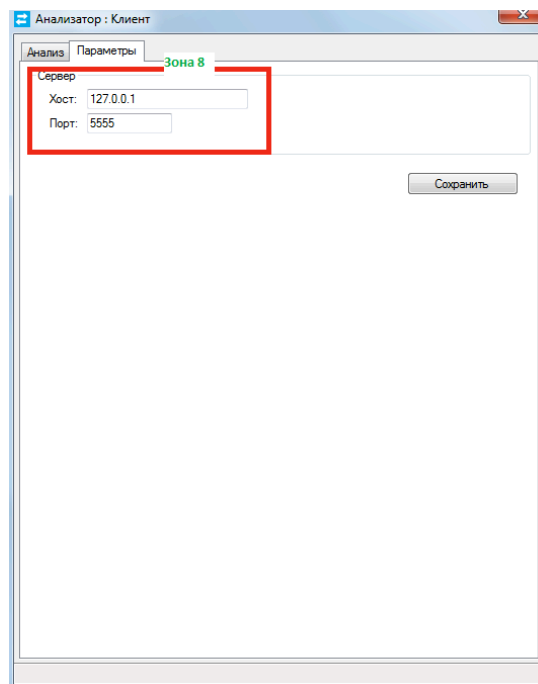


Рисунок 3 – Пользовательский интерфейс. Вкладка «Параметры» (экранный снимок)

Модуль «Анализ работы алгоритма» реализует прием и обработку ответа с серверной части, а также осуществляет анализ данных об обнаружении аномалий заданным алгоритмом путем сравнения исходной модели, принятой с модуля «Генератора вектора параметров для анализа», и ответа с модуля «Обнаружения сетевых аномалий» серверной части. Обеспечивает вывод данных анализа на пользовательский интерфейс клиентской части.

Серверная часть включает модуль «Обнаружения сетевых аномалий», который реализует подключение клиентов к серверной части, а также обрабатывает входящий трафик указанным алгоритмом, подключаемым из внешних библиотек. Также модуль реализует передачу данных об обнаружении сетевых аномалий модулю «Анализ работы алгоритма».

ПРОЦЕСС ФУНКЦИОНИРОВАНИЯ ПРОГРАММНОГО КОМПЛЕКСА

Функционирование разработанного программного комплекса для исследования качества можно разбить на два основных процесса: работа клиента и работа сервера.

Алгоритм работы сервера можно описать в виде следующей последовательности шагов, которые представлены в виде блок-схемы на рисунке 4:

1. Начало работы приложения. Сервер инициализирует параметры – размеры буфера, порта, пакета и потока, принимающего сообщения клиента (блоки № 1-2).

2. Сервер находится в режиме ожидания запросов от клиента (блоки № 3-4).

3. Если клиент отправляет запрос к серверу, то осуществляется прием трафика. Сервер обращается к библиотекам соответствующего алгоритма: NwaLib, MathNet.Numerics.dll, TaskParallelLibrary, WaveletStudio.dll. Далее осуществляется обработка трафика соответствующим алгоритмом. Сервер формирует ответный трафик по результатам проведенного анализа и отправляет ответ клиенту. Сервер возвращается в режим ожидания запросов от клиента (блоки № 5-9).

4. Если обращений к серверу более не обнаруживается, то сервер ждет команду выхода из приложения (блок № 10).

5. Конец работы приложения (блок № 11).

Алгоритм работы клиента можно описать в виде следующей последовательности шагов, которые представлены в виде блок-схемы на рисунке 5:

1. Начало работы приложения. Клиент фиксирует выбранный алгоритм для обнаружения аномалий (блоки № 1-2).

2. Клиент фиксирует параметры модели: число интервалов и количество аномалий (блок № 3).

3. Клиент генерирует модель соответствующую выбранным параметрам и выводит на экран ее графическое представление (блоки № 4-5).

4. Клиент фиксирует параметры алгоритма, введенные пользователем: размеры окон $W1$ и $W2$. $W1$ должно превышать значение $W2$. Сумма $W1$ и $W2$ не должна превышать количество интервалов (блоки № 6-7).

5. Если параметры алгоритма верны, то работа приложения переходит к выполнению блоков № 8-14. Клиент формирует трафик согласно модели. Далее отправляет запрос серверу на обработку трафика и ожидает ответ от сервера. Клиент анализирует полученный ответ от сервера и отображает данные в виде таблицы. При этом выводит значение окна $W1$, значение окна $W2$, количество правильно обнаруженных аномалий S , ошибки первого рода $E1$ и ошибки второго рода $E2$. Далее клиент ожидает сообщение от пользователя о выходе из приложения, но если этого не происходит, то работа приложения возвращается к блоку № 2.

6. Если параметры алгоритма не верны, то работа приложения переходит к выполнению блока № 9. Выводятся сообщения об ошибках: «Неверные параметры окна. Суммарный размер должен быть меньше общего размера интервалов!», «Неверные параметры окна. Окно $W1$ должно быть больше окна $W2$!». Далее приложение возвращается к работе блока № 3.

7. Работа приложения завершается (блок № 15).

