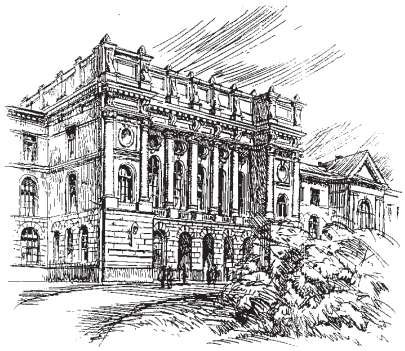
5(86)/2009



Научно-технические ведомости СПбГПУ

ИНФОРМАТИКА. ТЕЛЕКОММУНИКАЦИИ. УПРАВЛЕНИЕ

Санкт-Петербург. Издательство Политехнического университета

Федеральное агентство по образованию
Санкт-Петербургский государственный политехнический университет

НАУЧНО-ТЕХНИЧЕСКИЕ ВЕДОМОСТИ СПбГПУ

РЕДАКЦИОННЫЙ СОВЕТ ЖУРНАЛА

Алферов Ж.И., академик РАН; *Васильев Ю.С.*, академик РАН (председатель);
Костюк В.В., академик РАН; *Лопота В.А.*, чл.-кор. РАН;
Окрепилов В.В., чл.-кор. РАН; *Рудской А.И.*, чл.-кор. РАН;
Терещенко Г.Ф., академик РАН; *Федоров М.П.*, чл.-кор. РАН;
Фортов В.Е., академик РАН.

РЕДАКЦИОННАЯ КОЛЛЕГИЯ ЖУРНАЛА

Васильев Ю.С., академик РАН (главный редактор); *Арсеньев Д.Г.*, д-р техн. наук, профессор;
Бабкин А.В., д-р экон. наук, профессор (зам. гл. редактора);
Боронин В.Н., д-р техн. наук, профессор; *Глухов В.В.*, д-р экон. наук, профессор;
Дегтярева Р.В., д-р ист. наук, профессор; *Иванов А.В.*, д-р техн. наук, профессор;
Иванов В.К., д-р физ.-мат. наук, профессор; *Козловский В.В.*, д-р физ.-мат. наук, профессор;
Рудской А.И., чл.-кор. РАН (зам. гл. редактора); *Юсупов Р.М.*, чл.-кор. РАН.

СЕРИЯ "ИНФОРМАТИКА. ТЕЛЕКОММУНИКАЦИИ. УПРАВЛЕНИЕ"

РЕДАКЦИОННЫЙ СОВЕТ СЕРИИ

Юсупов Р.М., чл.-кор. РАН – председатель;
Абрамов С.М., чл.-кор. РАН;
Воеводин В.В., чл.-кор. РАН;
Заборовский В.С., д-р техн. наук, профессор;
Козлов В.Н., д-р техн. наук, профессор;
Фотиади А.Э., д-р физ.-мат. наук, профессор;
Черноруцкий И.Г., д-р техн. наук, профессор.

РЕДАКЦИОННАЯ КОЛЛЕГИЯ СЕРИИ

Юсупов Р.М., чл.-кор. РАН – председатель;
Арсеньев Д.Г., д-р техн. наук, профессор – зам. председателя;
Бабкин А.В., д-р экон. наук, профессор – зам. председателя;
Болдырев Ю.Я., д-р техн. наук, профессор;
Давыденко М.О., технический секретарь;
Карпов Ю.Г., д-р техн. наук, профессор;
Коротков А.С., д-р техн. наук, профессор;
Макаров С.Б., д-р техн. наук, профессор;
Устинов С.М., д-р техн. наук, профессор;
Цикин И.А., д-р техн. наук, профессор;
Шабров Н.Н., д-р техн. наук, профессор.

Журнал выходит под научно-методическим руководством Российской академии наук с 1995 года.

Журнал включен в Перечень ведущих рецензируемых научных журналов и изданий, в которых должны быть опубликованы основные результаты диссертаций на соискание ученой степени доктора и кандидата наук.

Журнал издается в четырех сериях:
Наука и образование;
Физико-математические науки;
Экономические науки;
Информатика, телекоммуникации, управление.

Журнал зарегистрирован в Госкомпечати РФ. Свидетельство № 013165 от 23.12.94.

Подписной индекс **18390** в каталоге "Газеты. Журналы" Агентства "Роспечать".

Журнал включен в базу данных "Российский индекс научного цитирования" (РИНЦ), размещенную на платформе Национальной электронной библиотеки на сайте http://www.elibraru.ru.

При перепечатке материалов ссылка на журнал обязательна.

Точка зрения редакции может не совпадать с мнением авторов статей.

Адрес редакции и издательства: Россия, 195251, Санкт-Петербург, ул. Политехническая, д. 29.
Тел. редакции серии (812) 552-62-16.

© Санкт-Петербургский государственный политехнический университет, 2009

Содержание

Информатика и информационные технологии

А.Д. Тазетдинов. *Об использовании графов понятий для структурирования и управления информацией в автоматизированных обучающих системах* 7

А.К. Соколовский. *Оценка привлекательности проектов в сфере информационных технологий* 13

А.А. Сухов. *Проблема человеко-машинного интерфейса в информационных системах* 18

Л.А. Хворова, В.М. Брыксин, Н.В. Гавриловская. *Использование информационных технологий при прогнозировании урожайности зерновых культур* 23

Телекоммуникационные системы и компьютерные сети

Ю.Я. Болдырев, В.А. Варгаузин, Е.П. Петухов, И.А. Цикин. *Технология сетевого доступа к программной среде MATLAB на основе кластерной вычислительной системы* 31

А.Н. Фирсов. *Оценка надежности асинхронных распределенных систем, устойчивых к произвольным отказам* 38

Проблемы передачи и обработки информации

Ю.Е. Сидоров, А.В. Шумилов. *Исследование помехоустойчивости рангового обнаружителя сигналов при априорной неопределенности* 45

В.А. Грудинин, Т. П. Белякова. *Разработка акустического метода оперативной диагностики состояния покрышек грузовых автомобилей* 50

Б.Б. Казаков, С.А. Ивановский. *Эффективные методы выделения движения объектов в последовательности изображений* 55

А.В. Томилин, М.С. Стуков, А.А. Стенин. *Применение диверсификации в задачах динамической защиты памяти* 60

В.С. Тутыгин, С.В. Шедов. *Новые алгоритмы обнаружения и определения параметров зашумленных сигналов* 64

Радиотехника, антенны, СВЧ-устройства

А.В. Ненашев, В.А. Охотников. *Разработка метода анализа и моделирования нелинейных радиотехнических устройств* 73

В.А. Карелин, И.В. Карелин. *Получение максимально правдоподобных оценок малых отклонений частот высокостабильных гармонических сигналов* 79

В.Н. Коновалов, И.В. Коновалов. *Подход к многоуровневой оптимизации комбинационных логических схем* 83

Ю.В. Серединский, В.П. Горбунов, Ю.Н. Новиков. *Применение ферритов для ослабления влияния кабелей на характеристики излучения антенных систем* 88

В.А. Сороцкий, М.А. Уткин. *Влияние электромагнитной связи антенн радиопередающих устройств на процессы в ключевых генераторах* 93

В.А. Сороцкий, М.А. Уткин. *Исследование характеристик многоячейковых тиристорных генераторов на цифровой модели* 99

Системный анализ и управление

Е.А. Макарова. *Управление поведением макроэкономических агентов на основе механизма самовыравнивания финансовых потоков* 105

В.Л. Токарев, С.В. Орлов. *Задача автоматизации процесса формирования лингвистических моделей для поддержки принятия решений* 111

П.Н. Победаш. *Анализ двухкритериальной модели оптимального управления реальными инвестициями с неопределенным спросом* 114

Н.А. Попцов, Д.Б. Бекурин, М.А. Марценюк. *Нейросетевое управление локомоциями самообучающейся модели биообъекта* 119

Вычислительные машины и программное обеспечение

М.Ю. Кудрин, Е.В. Соколов, А.Г. Тормасов. *Выявление состояний гонки с помощью графа совместного исполнения потоков* 123

Д.А. Крупенко. *Адаптация системы проектирования AUTOCAD для разработки конструкций микросборок СВЧ* 129

К.С. Солнушкин. *Система учета доступа к лицензионному программному обеспечению в центрах коллективного пользования* 133

А.А. Сухов. *Оценка эмоциональной составляющей интерфейсов тренажеров* 140

Д.Ю. Братин. *Особенности языка FDL v. 1.0 для описания фронтенда веб-приложений* 145

Математическое моделирование: методы, алгоритмы, технологии

Т.М. Гильмуллин, И.В. Аникин. *Моделирование объектов информационной безопасности для задачи оценки рисков* 151

А.С. Зраенко, К.А. Аксенов, Ван Кай. *Коалиционная модель мультиагентного процесса преобразования ресурсов* 156

М.Ю. Арзуманян. *Модели описания предприятий и их мониторируемые свойства* 161

А.П. Тейхриб, Л.Г. Доросинский. *Сравнительный обзор моделей логического времени* 168

И.И. Икрянов. *Модели движения индикаторов курса и глиссады с маятниковой системой стабилизации оптических осей* 174

Электроника, технологии производства материалов электронной техники

В.С. Нагорный. *Базовые рецептуры наноструктурированных серебром рабочих жидкостей для электрокаплеструйных технологий* 179

Информационные и телекоммуникационные технологии в образовании

Ю.А. Ветринский. *Опыт использования технологии смешанного обучения в практике преподавания технических дисциплин* 185

Н.Б. Кунтурова. *Методика формирования структурных групп кандидатов при поступлении в магистратуру с прогнозируемой успешностью их обучения* 190

Н.Б. Кунтурова. *Особенности оценивания профессиональной компетентности специалистов в процессе их обучения в вузе* 194

Сведения об авторах 199

Аннотации 203

Contents

Informatics and information technologies

A.D. Tazetdinov. *About using graphs of concepts for structurization and management information in the automated training systems* 7

A.K. Sokolovsky. *Evaluation of project attractiveness in information technology field* 13

A.A. Suhov. *Problems of human-machine interfaces in information systems* 18

L.A. Khvorova, V.M. Bryksin, N.V. Gavrilovskaya. *Use of information technologies in forecasting grain yields* 23

Telecommunication systems and computer networks

Yu.Ya. Boldyrev, V.A. Vargauzin, E.P. Petukhov, I.A. Tsikin. *Technology for remote network access to matlab software environment hosted on computer cluster system* 31

A.N. Firsov. *Reliablity estimation of asyncronous bysinatine fault-tolerant distributed systems* 38

Problems of transfer and information processing

Yu.E. Sidorov, A.V. Shumilov. *Nonparametric radio signal detector* 45

V.A. Grudinin, T.P. Belyakova. *Active control system of truck tires intactness status* 50

B.B. Kazakov, S.A. Ivanockij. *Effective methods of optical flow based object tracking in video sequences* ... 55

A.V. Tomilin, M.S. Stukov, A.A. Stenin. *Dynamic memory protection by the program diversity* 60

V.S. Tutygin, S.V. Shedov. *New algorithms of detection and parameters definition of noised signals* 64

Radio engineering, aerials, SHF devices

A.V. Nenashev, V.A. Okhotnikov. *Working out the method of the analysis and modeling of nonlinear radio engineering devices* 73

I.V. Karelin, V.A. Karelin. *Optimal estimation of high stability harmonic wave small frequency deflection* ... 79

V.N. Konovalov, I.V. Konovalov. *Approach to multilevel optimization of combinational circuit* 83

Y.V. Seredinskiy, V.P. Gorbunov, Y.N. Novikov. *The use of ferrite materials to decrease cables influence on antenna systems radiation characteristics* 88

V.A. Sorotsky, M.A. Utkin. *An influence of radio transmitters antennas electromagnetic interconnection on the processes in switch generators* 93

V.A. Sorotsky, M.A. Utkin. *Research of multicell thyristor generators characteristics using digital model* ... 99

The system analysis and management

E.A. Makarova. *Control of macroeconomic agent behavior on basis of self-equalizing mechanism of financial flows* 105

V.L. Tokarev, S.V. Orlov. *The problem of automations of process of formation of linguistic model for integrated decision-making support* 111

P.N. Pobedash. *An analysis of two-criteria optimal control problem of real investment under uncertainty by discrete principle of maximum* 114

N.A. Poptsov, D.B. Bekurin, M.A. Martsenyuk. *Neuronet operating by locomotion system of self-training bio-object model moving to target* 119

Computers and the software

M.U. Kudrin, E.V. Sokolov, A.G. Tormasov. *The race condition detection using the graph of joint thread execution* 123

D.A. Krupenko. *Methods of adaptation of system of designing AUTOCAD for working out of the design of microassemblages RF* 129

K.S. Solnushkin. *The system of accounting for access to the licensed software in center for collective use* 133

A.A. Suhov. *The evaluation of the emotional component of the computer simulators* 140

D.Y. Bratin. *Specialities of the FDL v. 1.0 the frontend description language* 145

Mathematical modelling: methods, algorithms, technologies

T.M. Gilmullin, I.V. Anikin. *Information security objects modeling for risk assessment* 151

A.S. Zraenko, K.A. Aksyonov, Van Kai. *Coalition model of multi-agent process of resources transformation* ... 156

M.Y. Arzumanyan. *Enterprise descreptional models and their monitored values* 161

A.P. Teihrib, L.G. Dorosinsky. *Comparative review of models of logical time* 168

I.I. Ikryanov. *Models of movement of indicators of the course and glidepath with pendular system of stabilization of optical axes* 174

Electronics, technologies of producing the materials of electronic technics

V.S. Nagorniy. *Base compoundings nanostructured silver of working liquids for electrical drop-jet technologies* ... 179

Information and telecommunication technologies in education

Yu.A. Vetrinsky. *The experience of blended learning technology use in practice of technical subjects teaching* ... 185

N.B. Kunturova. *Method of shaping of the structured groups of candidates at the entering in magistracy with predictable success of their education* 190

N.B. Kunturova. *Peculiarities of professional competence of specialists in process of education in high school* ... 194

About the authors 199

Abstracts 203

УДК 004.942

А.Д. Таземдинов

ОБ ИСПОЛЬЗОВАНИИ ГРАФОВ ПОНЯТИЙ ДЛЯ СТРУКТУРИРОВАНИЯ И УПРАВЛЕНИЯ ИНФОРМАЦИЕЙ В АВТОМАТИЗИРОВАННЫХ ОБУЧАЮЩИХ СИСТЕМАХ

Интенсификация образования и активное развитие дистанционного и интернет-образования сделали актуальной возможность автоматизированного управления обучающей (входной) информацией. Все большую роль и распространение получают автоматизированные обучающие системы (АОС). Однако в отличие от автоматизированных систем управления неживыми объектами, где точно известны все параметры входной информации и управляющих сигналов (формат, структура, тип, сила и т. д.), в образовательной сфере крайне мало количественно измеряемых параметров входной информации, на основании которых можно строить управляющие воздействия [1, 2]. Недостаточное внимание, уделяемое разработчиками АОС этой проблеме, приводит к неудовлетворенности обучающихся и преподавателей этими системами.

Если уровень знаний в значительной степени зависит от личных усилий и способностей, а также от психофизиологических особенностей личности обучаемых, то структура знаний отражает особенности организации учебного процесса. Чем лучше структурирована информация, предъявляемая на учебном занятии, тем проще она запоминается и дольше сохраняется в памяти. В связи с этим одним из ключевых моментов в создании эффективных механизмов управления АОС становится анализ содержания обучающей информации и ее структурирование. В то же время многие из параметров обучающей информации имеют категориальный (качественный) характер, и требуются новые подходы для нахождения способа их измерения.

Графы понятий

Обычно, в том числе и в АОС, учебный материал курса обучения или дисциплины представляет собой массив информации, включающий в себя набор элементов (хрестоматии, контенты, методические указания и т. д.), который делят линейно на логические и относительно законченные части следующим образом:

1. Раздел.
   1.1. Тема.
      1.1.1. Параграф.

Однако восприятие информации обучающимся осуществляется посредством более мелких элементов – понятий (концептов) [3]. Рассмотрим процесс усвоения знаний с позиции теоретико-множественного подхода. Любая предметная область характеризуется ограниченным количеством описывающих ее понятий. Поэтому возьмем некоторую предметную область, представленную множеством понятий A из k элементов, $A = \{a_1, a_2, ..., a_k\}$, где $k \in N$ множеству натуральных чисел. Множество понятий, используемое преподавателем при изложении учебного материала (УМ), можно определить как подмножество B на множестве понятий A о данной предметной области в целом, такое, что $B \subseteq A$. Количество понятий, знаемых обучаемым в изучаемой предметной области, определяется как подмножество C на множестве понятий A, такое, что $C \subseteq A$. Для преподавателя и обучаемого возможны следующие отношения на множестве A:

$C = B$ – когда преподаватель и обучаемый обладают одинаковым набором понятий о предметной области;

$C \subset B$ – обучаемый знает о предметной области меньше преподавателя, но оперирует общими

с преподавателем понятиями (обратная ситуация не рассматривается);

$C \cap B \neq \varnothing$ – существует некоторое общее совпадающее множество понятий обучаемого и преподавателя о данной предметной области;

$C \cap B = \varnothing$, $C \neq \varnothing$, $B \neq \varnothing$ – обучаемый и преподаватель имеют непересекающиеся знания о данной предметной области;

$C = \varnothing$ – обучаемый ничего не знает о данной предметной области.

Во всех случаях, кроме первого, обучаемому необходимо изучать новые понятия. Следует различать такие результаты обучения, как "знаю" и "понимаю". Если "знаю" – это могу повторить или узнать, то "понимаю" следует рассматривать как возможность оперировать изученными понятиями в нестандартных ситуациях. Обычно преподаватель, объясняя новое понятие, старается использовать уже известные обучаемым понятия. Если преподаватель выясняет, что его объяснение не понято, он может, в отличие от АОС, переключаться на более простой язык. И так до тех пор, пока результат не будет достигнут. Ключ к пониманию обучаемым УМ лежит в умении преподавателя найти то понятийное множество A, на котором $C = B$. В этом случае преподаватель гарантированно может объяснить значение нового понятия.