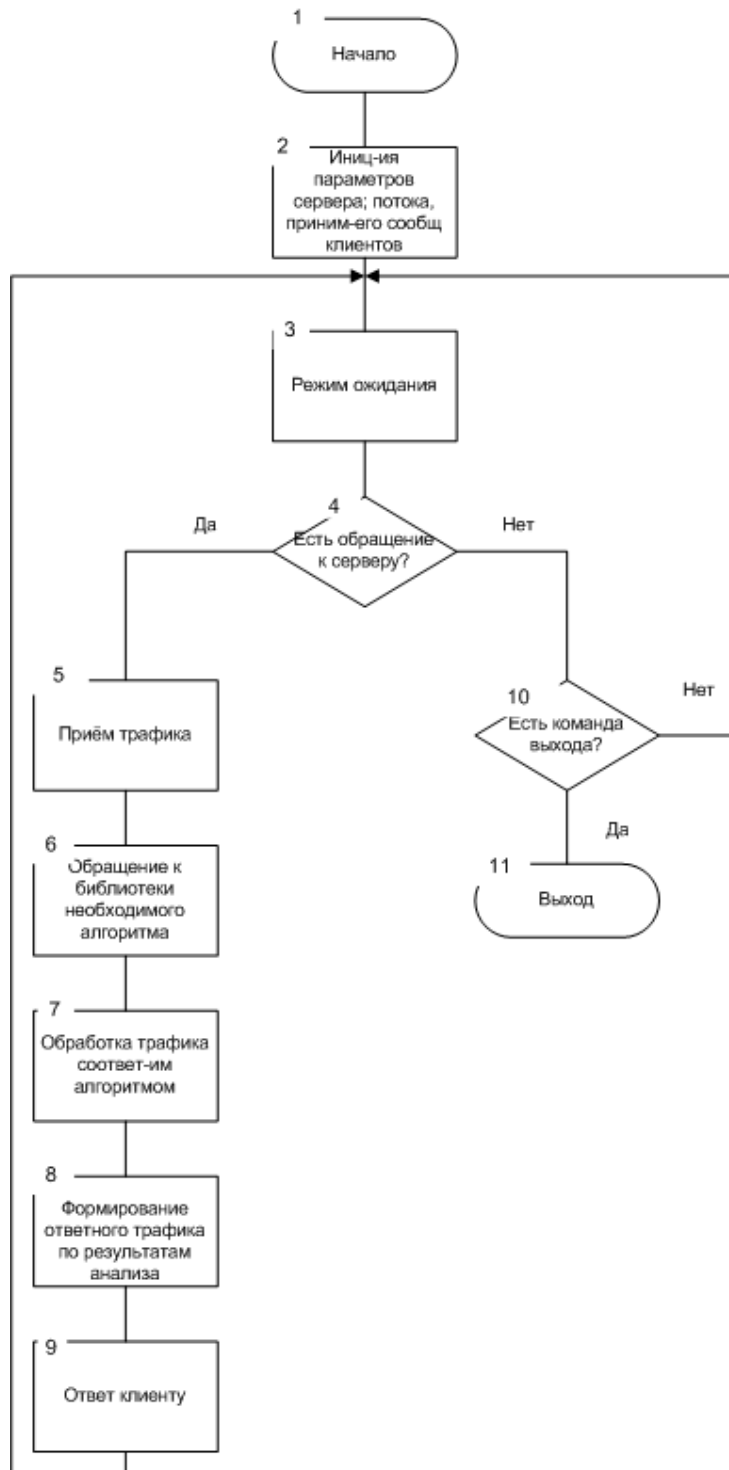


Рисунок 4 – Блок-схема алгоритма функционирования серверной части программного комплекса

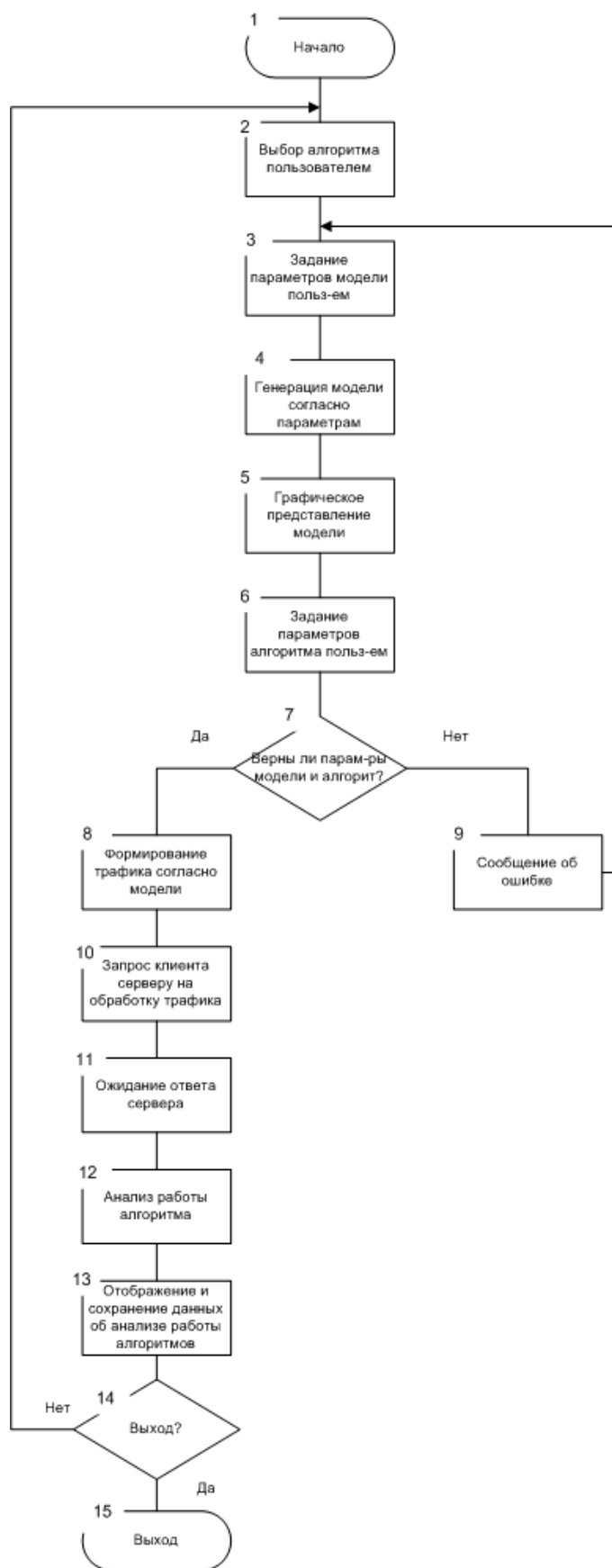


Рисунок 5 – Блок-схема алгоритма работы клиентской части программного комплекса

РЕЗУЛЬТАТЫ ЭКСПЕРИМЕНТАЛЬНОГО ИССЛЕДОВАНИЯ КАЧЕСТВА АЛГОРИТМОВ ОБНАРУЖЕНИЯ СЕТЕВЫХ АНОМАЛИЙ

С помощью разработанного программного комплекса было проведено множество экспериментов, направленных на:

- оценку влияния зависимостей входных параметров алгоритмов БД и ДВП на точность обнаружения аномалий;
- выбор наиболее качественного при классификации алгоритма обнаружения сетевых аномалий.

В результате проведения экспериментов для разных входных данных алгоритмов выявления сетевых аномалий при весовом коэффициенте $\beta = 2$ с учетом формул 1-3 были получены значения, представлены в таблице 3.

Таблица 3 – Результаты работы алгоритмов

№ этапа эксперимента	Алгоритмы	<i>Se</i> , %	<i>Sp</i> , %	<i>F-мера</i>
1 (I= 500; N= 4)	ДВП	60,60	66,67	61,73
	БД	25	25	25
2 (I= 300; N= 10)	ДВП	15,87	9,09	13,81
	БД	42,83	30	39,47
3 (I= 450; N= 15)	ДВП	14,29	11,11	13,51
	БД	44,44	50	45,45

Здесь N – общее число поданных сетевых аномалий; I – количество интервалов.

Анализ полученных данных показывает, что в среднем метрика качества – F-мера алгоритма обнаружения сетевых аномалий БД – на 20% превышает F-меру алгоритма ДВП. Кроме того, в двух случаях из трех алгоритм БД показал лучшее значение чувствительности, что указывает на меньшее число ошибок первого рода по сравнению с алгоритмом ДВП, что позволяет сделать вывод о более высоком качестве алгоритма при его использовании для решения задач классификации сетевых аномалий.

ЗАКЛЮЧЕНИЕ

Разработанный программный комплекс позволяет исследовать характеристики алгоритмов обнаружения сетевых аномалий с целью оценки качества алгоритма и выбора наиболее подходящего для решения задач классификации. Полученные при его использовании результаты показывают, что наиболее чувствительным при обнаружении аномалий является алгоритм Бродского-Дарховского. Также при его использовании обнаруживается меньше ошибок 1-ого и 2-ого родов, чем при использовании алгоритма на основе дискретного вейвлет-преобразования с применением статистических критериев. Кроме того, алгоритм Бродского-Дарховского имеет лучшую величину F-меры.

СПИСОК ЛИТЕРАТУРЫ

1. Шелухин О.И., Филинова А.С. Обнаружение сетевых аномальных выбросов трафика методом разладки Бродского-Дарховского // Т-Comm – Телекоммуникации и транспорт, 2013. – № 10. – Том 7. – С. 116-118.
2. Шелухин О.И., Панкрушин А.П. Оценка достоверности обнаружения аномалий сетевого трафика методами дискретного вейвлет-преобразования // Т-Comm – Телекоммуникации и транспорт, 2013. – № 10. – Том 7. – С. 110-113.

3. Аграновский А.В., Милославская Н.Г., Толстой А.И., Селин Р.Н. Выявление угроз безопасности как способ предотвращения атак на компьютерные системы // Безопасность информационных технологий, 2008. – № 1. – С. 5-15.

Микова Софья Юрьевна

Волгоградский государственный университет, г. Волгоград
Студентка кафедры информационной безопасности

Оладко Владлена Сергеевна

Волгоградский государственный университет, г. Волгоград
Кандидат технических наук, доцент кафедры информационной безопасности
E-mail: oladko.vs@yandex.ru

S.Yu. MIKOVA (*Student of the Department of Information Security*)

V.S. OLAD'KO (*Candidate of Engineering Sciences,
Associate Professor of the Department of Information Security*)
Volgograd State University, Volgograd

SOFTWARE FOR RESEARCH QUALITY ALGORITHMS TO DETECT NETWORK ANOMALIES

The article describes the approach to the assessment of the accuracy of anomaly detection algorithms occurring in information systems and technology networks. We identified the cause of network anomalies and algorithms that allow them to identify. We have developed software system architecture and the algorithms of the client and server. We investigated algorithms: the algorithm based on discrete wavelet transform using statistical criteria and anomaly detection algorithm Brodsky-Darhovsky. As a result of experiments and quality assessment algorithm was found having the best value of F-measure and sensitivity.

Keywords: *attacks; the information system; the attacker; false negative; false positive; the window size.*

BIBLIOGRAPHY (TRANSLITERATED)

1. Sheluxin O.I., Filinova A.S. Obnaruzhenie setevy'x anomal'ny'x vy'brosov trafika metodom razladki Brodskogo-Darhovskogo // T-Comm – Telekommunikacii i transport, 2013. – № 10. – Том 7. – С. 116-118.
2. Sheluxin O.I., Pankrushin A.P. Ocenka dostovernosti obnaruzheniya anomalij setevogo trafika metodami diskretnogo vejvlet-preobrazovaniya // T-Comm – Telekommunikacii i transport, 2013. – № 10. – Том 7. – С. 110-113.
3. Agranovskij A.V., Miloslavskaya N.G., Tolstoj A.I., Selin R.N. Vy'yavlenie ugroz bezopasnosti kak sposob predotvrashheniya atak na komp'yuterny'e sistemy' // Bezopasnost' informacionnyx texnologij, 2008. – № 1. – С. 5-15.

АДАПТАЦИЯ МЕТОДА МАКСИМАЛЬНОГО ПРАВДОПОДОБИЯ ВИТЕРБИ К ДЕКОДИРОВАНИЮ ХАОТИЧЕСКИХ ПОСЛЕДОВАТЕЛЬНОСТЕЙ

В статье рассмотрен подход к декодированию двоичной хаотической последовательности, закодированной с использованием дискретных хаотических отображений, с точки зрения критерия максимального правдоподобия.

Ключевые слова: декодирование; алгоритмы декодирования; хаос; хаотическая система связи.

Интерес к хаотическим системам передачи информации связан с хорошими свойствами сигналов, создаваемых хаотическими генераторами с точки зрения разработки защищенных систем связи и систем множественного доступа. В этих случаях можно использовать некоррелированность и непредсказуемость хаотических сигналов для построения алгоритмов кодирования.

Одной из возможностей является разработка алгоритмов кодирования и декодирования, основанных на дискретных хаотических отображениях. Для использования свойств сгенерированной хаотической последовательности в присутствии аддитивного гауссовского белого шума необходимо подобрать подходящий метод декодирования. В реальной системе, где длина N сообщения может достигать тысяч бит, процесс кодирования подразумевает практически бесконечную точность вычислений. В этом случае, данный метод можно применять для кодирования блоками из $D \ll N$ бит за один раз.

Прямое декодирование хаотической последовательности показывает во всех случаях худшие результаты, чем некодированная передача информационного сообщения, поскольку квадрат расстояния между каждым из отсчетов и пороговым значением может оказаться достаточно малым. Фактически при прямом декодировании мы лишаемся всех преимуществ системы, заключающихся в наличии избыточности в смежных символах, так как у каждого символа x_n есть $D - 1$ бит общих с x_{n-1} и x_{n+1} символов, $D - 2$ с x_{n-2} и x_{n+2} и т.д. В соответствии с этим предлагается метод декодирования, основанный на критерии максимального правдоподобия (МП) [1] со скользящим окном [2]. Он является адаптацией известного алгоритма Витерби [3]. Этот метод работает на базисе символической динамики [4], поэтому необходимо квантование фазового пространства $[0,1]$. В алгоритмах, описанных в работе [6], применялась символическая динамика при восстановлении траекторий S_I , поскольку каждая из ветвей обратной функции $f_1^{-1}(x)$ или $f_2^{-1}(x)$ соответствовала подинтервалу $\left[0, \frac{1}{2}\right)$ (закодирован бит $b = 0$) или $\left[\frac{1}{2}, 1\right]$ (закодирован бит $b = 1$).

Рассматриваемый алгоритм требует аналогичного деления на подинтервалы, однако для верного установления динамики системы их требуется большее количество. В соответствии с этим разделим интервал $[0,1]$ на непересекающиеся подинтервалы I_i с границами $\frac{i}{p}$ и $\frac{i+1}{p}$, $i = 0, \dots, p-1$ и центрами $c_i = \frac{i}{p} + \frac{i}{2p}$. Количество подинтервалов p выбирается четным, поэтому пороговая точка $x = \frac{1}{2}$ является верхней границей одного

подинтервала и нижней границей другого. Таким образом, зная подинтервал I_i , в который попадает отсчет, можно сказать, должен он быть декодирован как 0 или 1. Более того, если заменить исходную последовательность отсчетов последовательностью подинтервалов, в которых лежат соответствующие отсчеты (другими словами, хаотическая последовательность заменяется последовательностью центров c_i), то получим символическое представление последовательности, которое можно описать как марковский процесс первого порядка с соответствующей матрицей перехода \mathbf{T} [1, 7]. Элемент t_{ij} этой матрицы означает возможность перехода между интервалом i и интервалом j . В случае отображения сдвига Бернулли [6] каждый интервал отображается в два непрерывных интервала с равной вероятностью (пропорциональной длине исходного подинтервала отображенного на целевой подинтервал). Например, для случая $p = 4$ матрица перехода примет вид

$$\mathbf{T} = \frac{1}{2} \begin{pmatrix} 1 & 1 & 0 & 0 \\ 0 & 0 & 1 & 1 \\ 1 & 1 & 0 & 0 \\ 0 & 0 & 1 & 1 \end{pmatrix}.$$

В случае модифицированного логистического отображения и перевернутого логистического отображения матрицу перехода получить непросто. Она строится как отношение длины пересечения отображенного интервала $f(I_i)$ и целевого интервала I_j к длине изображения $f(I_i)$

$$t_{ij} = \frac{\text{length}[f(I_i) \cap I_j]}{\text{length}[f(I_i)]},$$

где $i = 0, \dots, p - 1$ и $j = 0, \dots, p - 1$.