Пусть у нас есть множество понятий A из k элементов. Преподаватель пытается объяснить их обучающемуся посредством множества понятий B из r элементов. Обучающийся воспринимает их посредством множества C из q понятий. Тогда любому $a_i \in A$, $i = \{1, 2, ..., k\}$ понятию можно сопоставить подмножество $D_i \subseteq B$, $D_i \neq \varnothing$ понятий, объединенных смысловыми связями. Образовавшуюся структуру математически и визуально можно описать с помощью графа G_i. Обучающийся любое понятие $d_j \in D_i$ воспринимает и понимает посредством подмножества собственных понятий $E_{ij} \subseteq C$, объединенных смысловыми связями в граф G'_{ij}. При этом подмножество E_{ij} в отличие от D_i может быть равно $\varnothing$. Тогда обучающемуся требуются дополнительные разъяснения, так как понятие a_i понято им не полностью, либо не понято совсем, либо понято неправильно. Третий вариант наиболее сложный. Именно он приводит к возникновению иллюзии полного понимания УМ, в то время как это понимание искажено, иногда существенно, недостаточной информацией или неверными посылками [4]. Таким образом, входная информация воспринимается обучающимся в виде графа понятий, представляющего собой многоуровневую структуру с нелинейной топологией, которая может совпадать или не совпадать с делением на разделы, темы, параграфы.

Каждая вершина такого графа содержит одно понятие (семантическую единицу информации), которое в свою очередь может быть также представлено в виде графа, имеющего свои собственные исходные вершины (свои семантические единицы информации, изученные на предыдущих этапах обучения). Под семантической единицей информации понимается сложное или простое понятие, а также конкретные формулы, теоремы, определения, аксиомы, леммы, следствия, законы, правила, события и факты, рассматриваемые в контексте УМ [5, с. 70]. Каждая дуга графа есть не что иное, как символ отношения (связи) между понятиями, которые она соединяет. Связи между понятиями могут быть не только прямыми, но и транзитными (иерархическими, косвенными), когда путь между двумя вершинами (понятиями) содержит больше одной дуги. Далее, поскольку речь идет о понятиях учебного материала и их связях, а графовая модель является лишь математическим аппаратом и средством визуализации этих связей, для обозначения дуг и путей на графе понятий будут использоваться термины "связь" и "длина связи" (минимальная длина соответствует одной дуге графа).

Идея представления понятий (концептов) УМ и их связей в виде графа не нова. Ей посвящены работы А.И. Уемова [6], В.М. Мизенцева, А.В. Кочергина [7], Л.П. Леонтьева, О.Г. Гохмана [5], В.Б. Швыркова [8], И.О. Александрова [9], Н.М. Леоновой [10] и многих других ученых. Основная задача, решаемая в их работах, – определить оптимальный объем учебного материала, выдаваемого за один раз, так как чем больше семантического материала содержится в порции УМ, тем труднее он усваивается. В педагогической литературе известны две методики расчета семантического веса УМ: В.М. Мизенцева [7] и Л.П. Леонтьева [5]. В основе обеих методик лежат сформулированные в предметной области теории систем идеи А.И. Уемова, в соответствии с которыми информационная мера сложности графовой модели определяется длиной (количеством) дуг графа. Последние находятся как от-

ношения между его элементами (вершинами) и конфигурацией графовой модели, определяемой показателем относительной энтропии как мерой неопределенности системы. При этом управление тем лучше, чем больше мы можем выдать информации за стандартное время или чем меньше мы потратим времени на выдачу стандартного объема информации $u(\Delta t) \to \max$ или $u(\Delta x) \to \min$.

Единица времени Δt и единица информации Δx – это единственные количественные величины (или параметры), которые могут быть реально измерены в рамках АОС и которые могут служить и служат формальными показателями эффективности функции управления любой АОС. Подобное ограничение в реальной жизни встречается редко. Обычно таких параметров гораздо больше, и они оказывают существенное влияние на процесс понимания и запоминания информации.

В качестве параметров обучающей информации можно рассматривать: объем; структурированность; связность смысловых элементов; язык изложения (насыщенность специальными терминами, стиль и т. п.); степень новизны; степень сложности материала (использование сложных логических конструкций и выводов и т. п.); форму представления (видео, аудио, графическая, текстовая и т. д.); последовательность подачи (изложения или изучения).

Современные исследования в области психолингвистики и когнитивной психологии, посвященные проблеме понимания [3, 8, 9], свидетельствуют о важности языковых конструкций, способа изложения, связности смысловых элементов, структурированности и форме представления новой информации для ее понимания и запоминания человеком. Поэтому информация, организованная в виде графа понятий, обеспечивает более целостное представление об УМ, характеризующееся большей связанностью и структурированностью материала (от общих понятий до частных фактов) и меньшим объемом сопутствующей второстепенной информации [11].

Отсутствие учета многих параметров обучающей информации связано не только с ограниченными возможностями автоматизированных систем по формированию нестандартных творческих реакций на ответы обучающихся. АОС не могут произвольно менять форму представления УМ, эмоциональную окраску текста, структуру смысловых связей или множества понятий языка изложения. Пока на это способен только человек. В первую очередь это связано со сложностью их количественного измерения. Поэтому введем некоторые дополнительные по отношению к объему и времени величины обучающей информации, которые могут быть количественно измерены при использовании графа понятий для структурирования и представления УМ.

Способы измерения параметров

Язык изложения – один из наиболее важных параметров обучающей информации, ответственных за понимание УМ. Для измерения сложности языка изложения разделим граф УМ на подграфы по уровням иерархии, каждый из которых объединит понятийные множества, необходимые для объяснения понятий вышележащего уровня. Графы первого уровня формируются из понятий множества A; второго уровня – из понятий множеств B и C; третьего уровня – из понятий множества B_2, необходимого для объяснения понятий множеств B и C; графы j-го уровня – из понятий множества B_j, необходимого для объяснения понятий множества B_{j-1}. В этом случае степень сложности языка изложения будет измеряться в количестве уровней графа, где нулевому уровню сложности соответствуют графы второго уровня. То есть чем сложнее язык изложения, тем больше шагов рекурсивного спуска требуется для объяснения предыдущих понятий. Введение этого признака позволяет выполнить анализ и итерационную оптимизацию графа понятий УМ, заключающуюся в замене множества B для тех понятий a_i, у которых $j(a_i) > 0$ так, чтобы $j \to 0$.

Связность смысловых элементов определяется как отношение количества элементов множества A к количеству дуг графа или графов первого уровня:

$$s = n(k-1)/\sum_{i=1}^{n} m_i, \qquad (1)$$

где n – количество графов первого уровня; k – количество элементов множества A; $m_i = (V_i = (v_1, v_2, ..., v_m))$ – число дуг i-го графа.

Для минимального графа УМ, который состоит из графов первого и второго уровня, связность смысловых элементов рассчитывается по формуле

$$s = \frac{n(k-1) + \sum_{i=1}^{k} h(z-1)}{\sum_{i=1}^{n} m_i + \sum_{i=1}^{k}\sum_{g=1}^{h} m_g}, \quad (2)$$

где h – количество графов второго уровня, разъясняющих a_i понятие; $m_g = (V_g = (v_1, v_2, ..., v_g))$ – число дуг g-го графа.

Чем больше значение s, тем больше связанность $s \to \max$. Значение $s < 1$ говорит о том, что материал слабо связан и требует реструктуризации либо разделения на части.

Структурированность. Согласно коннекционистской теории [3] каждый путь на графе понятий УМ можно свернуть и представить в качестве единицы информации. Тогда максимальное теоретическое количество вновь образованных единиц информации при m вершинах вычисляется как сумма от одного до n размещений

$$k = \sum_{n=1}^{m} A_m^n. \quad (3)$$

То есть на четырех вершинах (понятиях) мы получаем 40 вновь образованных единиц информации. Безусловно, четыре понятия за один раз запомнить несложно, а 40 практически невозможно. Однако такая связанность может отражать реальную систему взаимоотношений понятий предметной области. Следовательно, чтобы избежать снижения качества запоминания УМ, требуется специальная технология предъявления такого графа обучающемуся, основанная на использовании структурной основы учебного материала.

Очевидно, что для вычисления степени структурированности графа необходимо выявить на множестве связей смысловую структуру учебного материала, условно разделив связи на основные и дополнительные. Возникает вопрос о топологии этой структуры. Определив в качестве критерия оценки топологии степень связанности графа, можно построить шкалу, где с одной стороны располагается несвязный граф (фактически тезаурус), а с другой – сильно связный граф. Среднее значение шкалы занимает топология типа дерево. Дерево, реализуя естественную иерархию связей понятий предметной области, позволяет обеспечить взаимодействие всех вершин графа понятий при минимальном количестве дуг, равном $n-1$, при n вершинах. Независимо от алгоритма обхода дерева УМ промежуточные вершины повторяются тем большее количество раз, чем ближе они расположены к корню. Таким образом, более общие понятия и структура учебного материала лучше понимаются и запоминаются.

Для множества понятий A определяются связи (дуги графа) первого и второго уровня. Связи первого уровня формируют иерархическую структуру, идеальной топологией которой является дерево. Связи второго уровня обеспечивают дополнительную естественно-смысловую связанность понятий и условно считаются второстепенными. Величина показателя вычисляется из связей первого уровня как показатель сформированности дерева (отсутствие циклов и несвязных частей) в процентном отношении (оптимальность графа по данному показателю $str \to 100\ \%$):

$$str = \begin{cases} (|k-1-m| \cdot 100)/(k-1), \text{если } |k-1-m| < k; \\ 0, \end{cases} \quad (4)$$

где m – число дуг графа.

Важность цельности (или сформированности) дерева можно увидеть при анализе потери одного или нескольких понятий (вершин) из дерева УМ. Потеря одной вершины не приводит к уменьшению количества цепочек понятий, но существенно сокращает количество связей. Количество оставшихся связей можно вычислить по формуле

$$L_{\text{УМ}} = \sum_{i=1}^{k} C_{n_i}^{m}, \quad (5)$$

где k – количество образовавшихся деревьев; n_i – количество вершин в i-м дереве; $m = 2$.

Последовательность изложения представляет собой путь на графе УМ, включающий по возможности все вершины графа. Правильная последовательность изложения проявляется в отсутствии объяснений неизученных понятий неизученными. (То есть отсутствие неразрешимых петель на графе УМ, когда для понимания одного нового понятия требуется изучить второе понятие.) При этом для понимания второго понятия требуется знание и понимание первого. Идеальная последовательность – это когда все понятия множества A объяснены, но ни одно из понятий не объяснялось два и более раз (т. е. про-

стой путь). Обычно в излагаемом преподавателем материале петель бывает немного, поэтому в качестве единицы измерения этого параметра предлагается использовать количество петель. Оптимизация графа УМ будет заключаться в исключении петель из последовательности изложения $loop \to 0$.

Однако количество вершин (понятий) k' множества понятий A', входящих в путь (последовательность изложения) может не совпадать с количеством понятий множества A $k' \neq k$. Этому может быть несколько причин. Первая: анализ знаний обучающихся (например, входной тест) показал, что часть понятий уже знакома обучающемуся и дополнительное повторение не требуется. Вторая – это наличие петель в последовательность изложения. Третья – комбинация первой и второй. Предлагается следующий способ вычисления количества петель:

$$loop = k - k'', \qquad (6)$$

где k'' – количество понятий множества понятий A'', являющегося пересечением множеств A и A', $A'' = A \cap A'$, $A'' \subseteq A$, $k'' \leq k$.

Такие параметры обучающей информации, как форма представления, степень новизны и сложности УМ обычно формируются эмпирически, исходя из представлений и опыта преподавателя, и зависят от решаемой задачи. В АОС закладываются изначально и редко изменяются. Следует различать сложность обучающей информации (использование сложных логических конструкций и выводов и т. п.) и сложность практических заданий и задач. Управление последними активно используется в адаптивных АОС, где группы сложности заданий сопоставляются с уровнем знаний обучающихся, показанным в определенных контрольных точках учебного курса [10].

Следующим важным механизмом повышения качества запоминания УМ является визуализация изучаемого графа понятий с выделением текущего понятия и всей цепочки связи этого понятия с корневой вершиной. Визуализация позволяет дать целостный системный взгляд на изучаемую предметную область и решить ряд важных дидактических задач.

Предлагается следующий способ визуализации изучаемых элементов.

- Визуализируются следующие структуры:
  граф понятий G;
  все подграфы G_i и G_{ij} (слои) графа G, отсортированные в порядке предъявления (изложения) УМ;
  текущий подграф G_{ij};
  связку графов $G \to G_i \to G_{ij}$.
- Изученные понятия отличаются цветом от неизученных.
- Каждое изученное понятие может быть окрашено в цвет, соответствующий уровню понимания этого понятия. Например: "хорошо" – синий; "удовлетворительно" – желтый; "плохо" – красный; "отлично" – серый. Глядя на граф, нужно в первую очередь видеть то, что требует дополнительной проработки, поэтому яркими цветами выделяются плохие оценки, а отличные не выделяются.
- Неизученные понятия окрашены в черный цвет (шрифт и обрамление).
- Текущая цепочка понятий выделяется жирным шрифтом и обрамлением, а текущее понятие – зеленым цветом.
- Преподаватель должен иметь возможность просмотра результатов изучения УМ каждого из учеников. Результаты должны быть показаны по желанию преподавателя в сокращенном виде (только в виде графа) или в полном (со всем набором визуализированных подграфов). И в полном, и в сокращенном виде используются все соответствующие цветовые окраски, предоставляя преподавателю информацию об уровне знаний каждого из обучаемых.
- Визуализация может служить не только средством получения информации и обратной связи, но и средством доступа (портального в случае web-реализации) к обучающим заданиям. Задания на странице могут быть сгруппированы по уровням сложности.
- При накоплении информации можно заметить, что одни и те же понятия используются разными УМ или дисциплинами. Происходит частичное пересечение графов. Не являясь средством организации междисциплинарных связей, предлагаемая технология тем не менее может служить средством визуализации этих связей.

Визуализация позволяет решить ряд дидактических задач. Наиболее важная из них – целостное восприятие учебного материала [12]. Обучаемый воспринимает УМ не как множество понятий, а как объект, постоянно видя, на каком этапе процесса обучения он находится, как

текущая изучаемая часть связана с другими частями УМ и что еще необходимо сделать. Точно так же преподаватель, подходя к анализу знаний обучаемого, видит целостную, объемную картину этих знаний, представленную визуализированным графом. В случае использования визуализации преподавателю нет необходимости проводить регулярные тестирования. Информация накапливается, и достаточно беглого взгляда на граф обучаемого для оценки ситуации. Кроме того, визуализация решает проблему портального доступа к web-контенту учебного материала.

Смысловое содержание учебного материала представляет собой целостное единство теоретической и фактической информации, а понимание отдельных слов непосредственно связано с правильным пониманием смысла учебного материала. Проблемы автоматизации управления в АОС заключаются в том, что отнесение обучающегося к слабому уровню обусловлено не только уровнем незнаний или плохими способностями обучающегося, но, прежде всего, несовпадением языков взаимодействия АОС или преподавателя с обучающимся. Результаты многочисленных исследований говорят о том, что понимание учебного материала – важнейший фактор, влияющий как на скорость запоминания, так и на длительность хранения информации в памяти. Поэтому одним из ключевых моментов повышения качества обучения становится анализ содержания обучающей информации и ее структурирование.

Отсутствие учета многих параметров обучающей информации в первую очередь связано со сложностью их количественного измерения. При стандартном делении учебного материала на разделы, темы, параграфы многие из параметров имеют категориальный (качественный) характер, поэтому введение дополнительных измеряемых параметров в случае использования графов понятий даст возможность более точно определять характеристики обучающей информации и тем самым более эффективно управлять процессом обучения. Кроме того, использование графов понятий для структурирования обучающей информации позволяет решить одну из важных дидактических задач – задачу целостного восприятия учебного материала. А визуальное отображение информации о траектории и результатах изучения УМ по каждому обучаемому в виде расцвеченного графа позволяет оказывать преподавателю существенную помощь в принятии оперативных решений по управлению процессом обучения.

СПИСОК ЛИТЕРАТУРЫ

1. **Тазетдинов А.Д.** Анализ математических моделей обучения в приложении к компьютерным обучающим системам репетиторского типа // Научно-технические ведомости СПбГПУ. 2008. № 3(60). С. 191–196.

2. **Тазетдинов А.Д.** Математическая модель усвоения знаний для компьютерных систем репетиторского типа // Научно-технические ведомости СПбГПУ. 2008. № 4(62). С. 134–140.

3. **Солсо Р.** Когнитивная психология. 6-е изд. СПб.: Питер, 2006. 589 с.

4. **Неволин И.Ф., Позина М.Б.** Процессы понимания и когнитивной самооценки в тестовых технологиях // www.nesterova.ru/nauch/testing.pdf.

5. **Леонтьев А.П., Гохман О.Г.** Проблемы управления учебным процессом (математические модели). Рига: Зинатне, 1984. 239 с.

6. **Уемов А.И.** Системный подход и общая теория систем. М.: Мысль, 1978. 272 с.

7. **Мизенцев В.П., Кочергин А.В.** Проблема аналитической оценки качества и эффективности учебного процесса в школе. Куйбышев: Изд-во Куйб. гос. пед. ин-та, 1986. 132 с.

8. **Швырков В.Б.** Введение в объективную психологию: Нейрональные основы психики: Избр. тр. М.: Изд-во ИП РАН, 2006. 592 с.

9. **Александров И.О.** Формирование структуры индивидуального знания. М.: Изд-во Ин-та психологии РАН, 2006. 560 с.

10. **Леонова Н.М., Марковский М.В.** Имитационные математические модели процессов адаптивного управления образовательной деятельностью: Монография / Под ред. А.Д. Модяева. М.: Изд-во МИФИ, 2006. 123 с.

11. **Тазетдинов А.Д.** Технология структурирования и визуализации учебной информации в репетиторских системах // Информационно-управляющие системы. 2009. № 1(38). С. 60–65.

12. **Тазетдинов А.Д.** Технология адаптации в компьютерных обучающих системах репетиторского типа // Экономика и управление. 2008. № 3. С. 215–219.

УДК 65.0+006

А.К. Соколовский

ОЦЕНКА ПРИВЛЕКАТЕЛЬНОСТИ ПРОЕКТОВ В СФЕРЕ ИНФОРМАЦИОННЫХ ТЕХНОЛОГИЙ

В современных экономических условиях большинство организаций, от государства и до коммерческих предприятий, постоянно решают важные вопросы о своем дальнейшем развитии и приобретении конкурентных преимуществ. Для успешного развития и функционирования организациям необходимо строить адекватную стратегию и выбирать механизмы для ее реализации. Хорошо известен принцип: "Невозможно управлять тем, что нельзя измерить". Поэтому организация, которая хочет не только выжить, но и преуспеть в информационную эпоху, нуждается в высокоэффективной системе по оценке привлекательности тех или иных возможностей развития.