Если интервал задан нижней границей a_l^i и верхней границей a_u^i , то длина интервала $\text{length}(I_i) = a_u^i - a_l^i$.

В процессе декодирования будем рассматривать последовательность центров подинтервалов d_k^i ($k = 0, \dots, L - 1$) в пределах скользящего окна длиной L , где $d_k^i = c_i$, если x_{n+k} лежит в I_i . Подобным образом можно сказать, что марковский процесс находится в состоянии $s_k = i$ в момент времени $k = 0, \dots, L - 1$, если x_{k+n} лежит в I_i . Начальное состояние s_0 полагают первоначально равным нулю, а по мере декодирования методом скользящего окна начальное состояние выбирается из ранее рассчитанных значений для каждого декодируемого блока длиной L . Таким образом учитывается избыточность всей последовательности.

Для применения алгоритма Витерби в виде, описанном в [1], необходимо взять блок из L принятых символов y_n, \dots, y_{n+L-1} и выполнить следующие действия.

1. Инициализация алгоритма

$$\begin{aligned}\delta_0^i &= \pi_i b_i(y_n) & 0 \leq i < p \\ \phi_0^i &= 0 & 0 \leq i < p\end{aligned}\quad (1)$$

2. Прямой проход

$$\delta_k^j = \max_{0 \leq i < p} \left\{ \delta_{k-1}^i t_{ij} b_j(y_{n+k}) \right\} \quad \begin{matrix} 0 < k < L \\ 0 \leq j < p \end{matrix} \quad (2)$$

$$\phi_k^j = \arg \max_{0 \leq i < p} \left\{ \delta_{k-1}^i t_{ij} b_j(y_{n+k}) \right\} \quad \begin{matrix} 0 < k < L \\ 0 \leq j < p \end{matrix} \quad (3)$$

3. Остановка

$$\hat{s}_{p-1}^n = \arg \max_{0 \leq i < p} \left\{ \delta_{p-1}^i \right\} \quad (4)$$

4. Обратный проход

$$\hat{s}_k^n = \phi_{k+1}^{\hat{s}_{k+1}^n} \quad k = L-2, \dots, 0, \quad (5)$$

где t_{ij} соответствует вероятности перехода марковского процесса, $b_j(y_k)$ соответствует вероятности перехода канала [8]

$$b_j(y_k) = p(y_k | c_j) = \frac{1}{\sqrt{2\pi\sigma}} e^{-\frac{(y_k - c_j)^2}{2\sigma^2}}, \quad (6)$$

и π_i – априорная вероятность x_n по отношению к интервалу I_i . Эти вероятности могут быть вычислены по ПРВ отображения как

$$\pi_i = F\left(\frac{i+1}{p}\right) - F\left(\frac{i}{p}\right), \quad (7)$$

где $F(x)$ – функция распределения.

Последовательность $d_k^{\hat{s}_k^n}$, $k = 0, \dots, L-1$, квантованная по критерию максимального правдоподобия, соответствует блоку принятых символов y_n, \dots, y_{n+L-1} . Для продолжения алгоритма выбираем $d_0^{\hat{s}_0^n}$, принимаем решение о принятом бите

$$\hat{b}_{n+1} = \left[d_0^{\hat{s}_0^n} + \frac{1}{2} \right] \quad (8)$$

и выполняем аналогичные действия над следующим блоком из L символов, y_{n+1}, \dots, y_{n+L} . Отметим, что в отличие от алгоритмов эвристического декодирования, нет необходимости в нормировании принятой последовательности y_n к интервалу $[0,1]$ [2].

В случае декодирования Витерби при $p = 32$ и $L = 10$ результаты практически аналогичны полученным, в случае рекурсивного эвристического декодирования [9] с 7 итерациями и $M = 4$. Тем не менее алгоритм максимального правдоподобия более сложен, чем эвристическое декодирование с точки зрения числа операций сложения и умножения, требуемых для обработки сравнимого количества параметров.

СПИСОК ЛИТЕРАТУРЫ

1. Kisel A., Dedieu H., Schimming T. Maximum likelihood approaches for noncoherent communications with chaotic carriers // IEEE Transactions on Circuit and Systems, 2001. – Vol. 48. – № 11. – P. 533-542.
2. Escribano F.J., Lopez L., Sanjuan M.A.F. Exploiting symbolic dynamics in chaos coded communications with maximum a posteriori algorithm // Electron. Lett, 2006. – Vol. 42. – P. 984-985.
3. Viterbi A. Error bounds for convolutional codes and an asymptotically optimum decoding algorithm // IEEE Transactions on information theory, 1967. – Vol. IT-13. – P. 260-269.
4. Schweizer J., Schimming T. Symbolic dynamics for processing chaotic signals – II: Communication and coding // IEEE Transactions of circuits and systems, 2001. – Vol. 48. – №.11. – P. 1283-1295.
5. Митин С.В. Кодирование двоичной последовательности с использованием дискретных хаотических отображений // Электронное научно-техническое издание «Наука и образование» (МГТУ им. Н.Э. Баумана), 2012. – № 6.
6. Papoulis A., Pillai S.U. Probability, random variables and stochastic processes. – McGraw-Hill, Inc., Boston, 2002. – 852 p.
7. Прокис Дж. Цифровая связь: пер. с англ. – М.: Радио и связь, 2000. – 800 с.
8. Митин С.В. Алгоритмы декодирования двоичной последовательности, закодированной с использованием дискретных хаотических отображений // Информационные системы и технологии, 2013. – № 3(77). – С. 91-101.

Митин Сергей Владимирович
Академия ФСО России, г. Орел
Доцент
E-mail: tician@yandex.ru

S.V. MITIN (*Associate Professor*)
Academy of Federal Agency of Protection of the Russian Federation, Orel

ADAPTATION MAXIMUM LIKELIHOOD VITERBI'S METHOD FOR DECODING OF CHAOTIC SEQUENCES

The article describes the approach to the decoding of binary chaotic sequence encoded using discrete chaotic maps, in terms of the criterion of maximum likelihood.

Keywords: *decoding; decoding algorithms; chaos; chaotic communication.*

BIBLIOGRAPHY (TRANSLITERATED)

1. Kisel A., Dedieu H., Schimming T. Maximum likelihood approaches for noncoherent communications with chaotic carriers // IEEE Transactions on Circuit and Systems, 2001. – Vol. 48. – № 11. – P. 533-542.

2. Escibano F.J., Lopez L., Sanjuan M.A.F. Exploiting symbolic dynamics in chaos coded communications with maximum a posteriori algorithm // *Electron. Lett*, 2006. – Vol. 42. – P. 984-985.
3. Viterbi A. Error bounds for convolutional codes and an asymptotically optimum decoding algorithm // *IEEE Transactions on information theory*, 1967. – Vol. IT-13. – P. 260-269.
4. Schweizer J., Schimming T. Symbolic dynamics for processing chaotic signals – II: Communication and coding // *IEEE Transactions of circuits and systems*, 2001. – Vol. 48. – №.11. – P. 1283-1295.
5. Mitin S.V. Kodirovanie dvoichnoj posledovatel'nosti s ispol'zovaniem diskretny'x xaoticheskix otobrazhenij // *E'lektronnoe nauchno-texnicheskoe izdanie «Nauka i obrazovanie» (MGTU im. N.E'. Baumana)*, 2012. – № 6.
6. Papoulis A., Pillai S.U. *Probability, random variables and stochastic processes*. – McGraw-Hill, Inc., Boston, 2002. – 852 p.
7. Prokis Dzh. *Cifrovaya svyaz': per. s angl.* – M.: Radio i svyaz', 2000. – 800 s.
8. Mitin S.V. Algoritmy' dekodirovaniya dvoichnoj posledovatel'nosti, zakodirovannoj s ispol'zovaniem diskretny'x xaoticheskix otobrazhenij // *Informacionny'e sistemy' i texnologii*, 2013. – № 3(77). – S. 91-101.

УДК 004.01; 621.391

А.П. ФИСУН, Р.А. ФИСУН

**МЕТОДИКА ОЦЕНКИ ЭФФЕКТИВНОСТИ СРЕДСТВ
КОМПЛЕКСНОГО ОБЕСПЕЧЕНИЯ ИНФОРМАЦИОННОЙ
БЕЗОПАСНОСТИ ИНФОРМАЦИОННЫХ СИСТЕМ
И ИНФОРМАЦИОННО-ТЕЛЕКОММУНИКАЦИОННЫХ СЕТЕЙ
ОБЪЕКТОВ ИНФОРМАТИЗАЦИИ ОРГАНОВ ПУБЛИЧНОЙ ВЛАСТИ**

В статье изложена методика оценки эффективности средств обеспечения информационной безопасности информационных систем (ИС) и информационно-телекоммуникационных сетей (ИТКС) объектов информатизации Федеральных органов государственной власти, органов государственной власти субъектов Российской Федерации, характеризующихся значительным многообразием и разнообразием количественных, качественных показателей и критериев эффективности этих технологий и функционирующих в сложных условиях неопределенности, риска, воздействия внешних и внутренних угроз.

Ключевые слова: *информационные системы (ИС); информационно-телекоммуникационные сети (ИТКС); информационно-телекоммуникационные технологии (ИКТ); объекты информатизации; информационная безопасность ИС; ИТКС; ИКТ; средства комплексного обеспечения информационной безопасности и защиты информации (СКОИБ); оценка эффективности СКОИБ.*

Проблема выбора эффективных вариантов средств комплексного обеспечения информационной безопасности ИС, ИТКС, ИКТ объектов информатизации как органов публичной власти органов публичной власти, так и любых органов, организаций, учреждений, предприятий, использующих современные информационно-телекоммуникационные технологии (ИКТ), связана не только со сложностью получения количественных экономических показателей СКОИБ, но и с объективной, интегральной технико-экономической оценкой функциональных показателей СКОИБ. Существующие методики оценки технико-экономической эффективности альтернатив зачастую рассчитаны на экономические показатели себестоимости средств, мероприятий и работ, сроков их окупаемости и другие затраты ресурса, что при всей важности получения формальных экономических расчетов не позволяет объективно интерпретировать полученные результаты с учетом эффекта выбранных или заданных функциональных показателей разрабатываемых, модернизируемых, выбираемых СКОИБ или их компонентов.

В качестве варианта оценки искомой эффективности можно рассмотреть методику, основанную на известных методиках вероятностной оценки обеспечения информационной безопасности (ОИБ) ИС, ИТКС, ИКТ объектов информатизации [1-6] и векторной оценки выигрыша [7] по выбранным основным функциональным и экономическим показателям СКОИБ, представленную следующими этапами:

– этап 1 – определение, формирование, разработка системы интегрального, комплексных и единичных показателей качества СКОИБ, ее компонентов (модели, алгоритма, комплекса, средства, компонента), отражающей объективные, основные свойства СКОИБ, и группирование [7-11] этих показателей на показатели функционирования СКОИБ, временные и экономические;

– этап 2 – выбор, определение идеальных требований к каждому из входящих в систему интегрального, комплексных и единичных показателей качества (СПК) для разрабатываемой, модернизируемой СКОИБ или ее компонента;

– этап 3 – нормализация множества показателей качества СПК и их критериев оптимальности, введение меры в пространство показателей и критериев СПК, выбор метода свертки векторного показателя эффективности СКОИБ в обобщенный скалярный показатель эффективности;

– этап 4 – окончательное определение и формулирование интегрального (обобщенного) показателя эффективности (критерия оптимальности) СКОИБ или ее компонента;

– этап 5 – осуществление оценки комплексных, единичных, являющихся частными к интегральному (обобщенному) показателям технико-экономической эффективности;

– этап 5 – осуществление оценки интегрального (обобщенного) показателя эффективности СКОИБ (ИКТ СКОИБ, алгоритма, компонента);

– этап 6 – определение эффективности по выбранному критерию (оптимальности, превосходства, применимости), полученного выигрыша в технико-экономических показателях СПК выбранной, разработанной, модернизируемой СИБ или ее компонентов;

– этап 7 – определение на основе методики вероятностной оценки ОИБ, значения всех показателей СПК для существующего (базового, типового) варианта СКОИБ, которые формируют в пространстве показателей соответствующую точку;

– этап 8 – определение для разрабатываемой, модернизируемой СКОИБ (ИС, ИТКС, ИКТ СКОИБ, алгоритма, компонента) на основе методики вероятностной оценки ОИБ, реализованные значения показателей СПК, которые формируют в пространстве показателей соответствующую точку;

– этап 9 – определение на основе методов нормализации, принятой меры и свертки комплексных, единичных (частных), интегрального (общего) показателей эффективности существующей (типовой, базовой) и разработанной, модернизированной СКОИБ (ИС, ИТКС, ИКТ СКОИБ, алгоритма, компонента) и оценка выигрыша в эффективности;

– этап 10 – формулировка выводов об эффективности по выбранным критериям (оптимальности, превосходства, применимости) варианта СКОИБ (ИС, ИТКС, ИКТ СКОИБ, алгоритма, компонента) по его технико-экономическим показателям.

Важным этапом оценки искомой эффективности является формирование интегральной иерархической СПК (ИСПК) разрабатываемой СКОИБ, отражающей большое разнообразие и многообразие взаимосвязанных компонентов, подсистем СКОИБ (комплексного обеспечения ИБ и безопасности информации, в том числе правового, организационного, технического, программно-аппаратного, криптографического, некриптографического обеспечения; управления (СУ) СКОИБ; других), сложность процесса функционирования, активное взаимодействие с внешней средой. Разработка такой ИСПК предполагает [1] определение и формирование множества локальных СПК, представляющих систему комплексных и единичных показателей (КЕ-показатели или СПК КЕ), характеризующих свойства компонентов СКОИБ, обеспечивающих ее эффективное функционирование. Объективность, полнота, релевантность ИСПК СКОИБ определяется научно-обоснованными критериями, требованиями, предъявляемыми к СКОИБ, ее компонентам и процессу функционирования, а также корректной декомпозицией ее структуры на КЕ-показатели.