Информационный бум последних десятилетий двадцатого века привел к тому, что многие фундаментальные постулаты промышленной конкуренции устарели [1].

В 1990 году Институт Нолана Нортона (Nolan Norton Institute) – научное подразделение KPMG (Kleiveld Peat Marwick Goerdeler) – предложило провести исследование по теме "Показатели деятельности организации будущего". Необходимость такого исследования назрела, ибо существующие подходы к оценке деятельности предприятия, опирающиеся в первую очередь на финансовые показатели, неизбежно устаревали и, как предполагали участники исследования, тормозили развитие организаций [1].

К концу двадцатого столетия система финансовых критериев оценки деятельности компании достигла высокого развития. Однако многие авторы подвергли критике широкое, если не сказать исключительное, использование финансовых показателей в бизнесе [2]. По сути, чрезмерный акцент на достижении и сохранении краткосрочных финансовых результатов может привести к излишним инвестициям в краткосрочные проекты и недостаточным инвестициям в долгосрочные, связанные с информационными технологиями (далее ИТ проект). В первую очередь речь идет о вероятности неверной оценки нематериальных интеллектуальных активов, которые и служат предпосылками будущего роста организации.

Можно привести широко известный пример из истории всемирно известного гиганта IBM, связанный с неверной оценкой перспективности ИТ проекта по производству операционных систем для персональных компьютеров, на тот момент весьма редких среди простых потребителей – физических лиц. Основываясь на текущих маркетинговых исследования и финансовых оценках, корпорация IBM сочла нецелесообразным разрабатывать собственную операционную систему. ИТ проект заключался в развитии операционной системы "Microsoft DOS" на базе персональных компьютеров, выпускаемых корпорацией IBM. Согласно контракту, заключенному в восьмидесятых годах между IBM и малоизвестной компанией Microsoft, корпорация IBM взяла на себя расходы на производство системы и пообещала продавать свои персональные компьютеры только с операционной системой MS DOS, отчисляя при этом Microsoft проценты с каждой проданной машины. Наверное, не стоит говорить о последствиях, которые произошли по результатам выполнения контракта.

Широко распространенные методы финансовой оценки ИТ проектов поддерживают в основном малорискованные проекты, для которых можно определить положительные финансовые показатели, такие, как чистый приведенный доход (NPV), индекс рентабельности инвестиций (PI), срок окупаемости инвестиций (PP), внутренняя норма прибыли (IRR) и др.

Вложения в проекты с длительным сроком возврата инвестиций отвергаются, что в корне неверно для ИТ проектов, которые в подавляющем большинстве приносят нематериальные выгоды; выгоды от знаний, приобретенных за счет новой информационной технологии, могут быть получены в далеком будущем.

Методы по оценке привлекательности ИТ проектов в условиях неопределенности становятся

еще более актуальной проблемой. Это связано с существенным развитием информационных технологий в целом, увеличением масштаба ИТ проектов, их значимости не только для отдельных организаций и граждан, но и для жизни общества в целом. Разработанность проблемы по комплексной оценке привлекательности ИТ проектов низка. Нет системных подходов, учитывающих комплексные и максимально полные данные о субъективных и объективных факторах модели оценки ИТ проекта с ориентацией на требования (и скрытые потребности) заинтересованных сторон. Учет этих факторов существенно влияет на степень точности конечных результатов процесса оценки привлекательности ИТ проекта, тех, что определяют выбор ИТ проекта для реализации.

Проведение оценок социотехнических систем [3] – сложная комплексная задача, которая подчас возникает на этапе, когда решается вопрос о реализации того или иного ИТ проекта. Для адекватного решения требуется применение специальных методов и инструментальных средств, ориентированных как на социальный, так и на технический аспект результатов реализации ИТ проекта.

В первую очередь необходимо выделить цели каждой заинтересованной стороны (табл. 1), которые необходимо достичь при помощи реализации ИТ проекта.

Во вторую очередь определить применимые технические подходы к оценке привлекательности ИТ проекта; они должны выбираться исходя из значимости заинтересованных сторон (ЗС) при принятии решения о реализации.

В третью очередь согласовать иерархический комплекс показателей, которые должны быть взаимосвязаны, измерены и приведены к интег-

Таблица 1

Возможные цели и их противоречия для основных заинтересованных сторон ИТ проекта

Заинтересованная сторона	Характеристика целей ЗС	Противоречия интересов ЗС
Организация* инвестора	Как правило, основной выгодоприобретатель от инвестиций в ИТ проект	Возможно противоречие между целями проекта и долгосрочными стратегическими целями этой организации. Наиболее часто инвестор заинтересован только в финансовой эффективности ИТ проекта
Организация функционального заказчика	Работают с результатами ИТ проекта для приращения ценности производимого продукта или услуги	Поддержат тот ИТ проект, который наилучшим образом удовлетворит их функциональные требования, невзирая на цену. Заинтересованы в наилучшей функциональности результатов проекта
Команда проекта, "внедренцы"	Люди, работающие над успешным внедрением информационных технологий	Ориентируются на краткосрочные критерии, установленные спонсором ИТ проекта. Обычно не берут в расчет более широкий контекст и дальнюю перспективу
Поставщики	Неоднородная группа, снабжающая ИТ проект необходимыми товарами и услугами	Так же как и команда проекта, акцентируют внимание на достижении краткосрочных целей
Прочие заинтересованные стороны (государство, общество, некоммерческие организации)	Их существование или деятельность затрагиваются реализацией или результатами ИТ проекта, но они не всегда получают прямую выгоду	Всегда есть заинтересованные стороны, которые поддерживают и которые препятствуют реализации ИТ проекта. Никогда нельзя полностью отбрасывать эту категорию заинтересованных лиц

* В широком смысле под организацией можно понимать государство, некоммерческие организации или бизнес.

ральному значению при помощи математического инструмента для сбалансированного достижения целей заинтересованных сторон в порядке их значимости.

Таким образом, для выбора ИТ проекта, наилучшим образом соответствующего требованиям заинтересованных сторон, необходимо создать методику количественной и качественной оценки. В связи с вышеизложенным создание методики, позволяющей определить количественно привлекательность ИТ проекта, а также ранжирование этих показателей – актуальная и практически востребованная задача.

Характерные проблемы, связанные с уже реализованными ИТ проектами, хорошо иллюстрирует пример американского опыта реализации масштабного интегрированного ИТ проекта по созданию электронного правительства [4]:

Программа оценки деятельности. Агентства обычно оценивают свои ИТ системы исходя из того, как хорошо они поддерживают процессы и нужды агентства, а не из того, насколько хорошо отвечают нуждам населения. Системы часто оцениваются по проценту рабочего времени вместо внутренних и внешних показателей выгоды, которую они приносят тем программам правительства, которые поддерживают.

Создание адекватной системы оценки – наиболее существенный элемент подготовки к реализации ИТ проектов в направлении построения электронного правительства.

Технологический рычаг. В 1990-х правительственные агентства использовали ИТ для автоматизации существующих процессов вместо того, чтобы создать более эффективные и результативные решения, которые сейчас сформированы по результатам уроков, полученных в ходе работы электронного бизнеса.

Островки автономии. Агентства в основном покупают системы, которые ориентированы на внутренние нужды и в меньшей степени – системы, способные взаимодействовать и обмениваться информацией с другими агентствами. Следовательно, граждане, чтобы получить необходимую услугу, вынуждены осуществлять поиск по разным агентствам, бизнес вынужден отправлять файлы с одинаковой информацией множество раз, а государственным агентствам непросто распространять информацию.

Сопротивление изменениям. Бюджетный процесс и культура агентства увековечивают абсолютную бюрократию. Бюджетный процесс не предоставляет механизма для инвестиций в информационные технологии, способные интегрировать деятельность различных агентств. Более того, культура агентств и боязнь реорганизаций создают сопротивление работе по интеграции и распространению использования информационных технологий.

Методика, предлагаемая как возможный вариант решения проблем, описанных выше, рассматривает измерение и количественную оценку ИТ проекта в соответствии с определением качества, заложенного в международные стандарты серии ИСО 9000, как степени соответствия требованиям (и предполагаемым потребностям) потребителей. В роли потребителей ИТ проекта можно рассматривать заинтересованные стороны, представленные в табл. 1.

Итак, в результате анализа зарубежного и отечественного опыта была предложена методика по оценке привлекательности ИТ проектов с учетом заинтересованных сторон и классификации информационных технологий. Процесс проведения оценки, демонстрирующий нелинейность принятия решения по выбору наиболее привлекательного ИТ проекта для реализации, включает в себя следующие этапы:

1. Определение "ключевой" заинтересованной стороны, которая проводит оценку и принимает необходимые решения (точка зрения на ИТ проект). Ранжирование значимости других заинтересованных сторон.

2. Определение иерархической структуры и перечня необходимых показателей для оценки привлекательности ИТ проекта.

Многие авторы (например, [5]) предлагают использовать именно иерархию показателей для измерения "отдачи от ИТ проекта", позволяющую учитывать все необходимые аспекты.

Иерархия и перечень показателей выбираются с учетом "ключевой" заинтересованной стороны и класса применяемых информационных технологий на основании разработанных таблиц для указанного сочетания "ключевая заинтересованная сторона – класс информационных технологий". Пример приведен в табл. 2. Ключевая заинтересованная сторона – "Инвестор", класс информационных технологий – "Специализированное программное обеспечение".

3. Взаимосвязи между показателями как одного, так и различных уровней иерархической

Таблица 2

Пример иерархии и перечня показателей для оценки привлекательности ИТ проекта

Иерархия	Категория показателей	Измеряемые показатели
Уровень 1	Отдача на инвестированный капитал (ROI)	В денежном выражении
Уровень 2	Сумма расходов и доходов	В денежном выражении
Уровень 3	Улучшение качества планирования и контроля	Время, необходимое для планирования. Затраты на планирование. Время руководителей, необходимое для контроля. Количество предупреждений. Затраты на контроль
Уровень 4	Качество принимаемых решений	Частота неудач и полного изменения решений. Количество рассмотренных альтернатив. Время, необходимое для принятия решения. Количество решений. Стоимость решений
Уровень 5	Ценность информации	Полезность (в терминах обоснованности, точности, ясности, частоты, достаточности, своевременности, достоверности, уместности, содержания и стоимости)
Уровень 6	Характеристики системы	Необходимое количество людей, оборудования и поддержки. Время ответа. Частота поломок. Входы и выходы. Количество форм. Количество операций. Количество хранилищ и их размер. Качество банка данных, размер и качество модели банка. Гибкость, простота, степень автоматизации. Масштаб бизнес-компонент, связанных с системой. Удовлетворенность потребителей. Уровень ошибок. Области постоянных проблем. Простота поддержания и модификации. Незапланированное влияние на деятельность компании. Экономия и затраты

структуры определяются при помощи математической модели, поддерживающей процесс количественной оценки привлекательности ИТ проектов. Она основана на применении логики антонимов [6]. За счет формализации этих связей гибко учитывается влияние показателей, определяющих дальнейшие сценарии проведения оценки (в том числе обращение оценки в ноль, что соответствует отказу от ИТ проекта).

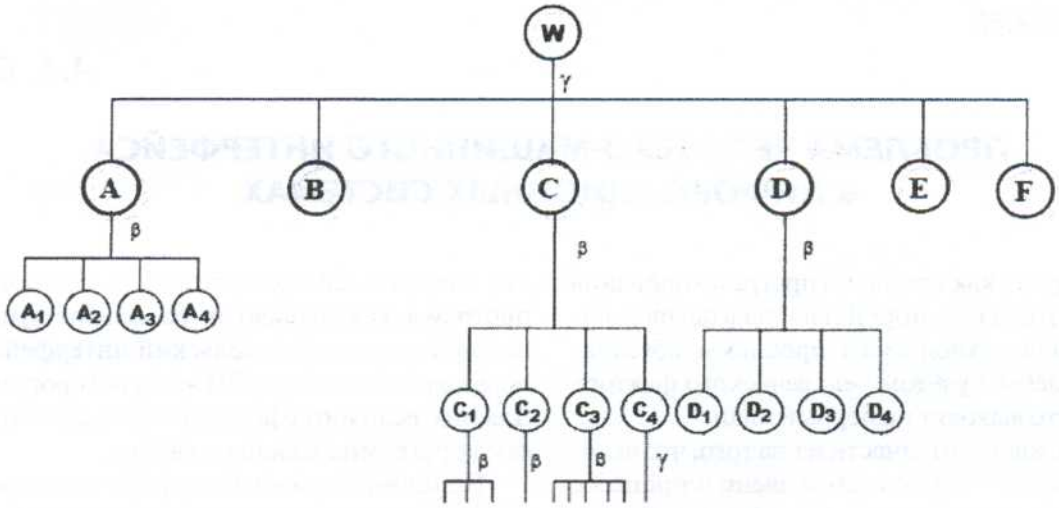
Для получения интегрального показателя привлекательности ИТ проекта необходимо построение графической модели иерархии показателей с учетом связей между ними (см. рисунок).

Для характеристики связей применяются операторы логики антонимов: β – "слабая связь", γ – "сильная связь". Дополнительно с использованием экспертных оценок каждому из показателей необходимо поставить в соответствие определенный весовой коэффициент, учитывающий его важность для интегральной оценки. Аксиоматика логики антонимов предлагает формулы, определяющие два типа связи между показателями:

$$H[A\beta B] = \rho_1 H[A] + \rho_2 H[B], \quad (1)$$

$$H[A\gamma B] = -\log_2\left[1-\left(1-2^{-\rho_1 H[A]}\right)\left(1-2^{-\rho_2 H[B]}\right)\right], \quad (2)$$

где ρ_i – весовой коэффициент i-го показателя.



Графическая модель показателей

Далее производится преобразование графической модели в формулу, которая позволяет получить интегральную оценку привлекательности ИТ проекта. Она составляется в соответствии с аксиоматикой логики антонимов, на основании выражений (1) и (2). Для графической модели, представленной на рисунке, интегральная оценка в общем виде выглядит следующим образом:

$$H[W]=H[A\ \gamma B\ \gamma C\ \gamma D\ \gamma E\ \gamma F]. \qquad (3)$$

4. Для корректного проведения оценки важно выбрать соответствующий экспертный совет, в который включаются представители основных заинтересованных сторон, а также эксперты и специалисты в предметных областях.

5. Принятие решения. По результатам оценки привлекательности ИТ проектов уполномоченное лицо или орган принимает решение о реализации проекта (либо об отказе от реализации).

Предложенная методика направлена на повышение достоверности оценки привлекательности ИТ проектов и ускорение процесса принятия управленческого решения за счет ранжирования результатов оценивания и выбора наиболее предпочтительного варианта в соответствии с требованиями заинтересованных сторон.

Развитие подходов к оцениванию сложных многомерных систем с противоинтуитивным поведением требует применения новых моделей, учитывающих комплекс потребностей, связанных с конкретным ИТ проектом и использующих современный аппарат теоретико-методологических положений для формализации логики оценки.

Однако не следует считать, что оценивание привлекательности ИТ проекта – это точный процесс, приводящий к гарантированно единственному результату. При проведении оценки многое зависит от конкретных параметров, относительный вес которых часто определяется экспертным путем.

СПИСОК ЛИТЕРАТУРЫ

1. **Каплан Р.С., Нортон Д.П.** Сбалансированная система показателей. От стратегии к действию. 2-е изд., испр. и доп. : Пер. с англ. М.: ЗАО "Олимп-Бизнес", 2003. 320 с.: ил.

2. **Chandler A.D., Scale Jr., Scope.** The Dynamics of Industrial Capitalism. Cambridge, Mass.: Harvard University Press, 1990.

3. Macroergonomics: theory, methods and applications / Ed. H.W. Hendrick, B.M. Kleiner. 2002.

4. **Itami H.** Mobilizing Invisible Assets. Cambridge, Mass.: Harvard University Press, 1987.

5. E-Government Strategy: Simplified Delivery of Services to Citizens / Office of management and budget. Executive officer of the President of the United States. 2002. Feb 27.

6. **Голюта Я.Я., Тисенко В.Н., Фальков Д.С.** Логика антонимов – теоретическая база для формирования комплексных оценок, в основе которых лежат экспертные оценки отдельных параметров // Моделирование интеллектуальных процессов проектирования и производства (CAD/CAM/98): Матер. 2-й междунар. науч.-техн. конф. Минск, 1998. С. 166–167.

УДК 004.514.62

А.А. Сухов

ПРОБЛЕМА ЧЕЛОВЕКО-МАШИННОГО ИНТЕРФЕЙСА В ИНФОРМАЦИОННЫХ СИСТЕМАХ

В то время как проблема программирования АСУТП отошла на второй план благодаря мощным SCADA-технологиям, проблемы, обусловленные слабым учётом человеческого фактора (ЧФ), часто выходят на первый план.

Происходит это отчасти из-за того, что человек-оператор – самое слабое звено в процессе управления производством, отчасти из-за слабой эргономической культуры проектировщиков.

Инженер-проектировщик, пользуясь ТЗ и техническими спецификациями, с помощью средств SCADA-системы "собирает" интерфейс. Основной своей задачей инженер считает полностью обеспечить оператора набором функциональных возможностей, предписанных проектными спецификациями. Из одних и тех же "интерфейсных кубиков" можно собирать различные пользовательские интерфейсы с одинаковой функциональностью. Разрабатывая ПИ, инженер руководствуется, как правило, лишь собственным опытом и знаниями, иногда применяет первое попавшееся проектное решение (интерфейсное) и оставляет его в таком виде до тех пор, пока не возникнут серьёзные нарекания. Иногда инженер из лучших побуждений просто "копирует" интерфейсные решения из аналогов или старых прототипов АРМ. Так размножаются грубые эргономические ошибки [1].

В продвинутых фирмах совместно с инженером-проектировщиком работают инженер-технолог и даже иногда художник-конструктор ("дизайнер"). Это серьёзное подспорье. Но технолог не знает, например, тонкостей восприятия информации человеком, а дизайнер заботится в основном лишь об эстетической и художественной стороне визуализации ПИ. Не всё, что "классно нарисовано", эффективно работает.

Разработчики программных комплексов иногда склонны рассматривать функциональность системы отдельно от её пользовательского интерфейса. При этом предполагается, что ПИ есть своего рода дополнение к функциональности системы. Со своей стороны, пользователи программ, как правило, не разделяют функциональность и пользовательский интерфейс. Для пользователей именно ПИ является программой. Для них, если интерфейс хороший, стало быть, и сама программа хороша и удобна.

Пользовательский интерфейс часто понимают только как внешний вид программы. Однако на деле пользователь воспринимает через ПИ всю систему в целом, а значит, такое понимание ПИ слишком узко. В действительности ПИ включает в себя все аспекты дизайна, которые оказывают влияние на взаимодействие пользователя и системы.

Проектирование ЧМИ превратилось в самостоятельную проблему, часто превосходящую по сложности проблему разработки кодов программы, и требует, как и процесс проектирования любой сложной системы, соответствующих методов, средств и усилий квалифицированных специалистов. Само применение термина "человеко-машинный интерфейс" представляет собой попытку разработчиков программного обеспечения отделить, по крайней мере концептуально, функциональное назначение программных продуктов от проблем, связанных с организацией взаимодействия пользователя с этими продуктами. Такое разделение – необходимое условие создания "дружественных" интерфейсов по отношению к пользователям программных продуктов. Само понятие дружественности – не одномерная величина, а вектор, содержащий согласно международной классификации семь компонент:

соответствие задачам, решаемым пользователем;
легкость применения;
управляемость;
соответствие ожиданиям пользователя;
устойчивость к ошибкам;
адаптируемость/индивидуализируемость;
легкость изучения.

А согласно [3]:

«4.3. Практичность (Usability) – набор атрибутов, относящихся к объему работ, требуемых для использования и индивидуальной оценки такого использования определенным или предполагаемым кругом пользователей.

Примечания. "Пользователи" могут интерпретироваться как большинство непосредственных пользователей *интерактивного программного обеспечения*. Круг пользователей может включать операторов, конечных пользователей и косвенных пользователей, на которых влияет данное программное обеспечение или которые зависят от его использования. Практичность должна рассматриваться во всем разнообразии условий эксплуатации пользователем, которые могут влиять на программное обеспечение, включая подготовку к использованию и оценку результатов.

Практичность, определенная в данном стандарте как конкретный набор атрибутов программной продукции, отличается от определения с точки зрения эргономики, где рассматриваются как составные части практичности другие характеристики, такие, как эффективность и неэффективность».

Современные подходы к проектированию человеко-машинного интерфейса (ЧМИ) базируются на определенной методологической основе, которая выделяет три ключевые проблемы организации процесса проектирования ЧМИ: идентификацию информации, необходимой для проектирования, определение и структурирование собственно процесса проектирования и, наконец, цели и порядок проведения эргономической экспертизы. Для иллюстрации сложности этого процесса приведем без комментариев две схемы, характеризующие процесс проектирования ЧМИ (рис. 1).

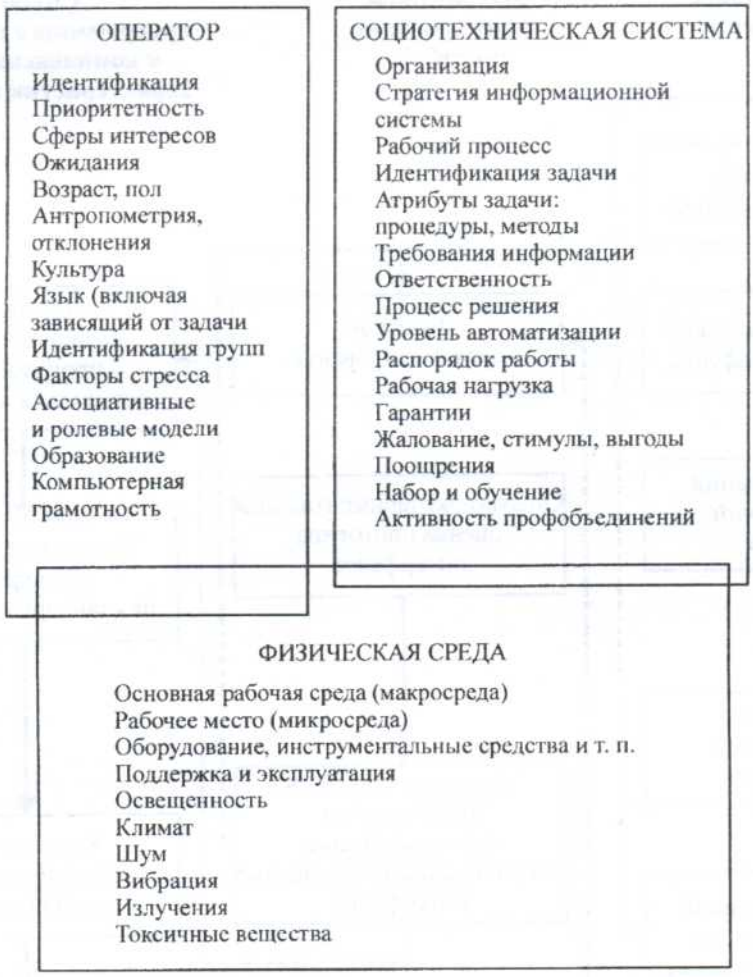


Рис. 1. Обобщенная структура информации, необходимой для проектирования

Среди задач, которые необходимо решать при проектировании интерфейса, можно выделить:

проектирование деятельности оператора по выполнению конкретных работ с помощью создаваемой системы;

проектирование архитектуры и функций интерфейса;

концептуализацию и реализацию форм диалога;

разработку техники общения оператора с интерфейсом, форматов отображения информации и т. д.

Перечисленные задачи взаимосвязаны и поэтому в процессе проектирования интерфейсов решаются в определенной последовательности.

Для значительного числа подходов к проектированию интерфейсов основные принципы таковы:

ориентация на потенциального пользователя (оператора) на всех этапах создания системы;

интерактивность и итеративность процесса проектирования;

опытно-экспериментальная проверка проекта интерфейса на каждом этапе его разработки.

В [2] предлагалась итеративная модель разработки пользовательского интерфейса, которая бы использовала постоянный циклический контроль качества и проверку функционирования

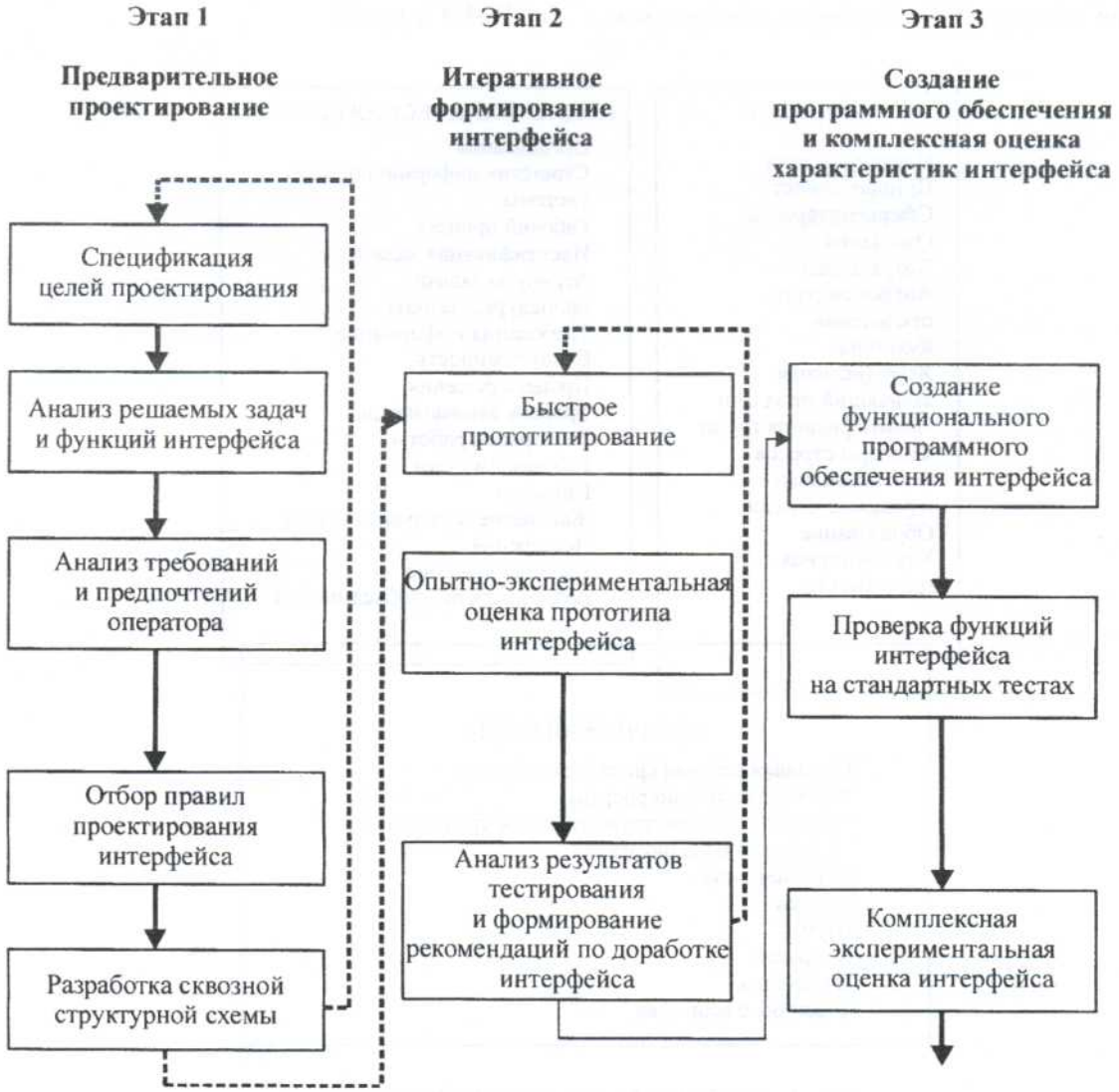


Рис. 2. Модель разработки пользовательского интерфейса

(рис. 2). Согласно [4] существует несколько методик оценки качества ЧМИ.

Методы оценки ЧМИ по показателям качества

Методы определения показателей качества ЧМИ различаются как по способам получения информации о ЧМИ (измерительный, регистрационный, органолептический, расчетный), так и по источникам получения информации (традиционный, экспертный, социологический).

Несмотря на многообразие методов, в большинстве случаев проверку качества всегда выполняет эксперт с большим опытом работы, который может помочь в улучшении эргономики ЧМИ. И привлечение специалистов было бы самым эффективным способом, однако проблема в том, что их очень мало, а подготовка собственных кадров может занимать длительное время и имеет смысл лишь тогда, когда фирма профессионально занимается созданием ЧМИ.

Автором также предлагается использовать итеративный подход к созданию интерфейса (т. е. цикл создания, оценки, изменения). Цикл должен повторяться до тех пор, пока результаты оценки не станут удовлетворительными. Однако для проверки качества интерфейса и пути возможного решения возникающих проблем желательно использовать нейронные сети.

Для этого необходимо спроектировать интерфейс в виде непрерывного автомата, который затем можно реализовать в виде нейронной сети. Покажем возможность этого.

Класс функций, вычисляемый с помощью нейронных сетей, замкнут относительно линейных операций. Действительно, пусть есть нейронные сети $S_1, S_2, ..., S_k$, которые вычисляют функции $F_1, F_2, ..., F_k$ от вектора входных сигналов x. Линейная комбинация $a_0 + a_1F_1 + a_2F_2 + ... + a_kF_k$ вычисляется сумматором с весами $a_0, a_1, a_2, ..., a_k$, на вход которого подаются выходные сигналы сетей $S_1, S_2, ..., S_k$. Разница в числе тактов функционирования этих сетей до получения ответа легко компенсируется "линиями задержки", составленными из связей с единичным весом.

Кроме того, класс функций, вычисляемый с помощью нейронных сетей, замкнут относительно унарной операции, осуществляемой нелинейным преобразователем сигнала, входящим в состав нейрона: если сеть S вычисляет функцию F, то, подавая выход этой сети на вход нелинейного преобразователя, получим на его выходе функцию $j(F)$.

Используя теорему Колмогорова, которая доказывает, что *каждая непрерывная функция n переменных, заданная на единичном кубе n-мерного пространства, представима в виде*

$$f(x_1, x_2, ..., x_n) = \sum_{q=1}^{2n+1} h_q \left[\sum_{p=1}^{n} \varphi_q^p(x_p) \right],$$

где функции $h_q(u)$ непрерывны, а функции $\varphi_q^p(x_p)$, кроме того, еще и стандартны, т. е. не зависят от выбора функции f, можно показать, что множество функций, вычисляемых нейронными сетями с заданной непрерывной нелинейной характеристической функцией, плотно в пространстве непрерывных функций от входных сигналов.

Если есть возможность с любой точностью приблизить любую непрерывную функцию и нет ограничений на способы соединения устройств, то можно сколь угодно точно имитировать работу любого непрерывного автомата. Покажем это.

Каждый автомат имеет несколько входов (n), несколько выходов (p) и конечный набор (s) параметров состояния. Он вычисляет $s + p$ функций от $n + s$ переменных. Аргументы этих функций – их входные сигналы (n) и текущие параметры их состояния (s). Значения функций – их выходные сигналы (p) и параметры состояния на следующем шаге (s). Каждый такой автомат можно представить как систему из $s + p$ более простых автоматов (рис. 3). Эти простые автоматы вычисляют по одной функции от $n + s$ переменных. Смена состояний достигается за счет того, что часть значений этих функций на следующем шаге становится аргументами – так соединены автоматы (см. рис. 3). Таким образом, без потери общности можно рассматривать сеть автоматов как набор устройств, каждое из которых вычисляет функцию нескольких переменных $f(x1, ..., xn)$. Этот простой, но фундаментальный факт позволяет использовать предыдущие результаты. Нейронные сети дают возможность с любой точностью вычислять произвольную непрерывную функцию $f(x1, ..., xn)$. Следовательно, с их помощью можно сколь угодно точно аппроксимировать функционирование любого непрерывного автомата.

Спроектированную таким образом сеть можно использовать для получения временных параметров срабатывания объектов управления

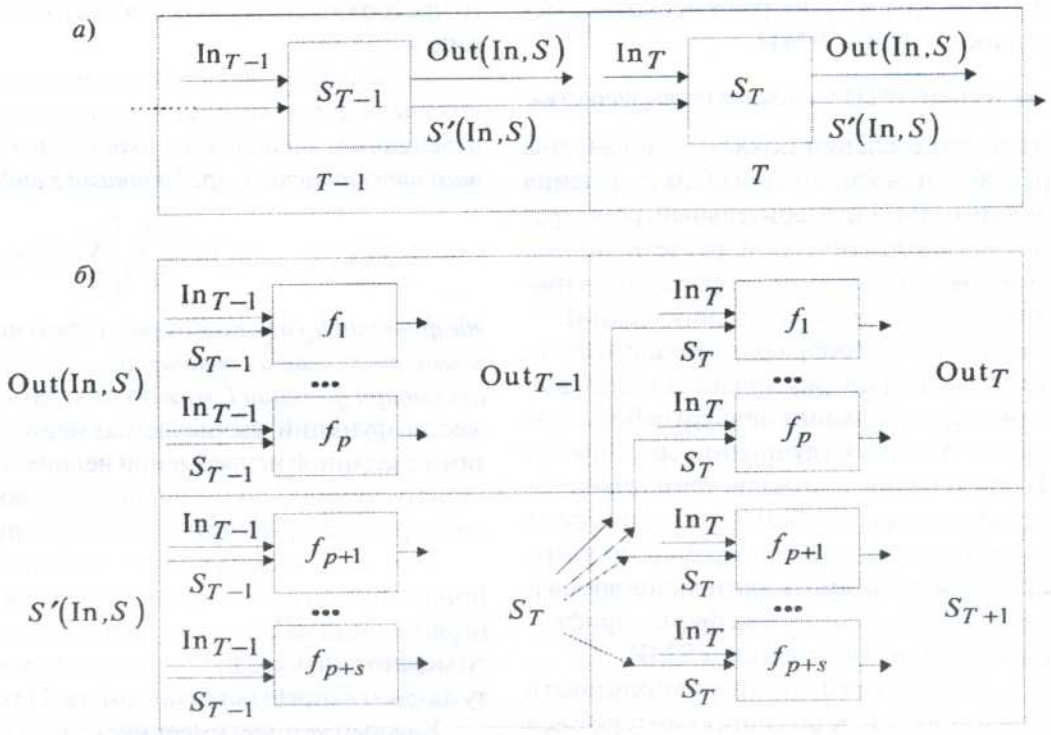


Рис. 3. Представление общего автомата с помощью модулей, вычисляющих функции многих переменных от входных сигналов: *а* – функционирование автомата; *б* – разложение его на отдельные модули

In, Out – входные и выходные сигналы; S – параметры состояния; T – дискретное время; Out (In, S) – зависимость выходных сигналов от значений входных и параметров состояния; S'(In, S) – зависимость состояния в следующий момент дискретного времени от входных сигналов и текущего состояния; $f_1 - f_p$ и $f_{p+1} - f_{p+s}$ – функции переменных (In, S) компонент вектора соответственно Out(In, S) и S'(In, S); T, $T\pm1$ – индексы, соответствующие моментам времени

ЧМИ. Также на основе нейронных сетей я предлагаю проводить проверку не только временных параметров, но также эмоциональной оценки интерфейса (цветовой гаммы, насыщенности), так как очевидно, что эти параметры очень сильно будут влиять на работу оператора. Использование правильно обученных нейронных сетей позволит применить опыт экспертов при оценке интерфейса, одновременно автоматизировав и стандартизировав процесс оценки.

В статье рассмотрен метод создания человеко-машинного интерфейса при помощи нейронных сетей. Данный метод позволит использовать плюсы экспертных оценок, одновременно автоматизировав и стандартизировав процесс оценки.

СПИСОК ЛИТЕРАТУРЫ

1. **Перевалов Я.М.** Проблемы проектирования пользовательских интерфейсов SCADA-систем: Докл. на IV Всерос. конф. «Разработка АСУТП в системе "Трейс Моуд": задачи и перспективы» // www.usability.ru

2. **Перевалов Я.М.** Технологии проектирования пользовательских интерфейсов АРМ АСУТП: Докл. на V Всерос. конф. «Разработка АСУТП в системе "Трейс Моуд": задачи и перспективы» // www.usability.ru

3. ГОСТ 28195-89. Оценка качества программных средств.