Свойства искомой СКОИБ ИС, ИТКС ИКТ объекта информатизации определяются внешними требованиями к ИС, ИТКС ИКТ, обеспечением одного из основных свойств информации, информационных услуг и продуктов – безопасностью информации (БИ), которая, в свою очередь, определяет требования к внутренним свойствам СКОИБ, ее компонентам, средствам обеспечения ИБ и БИ (СЗИ), средствам системы управления (ССУ), процессам функционирования СКОИБ, взаимосвязь которых представлен на рисунке 1.



Рисунок 1 – Система показателей качества СКОИБ

Обозначения на рисунке 1:

– $\vec{K}_\phi(k) = [t_{gom}(i); t_{nodf}(i); \vec{R}_\phi(i)]$ – ИСПК, характеризующий степень обеспечения ИБ и БИ;

– $t_{gom}(i)$ – время готовности СКОИБ на i -ом шаге функционирования СКОИБ;

– $t_{nodf}(i)$ – недоступности защищаемой информации для потенциального нарушителя для заданного процесса функционирования СКОИБ i -ом шаге функционирования СКОИБ;

– $\vec{R}_\phi(i)$ – вектор затрат ресурсов на обеспечение процесс функционирования СКОИБ i -ом шаге функционирования;

– $\vec{K}_{fcszi}(i) = [t_{gomcssi}(i); t_{nodfcssi}(i); \vec{R}_{fcszi}(i)]$ – СПК КЕ процесса обеспечения ИБ и ЗИ;

– $t_{gomcssi}(i)$ – время готовности СЗИ на i -ом шаге функционирования СКОИБ;

– $t_{nodfcssi}(i)$ – недоступности защищаемой информации для потенциального нарушителя для заданного процесса функционирования СЗИ на i -ом шаге функционирования СКОИБ;

– $\vec{R}_{fcszi}(i)$ – вектор затрат ресурсов на обеспечение процесс функционирования СЗИ на i -ом шаге функционирования.

– $\vec{K}_y(i) = [T_{uy}(i); \Delta \vec{K}_{fcszi}^y(i); t_{duy}(i); \vec{R}_y(i)]^T$ – СПК КЕ показателей управления СКОИБ,

представляющий нижестоящий уровень иерархии отношению с СПК КЕ процесса функционирования СКОИБ, включающий:

– $T_{uy}(i)$ – длительность цикла управления параметрами функционирования СЗИ при нарушении нормальных условий функционирования на i -ом шаге функционирования;

– $\Delta \vec{K}_{fcszi}^y(i)$ – вектор приращений значений СПК КЕ СЗИ на i -ом шаге функционирования СЗИ, обусловленных ошибками в контуре управления СКОИБ;

– $t_{дв}(i)$ – время доступности к сигналам управления на i -ом шаге функционирования СЗИ;

– $\vec{R}_y(i)$ – вектор затрат ресурсов управления на i -ом шаге функционирования СЗИ.

Основные компоненты СКОИБ – СЗИ и ССУ, представляющие материально-технический базис процессов обеспечения ИБ и ЗИ, управления СКОИБ, могут быть представлены соответствующей иерархией СПК КЕ СЗИ и ССУ (1), (2):

$$\vec{K}_{СИ}(i) = [M_{УСЗИ}(i); M_{БИСЗИ}(i); M_{РЗСИ}(i); M_{РСЗИ}(i)]^T, \quad (1)$$

$$\vec{K}_{СУ}(i) = [M_{УСУ}(i); M_{БИСУ}(i); M_{РЗСУ}(i); M_{РСУ}(i)]^T. \quad (2)$$

где $M_{УСЗИ}(i)$ и $M_{УСУ}(i)$ – матрицы единичных показателей и параметров устойчивости элементов СЗИ и ССУ, соответственно, на i -ом шаге функционирования СКОИБ; $M_{БИСЗИ}(i)$ и $M_{БИСУ}(i)$ – матрицы показателей комплексного обеспечения безопасности информации, используемыми СЗИ и ССУ на i -ом шаге функционирования СКОИБ; $M_{РЗСИ}(i)$ и $M_{РЗСУ}(i)$ – матрицы значений показателей и параметров разведзащищенности СЗИ и ССУ на i -ом шаге функционирования СКОИБ; $M_{РСЗИ}(i)$ и $M_{РСУ}(i)$ – матрицы затрат ресурсов на разработку и эксплуатацию СЗИ и ССУ на i -ом шаге функционирования СКОИБ.

Результатом первого этапа оценки является сформированное множество показателей качества разрабатываемой или модернизируемой СКОИБ, позволяющих осуществить сравнение эффективности такой СКОИБ с существующими, базовыми образцами.

На втором этапе обосновываются или выбираются идеальные, нормируемые, справочные требования (критерии) к СКОИБ для сформированных ИСПК и СПК КЕ. При этом в качестве методики обоснования затрат ресурса для разрабатываемой или модернизируемой СКОИБ и ее СЗИ, ССУ может быть применена методика количественной оценки угроз, БИ и ИБ ИКТ [1-6]. В качестве критериев к определению показателей СЗИ могут быть использованы потенциально возможные показатели и параметры СЗИ, ССУ, пределы их изменений и другие критерии, которые позволяют сформировать в пространстве разработанных показателей качества определенную идеальную точку, а также представляют потенциальную прогнозную оценку для последующих разработок и модернизации СЗИ и ССУ СКОИБ.

На третьем этапе осуществляются процедуры нормирования принятой системы ИСПК и ее СПК КЕ, выбор метода свертки векторного в скалярный показатель качества при сформированной ИСПК $\vec{Y} = (\vec{Y}_1, \dots, \vec{Y}_n, \dots, \vec{Y}_N)^T$ оцениваемого варианта СКОИБ и выбранном критерии оптимальности (3).

$$I_n = L(Y, Y_{тр}), \quad (3)$$

где целевая функция $L(Y, Y_{тр})$ представляет систему ИСПК и требований к ним, а также заданные критерии оптимальности (min , max , $min\ max$, $max\ min$ и другие). Для случая, когда целевая функция $L(Y, Y_{тр})$ отражает степень достижения показателем качества внешних требований к СКОИБ, она может выступать как частный показатель СПК КЕ эффективности используемых СЗИ, ССУ.

В силу значительного многообразия и разнообразия СЗИ и ССУ, обуславливающих такое же многообразие и разнообразие соответствующих их количественных и качественных показателей эффективности и критериев оптимальности, возникает проблема многокритериальной (векторной) оценки эффективности. Сложности решения этой

проблемы, в свою очередь, обусловлены необходимостью нормализации и последующей скаляризации, свертки, составляющих компонентов векторного показателя эффективности.