4. ГОСТ ИСО/МЭК 9126-93. Информационная технология. Оценка программной продукции. Характеристики качества и руководства по их применению.

5. О важности эргономики / Susan Dray. Переводчик: Александр Белышкин. Interaction, янв. 1995.

УДК 004.94

Л.А. Хворова, В.М. Брыксин, Н.В. Гавриловская

ИСПОЛЬЗОВАНИЕ ИНФОРМАЦИОННЫХ ТЕХНОЛОГИЙ ПРИ ПРОГНОЗИРОВАНИИ УРОЖАЙНОСТИ ЗЕРНОВЫХ КУЛЬТУР

Работоспособные технологии прогноза урожайности зерновых культур, интегрированные в сельскохозяйственное производство и основанные на математических моделях и космической информации, в условиях России отсутствуют. В связи с этим решение проблемы осуществления упреждающего прогноза по адекватной математической модели с использованием данных космосъемки весьма актуально.

Особое место в технологии прогноза урожайности зерновых культур должны занимать прогнозы, позволяющие давать оценку ожидаемой продуктивности посевов поэтапно. Это значительно увеличивает надежность прогнозов урожайности и эффективность использования прогностической информации. Методика поэтапного прогнозирования создавалась авторами в рамках разработки технологии определения года-аналога [4, 5].

Применение адаптированной модели биопродуктивности EPIC и космоснимков MODIS для прогнозирования урожайности зерновых культур на территории Западной Сибири

По мнению многих ученых, основу методического и информационного обеспечения работ по оценке урожайности зерновых культур должны составлять динамические имитационные модели биопродуктивности сельскохозяйственных культур, комплексы математических моделей прогнозирования урожайности с использованием данных дистанционного зондирования Земли.

Авторами разработана технология прогноза урожайности зерновых культур с использованием космических снимков и компьютерного моделирования, интеллектуальным ядром которого является модифицированная динамическая модель EPIC [6], разработанная Сельскохозяйственной научно-исследовательской службой департамента сельского хозяйства Соединенных Штатов (Soil and Water Research Laboratory USDA) [1, 3].

Технология прогноза урожайности зерновых культур реализована в виде комплекса программных средств MIDC (модель имитации развития культур) и апробирована в условиях зерновых провинций Западной Сибири – Новосибирской и Тюменской областей. По данной модели ежедневно рассчитываются основные параметры культуры (биомасса, листовой индекс, высота растения и др.). Это позволяет сравнивать их с данными наземных измерений и дистанционного зондирования и в случае необходимости производить корректировку. Используя фактические данные по сложившимся агрометеорологическим условиям, программный комплекс позволяет осуществлять прогноз урожайности за месяц до уборки.

Необходимо отметить, что модель EPIC разрабатывалась для оценки последствий эрозии почв на продуктивность посевов, что, естественно, наложило отпечаток на ее структурную и алгоритмическую реализацию. Поэтому на первый план вышли проблемы ее идентификации по имеющейся агрометеорологической информации и реализации цели исследования.

Этапы структурно-параметрической идентификации модели EPIC

Процедура идентификации модели EPIC и разработка на ее основе имитационно-моделирующего комплекса MIDC включала шесть этапов:

1) тестирование модели EPIC (выявлялась реакция по блокам и модели в целом на изменение основных агрометеорологических факторов);

2) модификацию блоков и структуры модели EPIC по результатам тестирования и на основе имеющейся агрометеорологической и спутниковой информации;

3) конструирование на базе технологии EPIC имитационно-моделирующего комплекса MIDC;

4) определение параметров, входящих в блоки: коротковолновой (КВР), длинноволновой (ДВР) радиаций, теплового режима почвы и функции суточной динамики температуры воздуха;

5) комплексную идентификацию модели по экспериментальным данным регионов;

6) апробацию методики прогноза урожайности в реальных условиях Западной Сибири.

Рассмотрим результаты отдельных этапов идентификации.

Тестирование модели EPIC проводилось с использованием данных, предоставленных Западно-Сибирским управлением гидрометслужбы по Новосибирской области и Тюменским центром гидрометслужбы по Тюменской области. Выяснилось, что предоставленные наборы агро- и метеоданных не полные, отсутствуют необходимые для модели величина приходящей КВР и температура поверхности почвы. В связи с этим возникла задача восполнения неполноты исходной агро- и метеоинформации за счет разработки блоков радиации (КВР и ДВР).

При отсутствии экспериментальных данных по приходящей коротковолновой радиации в EPIC генерируются данные, сильно отличающиеся от реальных. В результате возникают большие погрешности в расчетах [3]. Как показал опыт сбора необходимой агрометеоинформации, многие метеостанции коротковолновую радиацию не измеряют. Поэтому возникает проблема модификации и настройки модели КВР для регионов Западной Сибири.

Важный входной параметр модели – температура поверхности почвы (минимальная и максимальная). Как правило, такая информация также отсутствует на многих метеостанциях. В связи с этим возникла задача расчета теплового режима почвы (определение минимальной и максимальной температуры на поверхности почвы и расчет ее температурного профиля).

Выявленные недостатки при тестировании модели EPIC потребовали разработки дополнительного программного модуля, состоящего из трех взаимосвязанных блоков: блока ДВР, или эффективного излучения (для определения температуры на поверхности почвы в ночное время), блока КВР (для определения температуры на поверхности почвы в дневные часы) и блока теплового режима почвы. Блочная структура модели EPIC делает ее гибкой и дает возможность включать новые блоки (рис. 1).

Определение температуры на поверхности почвы. Как известно, днём поверхность почвы нагревается прямой солнечной радиацией, а ночью – охлаждается в результате эффективного излучения, и от неё охлаждаются прилегающие слои воздуха [2]. Причем минимальная температура воздуха наблюдается перед восходом Солнца, а максимальная температура отмечается через два-три часа после полудня.

По экспериментальным данным установлена зависимость между максимальной температурой почвы и суммарной радиацией:

$$T_{p\max} = 1{,}3 \cdot T_{v\max} + 0{,}5 \cdot Q + 0{,}5,$$

где $T_{p\max}$ – максимальная температура на поверхности почвы; $T_{v\max}$ – максимальная температура воздуха; Q – суммарная коротковолновая радиация.

В ночное время наблюдается взаимосвязь между минимальной температурой почвы и эффективным излучением. В качестве аппроксимирующей математической функции было выбрано уравнение линейной двухфакторной регрессии

$$T_{p\min} = 0{,}8 \cdot T_{v\min} + 1{,}5 \cdot E + 4{,}6,$$

где $T_{p\min}$ – минимальная температура на поверхности почвы; $T_{v\min}$ – минимальная температура воздуха; E – эффективное излучение.

Параметры уравнения регрессии оценивались по выборочному, но достаточно большому временному ряду экспериментальных значений температуры на поверхности почвы и расчетным значениям E из модели ДВР [2] по методу наименьших квадратов. С помощью корреляционного анализа осуществлена проверка адекватности полученной модели.

Температура профиля почвы рассчитывается с использованием уравнения теплопереноса.

Суточная динамика температуры воздуха. При расчетах длинноволновой радиации и температурного профиля почвы необходимо использовать данные по температуре воздуха и верхнего слоя почвы с шагом времени в один час и менее. Однако в предоставляемых данных в большинстве случаев присутствует лишь информация с шагом в одни сутки. Для решения данной проблемы предложено воспользоваться кусочно-



Рис. 1. Блок-схема модернизированной модели EPIC
(серым цветом выделены измененные блоки)

непрерывной функцией $F(t, j)$, где t – время от полуночи в часах, j – номер дня от начала года, построенной на основе минимальной и максимальной температур воздуха исходя из следующих соображений:

минимальная температура наблюдается на рассвете, а максимальная – через два часа после полудня;

к закату температура падает до значения $T_{\min j+1} + k(T_{\max j} - T_{\min j+1})$, где коэффициент $k \in [0; 1]$ рассчитывается по экспериментальным данным;

в интервалах $[t_{d,j}; t_m + 2]$, $[t_m + 2; t_{s,j}]$ и $[t_{s,j}; t_{d,j+1}]$ функция $F(t, j)$ задается квадратичной зависимостью $at^2 + bt + c$, где t_d, t_m, t_s – время рассвета, полудня и заката соответственно. При этом $F(t, j)$ имеет следующие точки экстремума для соответствующих интервалов: t_{m+2} – максимум для интервала $[t_{d,j}; t_{m+2}]$; t_{m+2} – максимум для $[t_{m+2}; t_{s,j}]$; t_d – минимум для $[t_{s,j}; t_{d,j+1}]$. Исходя из вышеперечисленных соображений составляем три системы линейных уравнений для поиска коэффициентов a, b, c. При этом учитываем, что в точках стыковки функция должна быть непрерывна:

$$1)\begin{cases} F(t_d, j) = T_{\max j} \\ F(t_m + 2, j) = T_{\max j} \text{ при } t \in [t_{d,j}; t_m + 2]; \\ F'(t_m + 2, j) = 0 \end{cases}$$

$$2)\begin{cases} F(t_m + 2, j) = T_{\max j} \\ F(t_{s,j}, j) = T_{\min j+1} + k\left(T_{\max j} - T_{\min j+1}\right) \\ \text{при } t \in [t_m + 2; t_{s,j}]; \\ F'(t_m + 2, j) = 0 \end{cases}$$

$$3)\begin{cases} F(t_{s,j}, j) = T_{\min j+1} + k\left(T_{\max j} - T_{\min j+1}\right) \\ F(t_{d,j+1}, j) = T_{\min j+1} \\ \text{при } t \in [t_{s,j}; t_{d,j+1}]; \\ F'(t_{d,j+1}, j) = 0 \end{cases}$$

При нахождении a, b, c учитывалось, что интервал $[t_{s,j}; t_{d,j+1}]$ с учетом смены дат разбивается на два интервала: $[0; t_{d,j}]$ и $[t_{s,j}; 24]$. Таким образом получаем четыре набора коэффициентов a, b, c для соответствующих интервалов.

На рис. 2 приведен пример суточной динамики расчетной и фактической температуры воздуха, °С, для с. Ключи Алтайского края (03.07.2000). Погрешность вычислений составляет 4,7 %.

Корректировка листового индекса и биомассы по спутниковым данным. Значения вегетационного индекса (*NDVI*), рассчитываемого по космоснимкам сканера MODIS с космических аппаратов TERRA и AQUA, хорошо коррелируют с такой характеристикой растительного покрова, как листовой индекс (*LAI*). Авторами алгоритма MOD15 (NASA) для обработки изображений MODIS были получены кусочно-линейные зависимости между *LAI* и *NDVI* для различных ландшафтов. Расчет индекса *NDVI* проводился по контурам тестовых полей.

На основе модельных и фактических данных были получены зависимости, приведённые ниже, биомассы и массы корней от листового индекса, рассчитанные на основе индекса *NDVI*, что позволяет производить корректировку моделируемых значений:

$BIOM = 1{,}09 \cdot LAI_{NDVI} - 0{,}57$ при росте *LAI*;

$BIOM = BIOM_{DM} + 0{,}26(LAI_{DM} - LAI_{NDVI})$ при падении *LAI*;

$RW = 0{,}31 \cdot BIOM + 0{,}03$,

где *BIOM* и *RW* – моделируемые значения биомассы и массы корня; LAI_{NDVI} – листовой индекс, полученный из индекса *NDVI*; LAI_{DM} и $BIOM_{DM}$ – моделируемый листовой индекс и биомасса в день, когда *LAI* максимален.

Данная модель позволяет производить корректировку в процессе расчетов. Например, корректировку листового индекса *LAI* можно выполнять по данным дистанционного зондирования через индекс *NDVI* (нормализованный вегетационный индекс), который линейно связан с *LAI*.

Создание технологии прогноза урожайности зерновых культур по отдельным областям Западной Сибири с использованием космических снимков и компьютерного моделирования включало следующие этапы:

1. Адаптацию модели биопродуктивности EPIC и программных средств для расчета урожайности [1]. Данный этап включает в себя разработку не только базы наземных данных и обработанных данных дистанционного зондирования, но и блок-схем алгоритмов расчетов по прогнозу урожайности.

2. Разработку методики сбора и обработки исходных наземных данных и данных дистанционного зондирования. Этап предполагает сбор необходимой агрометеоинформации для модели в гидрометеослужбах регионов и в результате наземного обследования тестовых полей.

3. Разработку алгоритмов и программных средств поддержки модели, включая проверку адекватности. Этап предполагает усовершенствование модели с включением вновь разработанных блоков: приходящей КВР и ДВР, теплового режима почвы.

4. Апробацию методики прогноза урожайности в реальных условиях Западной Сибири.

В течение вегетационного периода для контуров тестовых полей по космоснимкам вычисляется индекс *NDVI*, по которому рассчитывается листовой индекс *LAI* и передается в модель для корректировки расчетов.

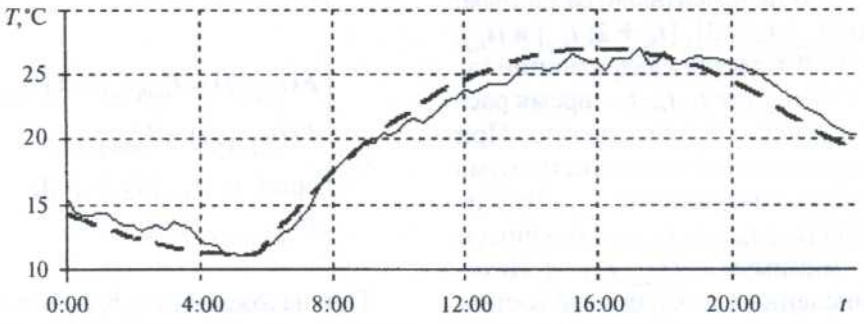


Рис. 2. Суточная динамика
расчетной (———) и фактической (— —) температуры воздуха

По заказу Департамента агропромышленного комплекса администрации Тюменской области в 2004 году проводились работы по мониторингу состояния сельскохозяйственных полей и прогнозированию урожайности яровых зерновых культур. Для каждого административного района Тюменской области, имеющего метеостанцию, по модифицированной модели EPIC было проведено моделирование роста и развития зерновых культур (яровая пшеница, ячмень, овес, горох). На этом этапе осуществлен подбор параметров модели для каждого района (определялись усредненные сроки сева и почвенные характеристики модели) и общие для всех районов варьируемые характеристики (индекс урожайности и другие параметры). Определение даты начала уборки проводилось по двум критериям: прекращению наращивания биомассы и падению листового индекса до нуля. Погрешность расчетов урожайности по модели для Тюменской области в целом не превышает 1 ц/га.

Для регионов Западной Сибири также проводилась коррекция листового индекса с использованием зависимости *LAI* от *NDVI* по данным сканера MODIS. Так, например, для Коченёвского района Новосибирской области в 2005 году урожайность по модели без коррекции листового индекса составила 13,29 ц/га, с коррекцией – 15,35 ц/га. Фактическая урожайность 15,7 ц/га.

Динамика прогнозируемой урожайности в зависимости от фактических метеоусловий представлена на рис. 3. Видно, что колебания в расчетных значениях биологической урожайности яровой пшеницы минимальны в конце июля. Это подтверждает возможность прогноза урожайности яровой пшеницы на юге Западной Сибири за месяц до начала уборки.

Технология определения года-аналога

Модели любого уровня, осуществляющие прогноз, должны опираться на предполагаемые погодные сценарии прогнозного года. Отсутствие фундаментальной теории, объясняющей наличие связи между метеорологическими процессами в отдельные сезоны и урожайностью, обусловливает применение в методиках в основном аппарата математической статистики, в частности кластерного и дискриминантного анализа. В результате использования данных процедур каждый год на основе различных характеристик может быть отнесен к одному из нескольких кластеров.

Процедура определения года-аналога осуществляется поэтапно [5]. До сева яровых (в мае) по уже известным погодным данным определяется год-аналог для данного года, по погодному сценарию которого осуществляется прогноз урожайности по разработанной эмпирической модели урожайности и динамической модели MIDC [1]. Уточнение оценки урожайности производится через две декады после сева и в другие периоды вегетационного периода с корректировкой года-аналога.

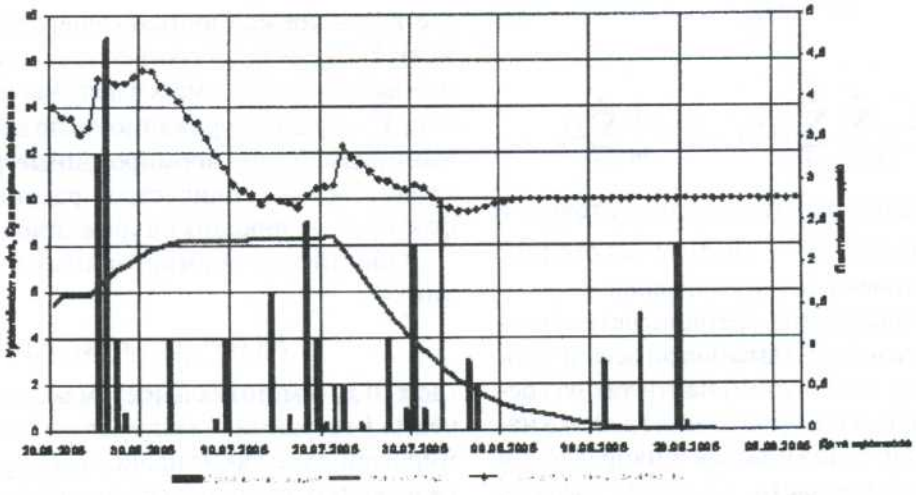


Рис. 3. Динамика прогнозируемой урожайности, листового индекса и осадков для метеостанции Здвинск Новосибирской области

Идентификация технологии оценки урожайности по эмпирической модели с использованием данных года-аналога была произведена на основе следующих агроклиматических показателей по Алтайскому краю [4]: суммы эффективных температур за вегетационный период и за первые две декады вегетационного периода; суммы осадков за вегетационный период и за первые две декады вегетационного периода; количества дней с осадками за вегетационный период и за первые две декады вегетационного периода; суммы осадков за зимний, весенний и осенний периоды, урожайность пшеницы.

Рассмотрим технологию определения года-аналога и оценку урожайности яровой пшеницы для конкретного объекта – 1997 года (считаем, что погодный сценарий данного года известен на момент определения года-аналога).