Процесс нормализации векторного показателя или критерия может быть выполнен в несколько известных шагов линейного преобразования (4) [7, 11]:

- шаг 1 – переход к единой размерности (безразмерности);
- шаг 2 – сведение к одной точке отсчета и переход к равноценным шкалам (одному масштабу).

$$I_n^{(n)}(\bar{x}; k) = c_n I_n(\bar{x}; k) + d_n, \quad (4)$$

где $c_n = \frac{1}{(I_n^*(\bar{x}; k) - I_n^0(\bar{x}; k))}$ – масштабный коэффициент; $d_n = -\frac{I_n^0(\bar{x}; k)}{(I_n^*(\bar{x}; k) - I_n^0(\bar{x}; k))}$ – коэффициент сдвига, корректирующий начало отсчета; $I_n^{(n)}, I_n^*, I_n^0$ – нормированное, наибольшее и наименьшее значения критериев.

Преобразование (4) позволяет привести все показатели и критерии к нулевой точке отсчета, шкалу их изменения – к отрезку $[0,1]$, сделать все показатели и их критерии безразмерными. При изменении критерия с *max* на *min* в выражении (4) осуществляется замена знака при коэффициентах c_n и d_n на обратные.

Решение задачи оптимизации СЗИ и ССУ СКОИБ по векторному критерию требует нахождения решений, удовлетворяющих экстремуму одновременно по показателям всех компонент СКОИБ, при котором одновременно может быть получена оценка близости предложенного варианта СКОИБ (СЗИ, ССУ) к потенциально возможным и определен эффект по каждому из показателей ИСПК. Известны [7, 8, 11 и др.] следующие пути решения искомой задачи:

- поиск компромиссных решений, оптимальных по Парето;
- поиск решений, оптимальных в смысле обобщенного скалярного критерия, полученного путем свертки, скаляризации, сформированных показателей всех компонент векторного критерия оптимизации (ВКО).

Поиск решений по Парето, как известно [7, 8, 11 и др.], затруднен необходимостью использования строгих математических методов оптимизации для широкого круга задач, а также отсутствием, как правило, единственности искомого решения, что дает основания использовать его для вспомогательного, предварительного уменьшения размерности исходного множества решений до этапа свертки ВКО.

Второй метод заключается в сведении векторной задачи оптимизации к скалярной путем формирования обобщенного критерия, значение которого для различных вариантов разрабатываемой СКОИБ представляет проекцию всех компонент ВКО на одну числовую ось. Это облегчает практический окончательный выбор варианта СКОИБ и её СЗИ и ССУ.

Среди известных групп методов скаляризации векторных критериев оптимизации (рисунок 2) [7, 8, 11 и др.] выбран метод идеальной точки для оценки обобщенного критерия оптимальности (5).



Рисунок 2 – Методы скаляризации векторных критериев оптимальности

$$I^{обобщ.}(\vec{x}) = \left[\sum_{n=1}^N (I_n^{(n)}(\vec{x}) - I_n^{(n)идеал.})^q \right]^{\frac{1}{q}}, \quad (5)$$

где $q = 1, 2, \dots$ – степень целевой функции; \vec{x} – вектор, оптимизируемых по ВКО параметров; $I_n^{идеал.}$ – идеальные значения критериев, принимаемые экстремальные значения n -х критериев, или внешние требования к их значениям разрабатываемых СЗИ, ССУ СКОИБ.

Для стохастических систем в выражении (3) выборочные значения критериев заменяются на их средние значения $\bar{I} = M\{\bar{I}\}$, где $M\{.\}$ – математическое ожидание.

На четвертом этапе с учетом выбранного метода свертки рассчитываются значения обобщенного показателя эффективности для разработанного и типового, базового вариантов СЗИ, ССУ СКОИБ и определяется эффективность разработанного варианта (6):

$$V_{эф} = \frac{(\bar{I}_{баз.образцаВТС} - \bar{I}_{нов.образцаВТС})}{\bar{I}_{баз.образцаВТС}} \cdot 100\%. \quad (6)$$

На пятом этапе формулируются выводы о технико-экономической эффективности разработки, модернизации внедрения предлагаемого варианта СЗИ, ССУ, если выигрыш в эффективности составляет величину более 5% как для модернизируемых, так и вновь разрабатываемых вариантов СЗИ, СУ СКОИБ.

СПИСОК ЛИТЕРАТУРЫ

1. Фисун А.П., Еременко В.Т. Оценка безопасности информации, обрабатываемой комплексами технических средств систем управления специальной связи // Сборник научных трудов Академии МВД России. – М.: Академия МВД РФ, 1994.

2. Фисун А.П. Моделирование и оценка угроз информационной безопасности // Сборник материалов 8-й Международной конференции «Информатизация правоохранительных систем» (МФИ-99). – Часть 2. – М.: Академия МВД РФ, 1999.
3. Фисун А.П., Лебедев В.Н. Определение и классификация угроз безопасности информационно-вычислительных сетей // Труды Академии управления МВД России «Компьютерные технологии и управление ОВД». – М.: Академия управления МВД РФ, 2000.– С. 67-78.
4. Фисун А.П., Минаев В.А. и др. Теоретические основы информатики и информационная безопасность: монография. – М.: Радио и связь, 2000. – 466 с.
5. Фисун Р.А. Моделирование угроз информационной безопасности информационно-телекоммуникационных систем: дипломный проект. – Орел, Академия ФАПСИ, 2003.
6. Фисун А.П. и др. Моделирование угроз информационной безопасности человеко-компьютерных систем / А.П. Фисун, К.А. Джевага, Р.А. Фисун, Л.А. Гращенко // Сборник материалов 2-й Всероссийской НПК «Методы и средства технической защиты конфиденциальной информации», 7-9 июня 2005 года. – Обнинск, ГОУ ГЦИПК, 2005.
7. Терентьев В.М. Методика обоснования требований к показателям качества АСМКРС. – Л.: ВАС, 1991.
8. Антушев Г.С. Методы параметрического синтеза сложных технических систем. – М.: Наука, 1989. – 88 с.
9. Феллер В. Введение в теорию вероятностей и ее приложения: пер. с англ. – В 2-х томах. – Т. 1. – М.: Мир, 1984. – 528 с.
10. Левин Б.Р. Вероятностные модели и методы в системах связи и управления.– М.: Радио и связь, 1985. – 312 с.
11. Венцель Е.С. Теория вероятностей. – М.: ВШ, 1998. – 576 с.
12. Ланнэ А.А., Улахович Д.А. Многокритериальная оптимизация. – Л.: ВАС, 1984. – 94 с.