По мнению специалистов, огромное влияние на формирование будущего урожая оказывают первые две декады вегетационного периода. Поэтому в работе особое внимание уделено исследованию именно этого периода.

Классификацию осуществляли по фактическим данным первых двух декад вегетационного периода 1997 года. Процедуру кластеризации проводили несколько раз с помощью программы статистической обработки данных Statistica 6.0 при различных значениях числа кластеров. Затем выбиралась лучшая группировка в смысле критерия минимума отношений средних внутрикластерных (J_1) и межкластерных (J_2) расстояний

$$J = J_1/J_2, \qquad (1)$$

где

$$J_1 = \frac{2}{m(m-1)}\sum_{i=1}^{m}\sum_{j=i+1}^{m}\bar{D}_{ij}, \quad J_2 = \frac{1}{m}\sum_{i=1}^{m}\bar{D}_{ii}.$$

Здесь $\bar{D}_{ii}$ – среднее расстояние между точками внутри i-го класса; $\bar{D}_{ij}$ – среднее расстояние между парами точек i-го и j-го классов.

Для графической интерпретации результатов используется критерий "каменистой осыпи", согласно которому необходимо найти такую абсциссу на графике, в которой линия графика начинает визуально сглаживаться в направлении правой, пологой его части. Согласно графику лучший вариант – разбиение на шесть кластеров (рис. 4).



Рис. 4. Результаты классификации по первым двум декадам вегетационного периода

Анализ проведенной классификации позволяет предположить, что для 1997-го годами-аналогами будут 1973, 1976, 1981, 1988. Урожайность 1997 года по фактическим данным этих лет будет находиться в пределах 13,4–17,5 ц/га. По погодному сценарию наиболее близкие годы – 1976, 1981, 1988. Следовательно, интервал для ожидаемой урожайности можно сузить: 13,4–15,9 ц/га.

На следующем этапе осуществлялся прогноз урожайности по эмпирической модели [4]:

$$Y_{j+1} = \begin{cases} Y_{\min} + (Y_j - Y_{\min})\prod_{i=1}^{k} H_i(x), \text{если } Y_j \ge \bar{Y}; \\ Y_{\min} + (Y_{\max} - Y_j)\prod_{i=1}^{k} H_i(x), \text{если } Y_j < \bar{Y}, \end{cases} \qquad (2)$$

где Y_{j+1} – урожайность текущего года (ожидаемая); Y_j, $Y_{\max}$ и $Y_{\min}$ – соответственно урожайность предыдущего года, максимальная и минимальная; $\bar{Y}$ – средняя урожайность по всему временному ряду; $H_i(x)$ – нормированные функции отклика; k – количество рассматриваемых факторов, влияющих на урожайность.

Общий вид нормированных функций отклика

$$H_i(x) = f_i(x_1, x_2, x_3, x_4),$$

где x_1 и x_4 – минимальное и максимальное значения i-го фактора, за пределами которых формирование урожая не происходит; x_2, x_3 – биологически оптимальные границы i-го фактора.

Идентификация структуры и параметров модели (2) проводилась путем спланированного

численного эксперимента, который включал подбор и корректировку структуры уравнений модели с целью минимизации отклонений расчетной урожайности от наблюдаемой. Значения параметров в уравнениях подбирались с помощью оптимизационных процедур.

В качестве параметров модели за первые две декады вегетационного периода, для которых осуществлялась процедура идентификации, выступали: пороговые значения суммы температур при переходе через 5 °С и пороговые значения суммы осадков.

Процент отклонения расчетных данных от фактических по всему временному ряду в среднем составил 20 %. Существенность различий в средней фактической урожайности и расчетной оценивали по критерию Стьюдента: $t_{рас} = 2{,}24$; $t_{кр} = 2{,}403$. При уровне значимости $\alpha = 0{,}02$ $t_{рас} < t_{кр}$. Следовательно, статистическая проверка показала адекватность модели (2) и что фактическая и прогнозируемая урожайность не отличаются статистически значимо.

Результаты численных расчетов урожайности по модели (2) выбранных годов-аналогов представлены в таблице.

Фактическая и расчетная величина урожайности годов-аналогов, ц/га

Год	Фактическая урожайность года-аналога	Расчетная урожайность по модели (3)
1976	13,7	11,4
1981	13,4	12,4
1988	15,9	12,9
1997	12,8	

Тогда урожайность 1997 года по расчетным данным аналогов будет находиться в пределах 11,4–12,9 ц/га, средняя урожайность для 1997 года составит 12,2 ц/га; по фактическим данным аналогов – в пределах 13,4–15,9 ц/га, средняя урожайность для 1997 года 14,6 ц/га (фактическая урожайность – 12,8 ц/га).

Сравнение приведенных результатов показывает, что оценка средней урожайности по расчетным данным аналогов дает ошибку 0,6 ц/га, по фактическим данным аналогов – 1,8 ц/га.

Аналогичная процедура проводилась по метеорологическим данным годов-аналогов за весь период вегетации. Сравнительный анализ различных разбиений и расчетов по модели (2) за первые две декады вегетационного периода показал, что годом-аналогом для 1997-го является 1981-й. Величина урожая по погодному сценарию 1981 года (фактическая) составит 13,4 ц/га, по модельным расчетам – 12,4 ц/га. Ошибка прогноза по модельным расчетам – 0,4 ц/га, по фактической урожайности 1981 года – 0,6 ц/га.

Итак, разработана технология прогноза урожайности зерновых культур, реализованная в имитационно-моделирующем комплексе программных средств и апробированная в условиях зерновых провинций Западной Сибири. Показано, что разработанный метод прогнозирования урожайности зерновых культур позволяет определять в течение вегетационного периода основные параметры биопродуктивности зерновых культур, давать заблаговременную (за месяц до уборки) оценку урожайности с точностью до 1 ц/га и определять сроки начала уборки зерновых культур, используя модель биопродуктивности, ежедневные данные метеостанций и данные космического мониторинга.

Анализ результатов применения технологии определения года-аналога показал хорошее совпадение фактической урожайности с расчетной по данным двух декад вегетационного периода. Это говорит, во-первых, о том, что действительно первые две декады определяют в значительной степени будущий урожай; во-вторых, что разработанная авторами эмпирическая модель может успешно применяться для текущего планирования и перспективного прогнозирования урожайности яровой пшеницы на территории Алтайского края.

СПИСОК ЛИТЕРАТУРЫ

1. **Брыксин В.М.** Применение адаптированной модели биопродуктивности EPIC и космоснимков MODIS для прогнозирования урожайности зерновых культур на территории Западной Сибири // Вестник НГУ. Сер. Информационные технологии. Новосибирск: Изд-во НГУ, 2007. Т. 5. Вып. 1. С. 20–26.

2. **Матвеев Л.Т.** Курс общей метеорологии. Физика атмосферы. Л.: Гидрометеоиздат, 1984.

3. **Хворова Л.А., Брыксин В.М, Масленников Д.В.** Оценка возможности использования программного комплекса EPIC для прогноза урожайности в Алтайском крае // Региональные проблемы развития Сибири и Дальнего Востока: IV Сибирский конгресс по прикладной математике ИНПРИМ-2000 / Ин-т математики. Новосибирск, 2000. С. 112.

4. **Хворова Л.А., Гавриловская Н.В.** Прогнозирование урожайности зерновых культур: методы и расчеты // Изв. АлтГУ. 2008. № 1 (57). С. 65–68.

5. **Хворова Л.А., Гавриловская Н.В.** Разработка алгоритма определения года-аналога для оценки урожайности зерновых культур в условиях Алтайского края // Изв. АлтГУ. 2007. № 1 (53). С. 66–67.

6. **Williams J.R.** The EPIC Model / U.S. Department of Agriculture, Agricultural Research Service, Grassland, Soil and Water Research Laboratory. Temple, 1984.

УДК 004.771

Ю.Я. Болдырев, В.А. Варгаузин, Е.П. Петухов, И.А. Цикин

ТЕХНОЛОГИЯ СЕТЕВОГО ДОСТУПА К ПРОГРАММНОЙ СРЕДЕ MATLAB НА ОСНОВЕ КЛАСТЕРНОЙ ВЫЧИСЛИТЕЛЬНОЙ СИСТЕМЫ

Современные сетевые информационные технологии позволяют реализовывать математические модели, требующие больших вычислительных ресурсов, без необходимости наличия мощного компьютера у конечного пользователя. При таком подходе разработанная модель может быть централизованно размещена на специализированном сервере, а на компьютере пользователя имеется лишь клиентское программное обеспечение, в задачи которого входит организация взаимодействия с моделью посредством визуального интерфейса. При этом может быть обеспечен одновременный доступ большого числа пользователей.

Одной из наиболее широко распространенных сегодня специализированных сред математического моделирования является MATLAB. Эта среда ориентирована на использование возможностей параллельных вычислений [1], которые в наиболее полной мере реализуются в кластерных многоядерных вычислительных системах (многоядерный компьютер с архитектурой SMP – Symmetric Multiprocessing или кластерная система типа MPP – Massively Parallel Processing) [2]. Примеры эффективного использования MATLAB на кластерной вычислительной системе можно найти в [3, 4].

Тем не менее существует целый ряд особенностей организации взаимодействия MATLAB с кластерной вычислительной системой, включая вопросы организации коллективного доступа, выбора типа планировщика для динамического распределения вычислительного ресурса, использования приложений MATLAB для проверки работоспособности кластерной вычислительной системы и т. п. Наряду с этим весьма важен механизм использования лицензий в разных режимах работы кластера.

Особенности функционирования MATLAB на кластерной вычислительной системе

Предварительно отметим, что кластерные вычислительные системы – это многопроцессорные комплексы, состоящие из так называемых узлов, каждый из которых состоит из двух процессоров и может, например, иметь свою (узловую) память. При этом следует иметь в виду, что каждый из современных процессоров состоит из нескольких "процессоров-ядер". Наиболее распространенные сегодня процессоры – это 4-ядерные системы AMD Barcelona и Intel Harpertown.

Пользователи взаимодействуют с кластерной системой в режиме удаленного доступа своих персональных компьютеров к одному из узлов кластерной системы, выделенному как "головной узел". При этом подразумевается, что пользователи не располагают средой MATLAB на своих компьютерах, а работают с ней в терминальном режиме (для ОС Windows – режим "удаленный рабочий стол" [5]).

При этом сам программный комплекс MATLAB устанавливается либо только на головном узле, либо на всех компьютерах кластерной системы.

Комплекс MATLAB включает средства разработки программ (язык MATLAB и среда моделирования Simulink) и большое число библиотек функций и моделей (Toolbox и Blocksets) для различных приложений. Особое место занимает библиотека программных средств PCT – Parallel Computing Toolbox (в ранних версиях MATLAB – Distributed Computing Toolbox), предоставляющая

пользователям программные средства повышения скорости решения задач с использованием любых остальных средств MATLAB.

Для управления ресурсами кластерной вычислительной системы MATLAB включает программную компоненту MDCS – MATLAB Distributed Computing Server (в ранних версиях MATLAB – Distributed Computing Engine). При этом MDCS может использовать разные планировщики заданий (встроенный в MATLAB планировщик Job Manager фирмы MathWorks, сторонний планировщик Windows из состава ОС Microsoft Windows Compute Cluster Server (WCCS) и т. п.).

Поскольку планировщики отличаются принципом управления вычислительными процессами, получившими в терминологии MATLAB термин "workers" (начиная с версии MATLAB R2009a они ассоциируются с процессорными ядрами), то процедура установки MATLAB зависит от выбора планировщика. При использовании стороннего планировщика требуемый вычислительный ресурс предоставляется только на время решения задачи. При этом достаточно одной установленной копии MATLAB. Например, при использовании планировщика задач из состава WCCS установка MATLAB производится только на головной узел, причем должен быть обеспечен доступ к установленной среде для любого из компьютеров кластерной вычислительной системы. Последнее достигается установкой MATLAB на сетевой диск головного узла кластера. Пример такой конфигурации представлен на рис. 1.

В случае использования встроенного в MATLAB планировщика Job Manager вычислительный ресурс резервируется и в случае, когда решение задачи не происходит. Для обеспечения работоспособности такого планировщика требуется установка комплекса MATLAB на каждый узел кластерной вычислительной системы.

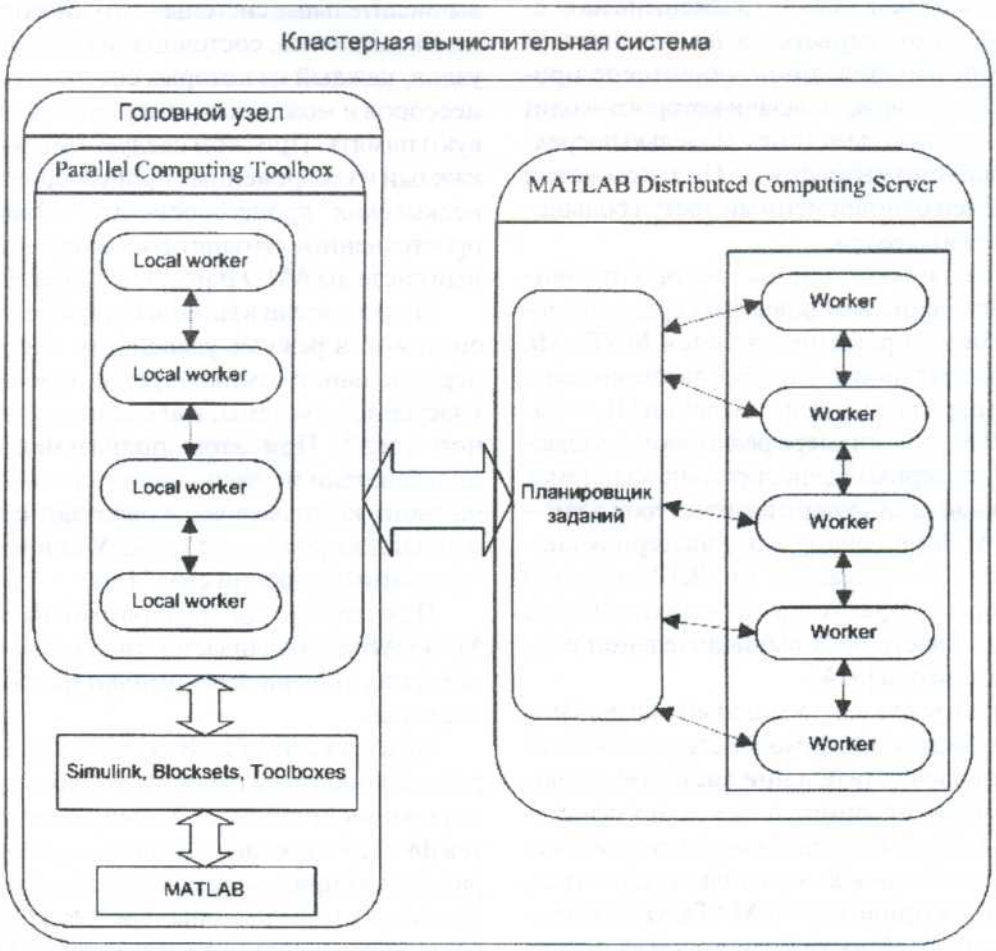


Рис. 1. Функционирование MATLAB на кластерной вычислительной системе

Выбор планировщика основывается на требованиях к его возможностям. Так, для организации многопользовательской работы особо важно количество одновременно решаемых задач. Встроенный планировщик Job Manager в версии R2009a не позволяет организовать систему очередей при обработке заданий: задача пользователя может быть решена только при наличии свободных ресурсов на момент ее постановки. В то же время сторонние планировщики изначально ориентированы на работу с очередью задач.

Важно отметить, что выбор планировщика также влияет на порядок использования лицензий в разных режимах работы кластера. В случае использования планировщика Job Manager оказываются постоянно задействованными лицензии "MATLAB Distributed Computation Engine". При использовании же стороннего планировщика лицензии будут запрошены только на время решения задач.

С другой стороны, при использовании планировщика Job Manager дополнительное время на запуск задач не требуется. Для небольших и быстро решаемых задач подобная конфигурация выглядит более предпочтительной, так как накладные расходы на запуск вычислительного процесса сторонним планировщиком могут быть существенными.

Пользовательский доступ к кластерной вычислительной системе

Пользователь в режиме удаленного доступа может располагать вычислительным ресурсом как исключительно одного головного узла (локальная конфигурация PCT на основе "local" на рис. 1), так и всей кластерной вычислительной системы (кластерная конфигурация PCT на основе "workers" на рис. 1).

В локальной конфигурации число доступных процессорных ядер и соответственно процессов (local workers) головного узла на момент выхода версии MATLAB R2009a не превышает четырёх.

Для доступа к большему ресурсу кластерной вычислительной системы требуется лишь перенастроить PCT с локальной конфигурации на конфигурацию такой системы. Для примера на рис. 2 такая конфигурация имеет имя "ccsconfig".

В результате для пользователя различие между локальной и кластерной конфигурациями PCT сводится лишь к различию числа доступных workers. Поэтому в терминологии MATLAB любая группа workers (включая локальную конфигурацию PCT) называется "кластером" (не путать с кластерной вычислительной системой на рис. 1) и не связывается с физической природой workers. Количество workers "кластера" в терминологии MATLAB называется размером "кластера" (ClusterSize). Пользователь MATLAB может всегда получить доступный ему через PCT ресурс с помощью команды `findResource`. Например, "кластер" с конфигурацией "ccsconfig" имеет размер 16 (рис. 3).

Специфика использования PCT

В терминологии MATLAB эффект ускорения вычислений при решении задачи группой workers интерпретируется как некий невидимый для

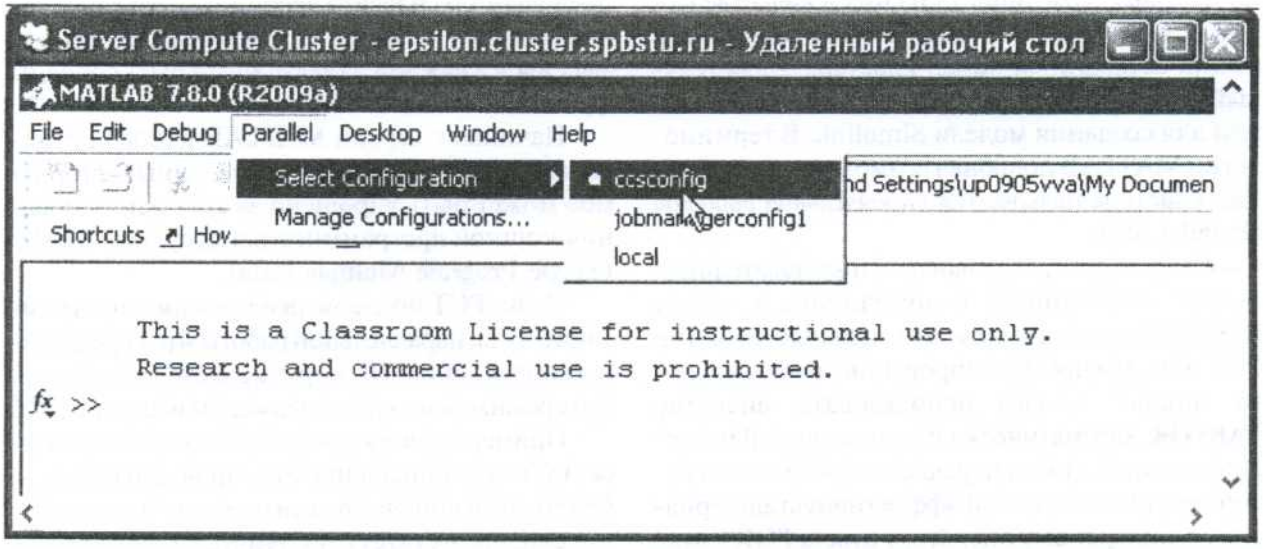


Рис. 2. Выбор конфигурации PCT для кластерной вычислительной системы

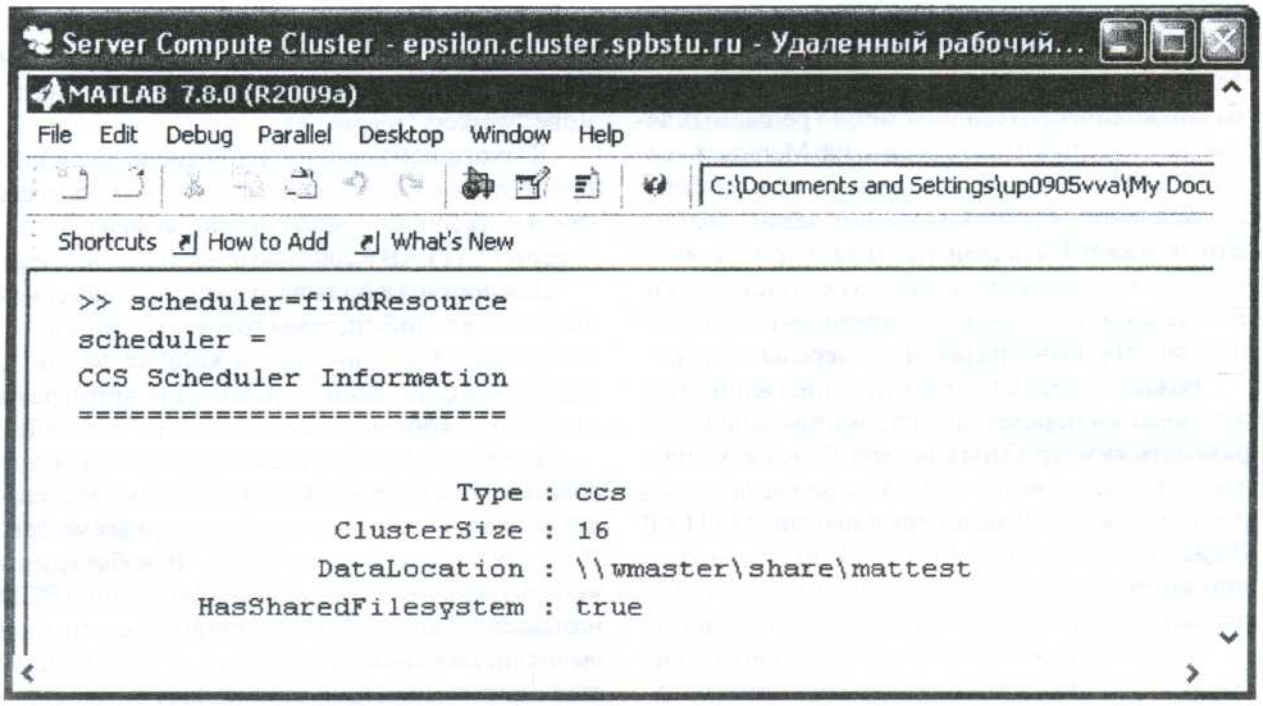


Рис. 3. Доступный ресурс в конфигурации PCT с кластерной системой

пользователя переход от одного MATLAB к нескольким (MATLABS). Группа этих workers называется *программным пулом*, а сами workers программного пула – "labs" как противопоставление слову "lab" в единственном числе. Размер пула не превосходит размера "кластера". Активизация программного пула производится командой **matlabpool**.

К примеру, имея "кластер" из 16 workers, можно использовать лишь 15 из них в качестве labs программного пула для исполнения программы, а один оставшийся worker кластера – для текущей интерактивной работы в редакторе MATLAB или для создания модели Simulink. В терминологии MATLAB подобная возможность разделения кластера называется *параллельной работой* (Parallel Job).

Открыв программный пул, пользователь получает возможность в дополнение к языку MATLAB задействовать специальные функции PCT. Так, для программирования "параллельного цикла" можно использовать оператор **PARFOR**, автоматически превращающий последовательный цикл в параллельный. Этот оператор представляет собой эффективную альтернативу стандартному оператору цикла **FOR** языка MATLAB. Для организации распределённых вычислений в многоядерной компьютерной среде (Disributed Job) в PCT имеется около 150 функций (алгоритмов), позволяющих распределять как вычисления, так и требуемый объём памяти по вычислительной кластерной системе. Эти функции в свою очередь могут быть вызваны функцией распределённых вычислений **dfeval** (Distributed FEVAL), работающей с распределенными массивами. Обработка каждой части распределённого массива ассоциируется с отдельной задачей. Результаты обработки отдельных массивов в каждой из задач могут быть собраны функцией конкатенации **cat**.

Начиная с версии MATLAB R2008b нетривиальное программирование подобных алгоритмов может быть упрощено за счёт использования мощной программной конструкции **SPMD (Single Program Multiple Data)**.

Итак, PCT поддерживает режимы программного пула, параллельной работы и распределённых вычислений. Проверить работоспособность этих режимов можно менеджером настроек PCT.

Пример проверки работоспособности (статуса) PCT с планировщиком ccs приведён на рис. 4. Статус этой конфигурации отмечен "галочкой", что свидетельствует о доступности этого ресурса (Find Resource Succeeded) и возможности ис-

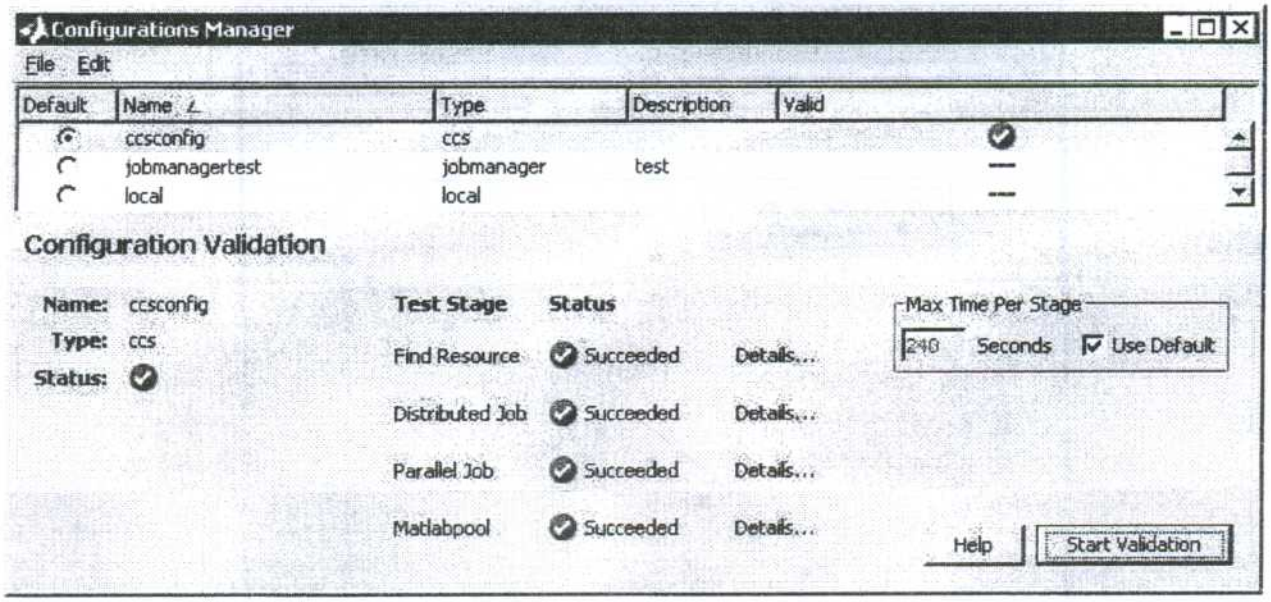


Рис. 4. Результат проверки статуса конфигурации PCT для кластерной системы с планировщиком ccs

пользования распределённых вычислений (Distributed Job Succeeded), параллельной работы (Parallel Job Succeeded) и программного пула (Matlabpool Succeeded).

Тестирование

Ниже рассмотрены результаты двух тестов, проведённых на 12-ядерной кластерной системе. Первый тест иллюстрирует эффективность использования оператора параллельного цикла **PARFOR**, второй – программной конструкции **SPMD**. Сущность основного результата тестов проста: эффективность использования PCT почти линейно растёт с ростом числа используемых workers в программном пуле.

Тест оценки эффективности оператора PARFOR. Для оценки эффективности оператора **PARFOR** в MATLAB был использован тест, основанный на компьютерном варианте карточной игры Blackjack с любым числом игроков, каждый из которых имеет большое число "рук". Алгоритм игры основан на методе статистического моделирования Монте-Карло. Тест может быть запущен либо из командной строки MATLAB командой `paralleldemo_parfor_bench`, либо из демонстрационной системы MATLAB.

На рис. 5 показана зависимость степени повышения скорости вычислений от числа workers в программном пуле при локальной конфигурации PCT в одной из ста игр, а на рис. 6 представлено процентное соотношение ситуаций выигрыша, который удалось получить при решении тестового набора из ста задач при использовании четырех вычислительных потоков.

Тест оценки эффективности программной конструкции SPMD. Хорошим примером использования программной конструкции **SPMD** является тест оценки эффективности алгоритмов векторных и матричных вычислений. Тест может быть запущен либо из командной строки MATLAB командой `paralleldemo_resource_bench`, либо из демонстрационной системы MATLAB. Этот тест впервые появился в версии MATLAB R2009a.

Эффект повышения скорости вычислений в зависимости от числа workers иллюстрирует рис. 7. В тесте используются алгоритмы быстрого преобразования Фурье (fft) и матричного умножения (matrix mult).

На рис. 8 приведена зависимость степени ускорения векторных и матричных вычислений от размера вектора (квадратной матрицы) для локальной конфигурации PCT. Используются алгоритмы быстрого преобразования Фурье (vector fft), матричного умножения (matrix mult), обращения матрицы методом треугольной декомпозиции (matrix LU), приведения матриц к диагональному виду (matrix SVD) и нахождения собственных векторов и чисел (matrix eig).

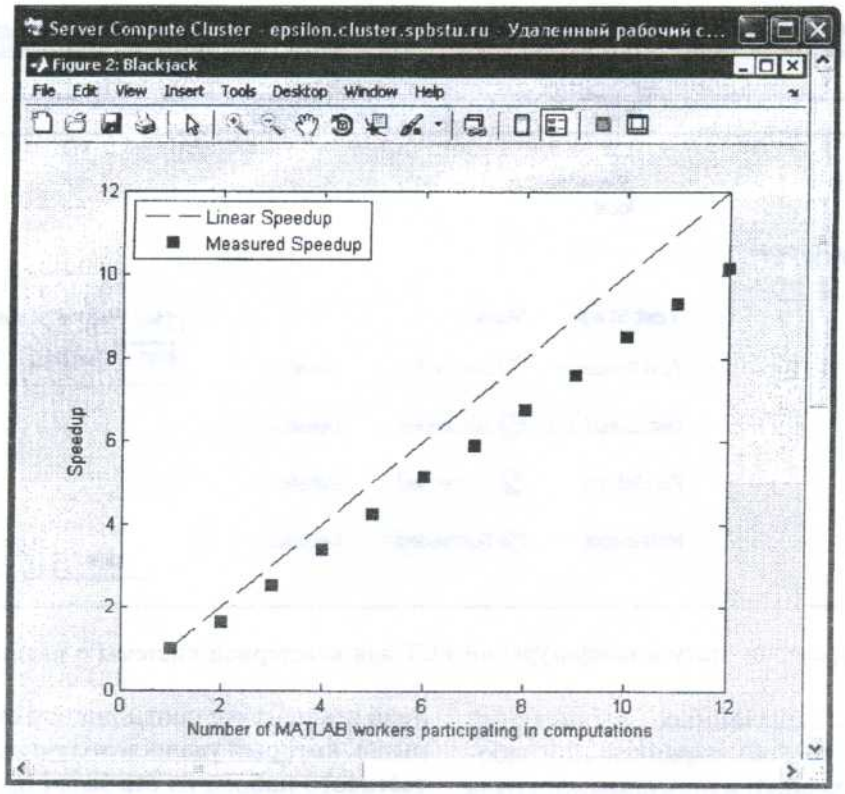


Рис. 5. Зависимость степени повышения скорости вычислений от числа workers в программном пуле

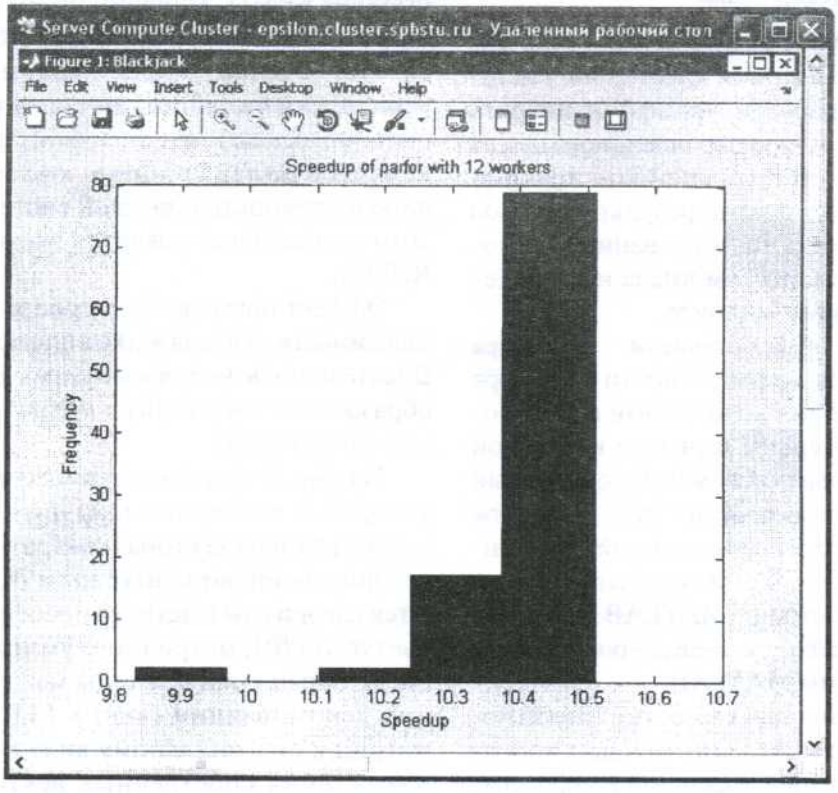


Рис. 6. Процентное соотношение выигрышных ситуаций

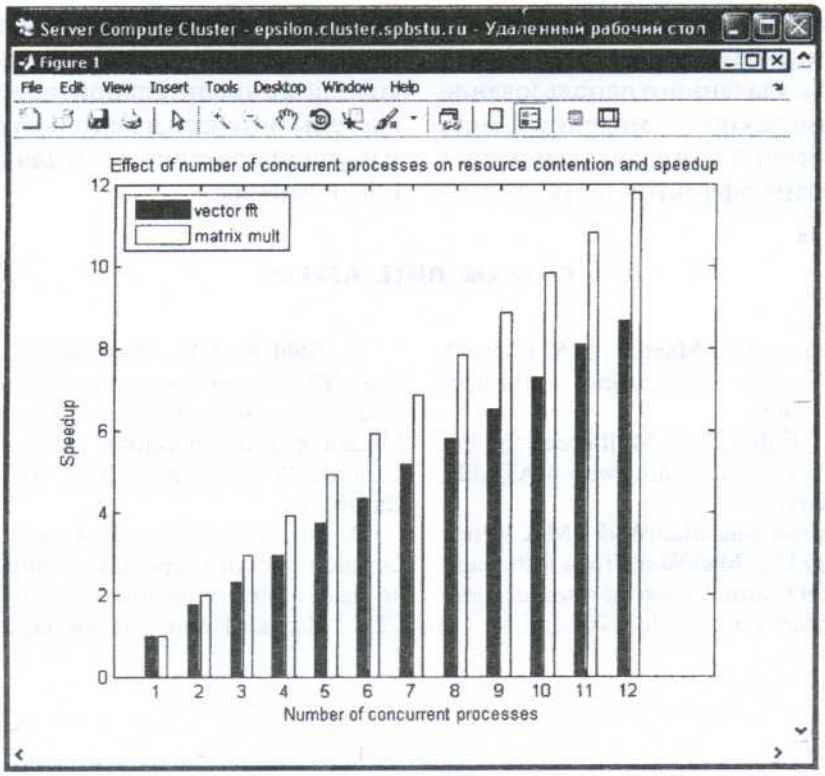


Рис. 7. Зависимость степени повышения скорости вычислений от числа workers в программном пуле

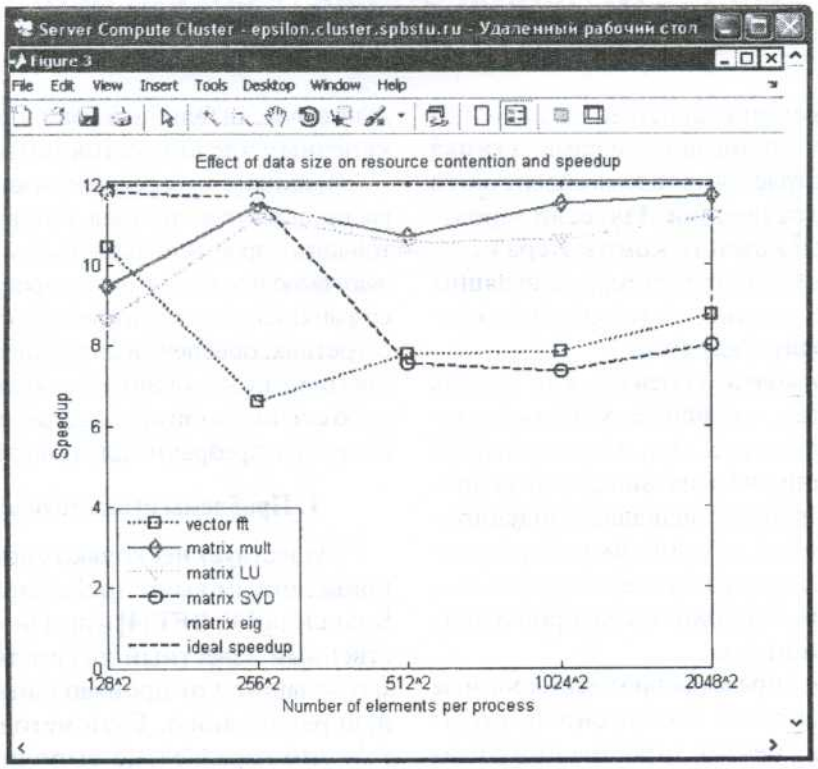


Рис. 8. Зависимость степени ускорения векторных и матричных вычислений от размера вектора (квадратной матрицы)

В итоге отметим, что рассмотренный программно-аппаратный комплекс продемонстрировал возможности удаленного использования среды математического моделирования MATLAB на кластерной вычислительной системе. Также показана эффективность средств MATLAB для создания параллельных программ, использующих ресурсы многопроцессорной вычислительной техники. Подобные системы интересны не только для инженерной и научной практики, но и для целей электронного обучения.