Фисун Александр Павлович

ФГБОУ ВПО «Госунiversитет – УНПК», г. Орел

Доктор технических наук, профессор, профессор кафедры «Электроника, вычислительная техника, информационная безопасность»

Тел.: 8 910 307 00 81

E-mail: fisun01@pisem.net

Фисун Роман Александрович

Отделение по Смоленской области Главного управления Центрального банка Российской Федерации по Центральному федеральному округу, г. Смоленск

Заместитель начальника отдела информационной безопасности

Тел.: 8 (4862) 42-15-56

A.P. FISUN (*Doctor of Engineering Sciences, Professor,
Professor of Department «Electronics, Computer Science, Information Security»
State University – ESPC, Orel*)

R.A. FISUN (*Deputy Head of Department of Information Security)
Department of the Smolensk region of the Main Department of the Central Bank of the Russian Federation in
the Central Federal District, Smolensk*)

METHODS OF ASSESSING THE EFFECTIVENESS OF AN INTEGRATED INFORMATION SECURITY OF INFORMATION SYSTEMS AND INFORMATION AND TELECOMMUNICATION NETWORKS OBJECTS OF INFORMATIZATION OF PUBLIC AUTHORITIES

The article outlines a method of assessing the effectiveness of information security of information systems (is) and information and telecommunications networks (ITN) formation of objects of Informatization of the Federal bodies of state power, bodies of state power of subjects of the Russian Federation, characterized by considerable diversity and

variety of quantitative, qualitative indicators and criteria of efficiency of these technologies and operating in difficult conditions of uncertainty, risk, exposure to external and internal threats.

Keywords: information systems (IS); information and telecommunications network (ITN) formation; information and telecommunication technologies (ICT); objects of informatization; information security IS; ITN; ICT; comprehensive information security and information protection (SCOB); performance evaluation SCOB.

BIBLIOGRAPHY (TRANSLITERATED)

1. Fisun A.P., Eremenko V.T. Ocenka bezopasnosti informacii, obrabaty'vaemoj kompleksami texnicheskix sredstv sistem upravleniya special'noj svyazi // Sbornik nauchny'x trudov Akademii MVD Rossii. – M.: Akademiya MVD RF, 1994.
2. Fisun A.P. Modelirovanie i ocenka ugroz informacionnoj bezopasnosti // Sbornik materialov 8-j Mezhdunarodnoj konferencii «Informatizaciya pravooxranitel'ny'x sistem» (MFI-99). – Chast' 2. – M.: Akademiya MVD RF, 1999.
3. Fisun A.P., Lebedev V.N. Opredelenie i klassifikaciya ugroz bezopasnosti informacionno-vy'chislitel'ny'x setej // Trudy' Akademii upravleniya MVD Rossii «Komp'yuterny'e texnologii i upravlenie OVD». – M.: Akademiya upravleniya MVD RF, 2000.– S. 67-78.
4. Fisun A.P., Minaev V.A. i dr. Teoreticheskie osnovy' informatiki i informacionnaya bezopasnost': monografiya. – M.: Radio i svyaz', 2000. – 466 s.
5. Fisun R.A. Modelirovanie ugroz informacionnoj bezopasnosti informacionno-telekommunikacionny'x sistem: diplomny'j proekt. – Orel, Akademiya FAPSI, 2003.
6. Fisun A.P. i dr. Modelirovanie ugroz informacionnoj bezopasnosti cheloveko-komp'yuterny'x sistem / A.P. Fisun, K.A. Dzhevaga, R.A. Fisun, L.A. Grashhenko // Sbornik materialov 2-j Vserossijskoj NPK «Metody' i sredstva texnicheskoj zashhity' konfidencial'noj informacii», 7-9 iyunya 2005 goda. – Obninsk, GOU GCIPK, 2005.
7. Terent'ev V.M. Metodika obosnovaniya trebovanij k pokazatelyam kachestva ASMKRS. – L.: VAS, 1991.
8. Antushev G.S. Metody' parametriceskogo sinteza slozhny'x texnicheskix sistem. – M.: Nauka, 1989. – 88 s.
9. Feller V. Vvedenie v teoriyu veroyatnostej i ee prilozheniya: per. s angl. – V 2-x tomax. – T. 1. – M.: Mir, 1984. – 528 s.
10. Levin B.R. Veroyatnostny'e modeli i metody' v sistemax svyazi i upravleniya.– M.: Radio i svyaz', 1985. – 312 s.
11. Vencel' E.S. Teoriya veroyatnostej. – M.: VSh, 1998. – 576 s.
12. Lanne' A.A., Ulaxovich D.A. Mnogokriterial'naya optimizaciya. – L.: VAS, 1984. – 94 s.

ТРЕБОВАНИЯ
к оформлению статьи для опубликования в журнале
«Информационные системы и технологии»

ОБЩИЕ ТРЕБОВАНИЯ

Объем материала, предлагаемого к публикации, измеряется страницами текста на листах формата А4 и содержит от 4 до 9 страниц; все страницы рукописи должны иметь сплошную нумерацию.

В одном сборнике может быть опубликована только одна статья одного автора, включая соавторство.

Плата с аспирантов за публикацию рукописей не взимается.

Аннотации всех публикуемых материалов, ключевые слова, информация об авторах, списки литературы будут находиться в свободном доступе на сайте соответствующего журнала и на сайте Российской научной электронной библиотеки – РУНЭБ (Российский индекс научного цитирования).

ТРЕБОВАНИЯ К СОДЕРЖАНИЮ НАУЧНОЙ СТАТЬИ

Научная статья, предоставляемая в журналы, должна иметь следующие **обязательные** элементы:

- постановка проблемы или задачи в общем виде;
- анализ достижений и публикаций, в которых предлагается решение данной проблемы или задачи, на которые опирается автор, выделение научной новизны;
- исследовательская часть;
- обоснование полученных результатов;
- выводы по данному исследованию и перспективы дальнейшего развития данного направления;
- библиография.

ТРЕБОВАНИЯ К ОФОРМЛЕНИЮ НАУЧНОЙ СТАТЬИ

Статья должна быть набрана шрифтом Times New Roman, размер 12 pt с одинарным интервалом, текст выравнивается по ширине; абзацный отступ – 1,25 см, правое поле – 2 см, левое поле – 2 см, поля внизу и сверху – 2 см.

Обязательные элементы:

- УДК
- заглавие (на русском и английском языках)
- аннотация (на русском и английском языках)
- ключевые слова (на русском и английском языках)
- список литературы, на которую автор ссылается в тексте статьи.

ТАБЛИЦЫ, РИСУНКИ, ФОРМУЛЫ

Все таблицы, рисунки и основные формулы, приведенные в тексте статьи, должны быть пронумерованы.

Формулы следует набирать в редакторе формул Microsoft Equation 3.0 с размерами: обычный шрифт – 12 pt, крупный индекс – 10 pt, мелкий индекс – 8 pt. **Формулы, внедренные как изображение, не допускаются!** Русские и греческие буквы, а также обозначения тригонометрических функций набираются прямым шрифтом, латинские буквы – *курсивом*.

Рисунки и другие иллюстрации (чертежи, графики, схемы, диаграммы, фотоснимки) следует располагать непосредственно после текста, в котором они упоминаются впервые. Рисунки, число которых должно быть логически оправданным, представляются в виде отдельных файлов в формате *.eps (Encapsulated PostScript) или TIF размером не менее 300 dpi.

СВЕДЕНИЯ ОБ АВТОРАХ

В конце статьи приводятся набранные 10 pt сведения об авторах в такой последовательности: фамилия, имя, отчество (полуужирный шрифт); учреждение или организация, ученая степень, ученое звание, должность, адрес, телефон, электронная почта (обычный шрифт). Сведения об авторах также предоставляются отдельным файлом и обязательно дублируются на английском языке.