СПИСОК ЛИТЕРАТУРЫ

1. **Gaurav Sharma, Jos Martin.** MATLAB: A Language for Parallel Computing: Springer (Springerlink.com), 2008.
2. **Luszczek P.** Enhancing Multicore System Performance Using Parallel Computing with MATLAB. MATLAB Digest, 2009.
3. Пресс-релиз компании MathWork "Max Planck Institute of Biochemistry Uses MathWorks Tools in its Quest to Cure Cancer". 2009. http://www.mathworks.com/company/pressroom/articles/article31269.html
4. **Ghidella J.R., Wakefield A., Grad-Freilich S.** et al. The Use of Computing Clusters and Automatic Code Generation to Speed Up Simulation Tasks: Modeling and Simulation Technologies Conf. and Exhibit. 20–23 August 2007. Hilton Head, South Carolina.
5. **Варгаузин В.А., Иванов Д.И., Цикин И.А.** Сетевой доступ к среде моделирования MATLAB на базе служб терминала: Учеб.-метод. пособие. СПб.: Изд-во Политехн. ун-та, 2008.

УДК: 004.75

А.Н. Фирсов

ОЦЕНКА НАДЕЖНОСТИ АСИНХРОННЫХ РАСПРЕДЕЛЕННЫХ СИСТЕМ, УСТОЙЧИВЫХ К ПРОИЗВОЛЬНЫМ ОТКАЗАМ

С течением времени зависимость людей от техники только увеличивается, а сама техника усложняется. Обостряется проблема обеспечения отказоустойчивости техники. Так, если вероятность сбоя отдельно взятого компьютера невелика, то в распределенных системах, состоящих из тысяч или даже миллионов таких компьютеров, сбои происходят постоянно.

Византийская модель сбоев [1], или модель произвольных сбоев, – наиболее общая и включает в себя все остальные. Она не накладывает никаких ограничений на действия сбойных процессов и благодаря этому позволяет моделировать такие типы сбоев, как поломка оборудования, потеря связи, неисправность оборудования, программные сбои, ошибки операторов и действия злоумышленников.

Особый интерес представляют асинхронные распределенные системы, так как они не накладывают никаких временных ограничений и благодаря этому идеально подходят для работы в сети Интернет. Наиболее универсальной моделью вычислительного узла является машина с конечным числом состояний (state machine).

В данной статье рассматриваются асинхронные распределенные системы, которые, во-первых, устойчивы к произвольным отказам; во-вторых, позволяют выполнять прикладные программы, требующие сохранять свое состояние между обращениями к ним; в-третьих, обеспечивают отказоустойчивость на программном уровне без необходимости в какой-либо специальной аппаратуре; в-четвертых, делают все это прозрачно для прикладных программ.

1. Проблемы существующих решений

Существует несколько систем, отвечающих приведенным выше требованиям: Rampart [2], SecureRing [3], BFT [4] и др. Они реализуют единственный известный на сегодня программный метод защиты от произвольных сбоев – активную репликацию. Суть метода заключается в том, что параллельно выполняется несколько копий (или, как их еще называют, реплик) одной и той же программы.

В вышеуказанных системах и практически во всех других работах, связанных с устойчивостью к произвольным отказам [1–8], фиксируется некоторое число f – максимальное количество процессоров (или вычислительных узлов), которое может выйти из строя из имеющихся n процессоров, не приведя к сбою системы в целом, и уже на его основе доказываются все утверждения. Это удобно для проектирования и доказательства корректности алгоритмов, однако в действительности не может быть никакой гарантии, что из строя не выйдет большее число процессоров, просто такая вероятность считается пренебрежимо малой и не учитывается.

Ни одна из упомянутых выше систем не предоставляет алгоритма для оценки вероятности сбоя получаемого решения, а именно в этом показателе больше всего заинтересован пользователь отказоустойчивых систем. Как правило, последнему предоставляется лишь возможность управлять необходимым уровнем надежности, изменяя параметр f: увеличивать для повышения надежности или уменьшать, экономя ресурсы. Такой подход в неявном виде предполагает, что все вычислительные узлы одинаковы по надежности. В сети же Интернет и распределенных системах, строящихся на ее основе, это требование, как правило, не выполняется. В одну распределенную систему могут быть объединены как серверы с устройствами бесперебойного питания, находящиеся под управлением профессиональных администраторов, с надежными и обновляемыми операционными системами и антивирусным программным обеспечением, так и рабочие станции, зараженные вирусами, с ненадежной аппаратурой и уязвимой операционной системой. Очевидно, что вероятность их отказа может отличаться на несколько порядков. А именно на использование в сети Интернет и в асинхронных распределенных системах ориентируются вышеперечисленные системы обеспечения отказоустойчивости.

Пренебрежение тем фактом, что вычислительные узлы иногда значительно отличаются по надежности, может повлечь ситуацию, когда увеличение параметра f приведет к уменьшению надежности системы в целом. Это продемонстрировано в разд. 5.

2. Общая схема предлагаемого решения

Очевидно, что на надежность отказоустойчивой системы в целом влияет надежность ее составных частей, и если надежности этих составных частей различаются, то необходимо знать каждую из них в отличие от случая, когда они одинаковы, – тогда достаточен один агрегированный параметр.

Наиболее удобным на практике способом получения значений надежности вычислительных узлов является сбор статистики об их отказах в течение достаточно длительного промежутка времени. Этому вопросу посвящен разд. 3.

В разд. 4 выводится формула для нахождения вероятности отказа системы в целом на основе данных о вероятности отказа ее компонент, а также предлагается эффективный полиномиальный алгоритм вычисления этой формулы.

В разд. 5 на основании полученной формулы оценивается надежность существующих решений в некоторых ситуациях, когда компоненты системы неоднородны по надежности. Способы повышения надежности в таких ситуациях предлагаются в разд. 6.

3. Оценка надежности вычислительных узлов на основе собираемых статистических данных

Обозначим вероятность возникновения сбоя на i-м вычислительном узле в некоторый единичный промежуток времени как β_i $(i = \overline{1,n})$. С целью упрощения получаемых формул этот единичный промежуток выбирается равным времени, которое необходимо отказоустойчивой системе для ликвидации последствий отказа. Его значение зависит от используемых алгоритмов обеспечения отказоустойчивости и восстановления, а также от времени, требующегося на передачу сообщений.

В реальных распределенных системах значение β_i может находиться на основе собираемой статистики о сбоях и следующей формулы:

$$\beta_i = \frac{1 + f_i}{\frac{1}{\beta} + t_i}, \qquad (1)$$

где f_i – количество зафиксированных сбоев на i-м вычислительном узле за время его работы в составе распределенной системы; t_i – время работы i-го вычислительного узла в составе распределенной системы; β – вероятность отказа вычислительного узла, о котором еще не собрано никакой статистики.

Значение β можно находить следующим образом:

$$\beta = \frac{1 + \sum_{i=1}^{m} f_i}{\frac{1}{\overline{\beta}} + \sum_{i=1}^{m} t_i}. \qquad (2)$$

Здесь m – общее количество вычислительных узлов в распределенной системе (не следует путать с n – количеством вычислительных узлов в группе, выделенной специально для решения конкретной задачи); $\overline{\beta}$ – задаваемая создателем распределенной системы на основе его опыта среднестатистическая вероятность отказа вычислительного узла, которая необходима в начальный момент работы распределенной системы, когда отсутствует любая статистическая информация.

Формула (1) обладает тем свойством, что при первоначальном подключении вычислительного узла к распределенной системе, когда о нем еще не собрано никакой статистики, значение β_i будет равно β-вероятности отказа среднестатистического вычислительного узла. При сборе же статистической информации в течение достаточно долгого времени вклад слагаемого $1/\beta$ в знаменатель и единицы в числитель окажется пренебрежимо мал: $\beta_i = \lim_{t_i \to \infty} \frac{1 + f_i}{\frac{1}{\beta} + t_i} = \lim_{t_i \to \infty} \frac{f_i}{t_i}$.

Аналогичное замечание относится и к формуле (2). В момент начала работы распределенной системы, когда не собрано никакой статистики, вероятность отказа среднестатистического вычислительного узла может определяться на основе опыта создателя системы. При накоплении статистики вклад этого слагаемого будет неуклонно уменьшаться и заменяться эмпирическими данными, полученными самой распределенной системой.

4. Формула и эффективный алгоритм вычисления вероятности сбоя отказоустойчивой системы

Имея данные о вероятности отказа каждого вычислительного узла в отдельности, можно рассчитать вероятность сбоя отказоустойчивого решения, объединяющего несколько вычислительных узлов.

Будем считать, что отказы вычислительных узлов – независимые события, ибо именно к этому стремится подсистема балансировки при создании групп и распределении нагрузки. Тогда вероятность того, что в наборе из n вычислительных узлов произойдет ровно k отказов, равна:

$$\alpha_n^k = \sum_{F \subseteq U, |F| = k} \left(\prod_{i \in F} \beta_i \right) \left(\prod_{j \in U \setminus F} (1 - \beta_j) \right), \qquad (3)$$

где F – неупорядоченное подмножество из множества всех вычислительных узлов U.

Смысл формулы (3) прост: суммируются вероятности возникновения всех возможных комбинаций из k выходящих из строя вычислительных узлов и $n-k$ узлов, продолжающих корректно работать.

Так как все приведенные выше отказоустойчивые системы гарантируют корректную работу лишь при условии, что не более чем $f = \left\lfloor \frac{n-1}{3} \right\rfloor$ вычислительных узлов выйдут из строя, то при возникновении более чем f сбоев вся система целиком может дать сбой. В [5] доказано, что ни один алгоритм не может обеспечить корректную работу системы из n вычислительных узлов более чем при $\left\lfloor \frac{n-1}{3} \right\rfloor$ отказах. Там же есть пример, когда $\left\lfloor \frac{n-1}{3} \right\rfloor + 1$ отказов приводят к сбою системы целиком. А так как рассматривается Византийская модель сбоев, когда вышедшие из строя узлы могут находиться под согласованным управлением злоумышленников, то необходимо рассматривать худший случай и считать, что при отказе более f узлов вся система дает сбой. Вероятность этого можно рассчитать:

$$\alpha_n = \sum_{k=\left\lfloor \frac{n-1}{3} \right\rfloor + 1}^{n} \alpha_n^k = \sum_{k=\left\lfloor \frac{n-1}{3} \right\rfloor + 1}^{n} \sum_{F \subseteq U, |F| = k} \left(\prod_{i \in F} \beta_i \right) \left(\prod_{j \in U \setminus F} (1 - \beta_j) \right). \qquad (4)$$

Воспользовавшись тем, что

$$\sum_{k=0}^{n} \sum_{F \subseteq U, |F| = k} \left(\prod_{i \in F} \beta_i \right) \left(\prod_{j \in U \setminus F} (1 - \beta_j) \right) = 1,$$

т. е. сумма вероятностей того, что откажет от нуля до n узлов, равна единице, так как это событие покрывает все возможные исходы, получим формулу, требующую суммирования меньшего числа слагаемых:

$$\alpha = 1 - \sum_{k=0}^{\left\lfloor \frac{n-1}{3} \right\rfloor} \sum_{F \subseteq U, |F| = k} \left(\prod_{i \in F} \beta_i \right) \left(\prod_{j \in U \setminus F} (1 - \beta_j) \right). \qquad (5)$$

Выражение

$$\sum_{k=0}^{\left\lfloor \frac{n-1}{3} \right\rfloor} \sum_{F \subseteq U, |F| = k} \left(\prod_{i \in F} \beta_i \right) \left(\prod_{j \in U \setminus F} (1 - \beta_j) \right) \qquad (6)$$

представляет вероятность того, что из строя выйдет не более чем $\left\lfloor \frac{n-1}{3} \right\rfloor$ вычислительных узлов (т. е. существующие отказоустойчивые решения могут гарантировать корректную работу системы в целом), поэтому его можно назвать надежностью полученного решения.

В случае, если вероятности отказа всех узлов примерно одинаковы и равны β, формулы (4) и (5) можно упростить, воспользовавшись тем, что

$$\left(\prod_{i\in F}\beta_i\right)\left(\prod_{j\in U\setminus F}(1-\beta_j)\right)=\beta^{|F|}(1-\beta)^{n-|F|}$$

и количество всевозможных комбинаций выбора k неупорядоченных элементов из множества мощностью n равно C_n^k:

$$\alpha=1-\sum_{k=0}^{\left\lfloor\frac{n-1}{3}\right\rfloor}C_n^k\beta^k(1-\beta)^{n-k}. \qquad (7)$$

Выражение (7) есть аналог формулы (5) для случая, когда вероятности сбоя вычислительных узлов примерно одинаковы. Хотя (7) и является частным случаем (5), но с практической точки зрения удобна тем, что требуется гораздо меньше операций для ее вычисления.

Применение формулы (5) "в лоб" на практике крайне нежелательно, ибо она требует перебора всех возможных подмножеств мощностью от нуля до $\left\lfloor \frac{n-1}{3} \right\rfloor$ множества из n элементов, а это в свою очередь требует порядка $\sum_{k=0}^{\left\lfloor\frac{n-1}{3}\right\rfloor}C_n^k$ операций. Кроме того, полученные подмножества необходимо еще перемножать, тогда получаем $n\sum_{k=0}^{\left\lfloor\frac{n-1}{3}\right\rfloor}C_n^k$ операций. А это выражение растет экспоненциально с ростом n.

Для получения эффективного способа вычисления общей надежности вернемся к формуле (3) и разобьем сумму, входящую в ее состав, на две части: первую, где n-й вычислительный узел входит во множество отказавших узлов, и вторую, где он продолжает работать без сбоев. Получим:

$$\alpha_n^k=\sum_{F\subseteq U,|F|=k,n\in F}\left(\prod_{i\in F}\beta_i\right)\left(\prod_{j\in U\setminus F}(1-\beta_j)\right)+$$

$$+\sum_{F\subseteq U,|F|=k,n\notin F}\left(\prod_{i\in F}\beta_i\right)\left(\prod_{j\in U\setminus F}(1-\beta_j)\right).$$

Вынесем множители с β_n за знаки сумм:

$$\alpha_n^k=\beta_n\sum_{F\subseteq U\setminus\{n\},|F|=k-1}\left(\prod_{i\in F}\beta_i\right)\left(\prod_{j\in(U\setminus\{n\})\setminus F}(1-\beta_j)\right)+$$

$$+(1-\beta_n)\sum_{F\subseteq U\setminus\{n\},|F|=k}\left(\prod_{i\in F}\beta_i\right)\left(\prod_{j\in(U\setminus\{n\})\setminus F}(1-\beta_j)\right).$$

Введем обозначение $U'=U\setminus\{n\}$ и рассмотрим выражение

$$\sum_{F\subseteq U',|F|=k-1}\left(\prod_{i\in F}\beta_i\right)\left(\prod_{j\in U'\setminus F}(1-\beta_j)\right).$$

Оно представляет собой не что иное, как вероятность того, что ровно $(k-1)$-й узел выйдет из строя в наборе U', состоящем из $n-1$ вычислительных узлов, и равняется α_{n-1}^{k-1}.

Аналогично

$$\sum_{F\subseteq U',|F|=k}\left(\prod_{i\in F}\beta_i\right)\left(\prod_{j\in U'\setminus F}(1-\beta_j)\right)=\alpha_{n-1}^k.$$

Получили рекурсивную формулу

$$\alpha_i^k=\beta_i\alpha_{i-1}^{k-1}+(1-\beta_i)\alpha_{i-1}^k. \qquad (8)$$

Базой рекурсии будут служить выражения:

$$\alpha_i^0=\prod_{j=1}^{i}(1-\beta_j)=\alpha_{i-1}^0(1-\beta_i)\text{ при }i>0$$

(вероятность того, что не будет сбоев ни на одном из i узлов);

$$\alpha_i^i=\prod_{j=1}^{i}\beta_j=\alpha_{i-1}^{i-1}\beta_i\text{ при }i>0$$

(вероятность того, что сбои произойдут на всех i узлах);

$$\alpha_0^0=1.$$

Применение формулы (8) напрямую опять окажется сопряжено с экспоненциальным по отношению к n ростом затрат вычислительных мощностей. Однако можно заметить, что большинство значений α_i^k будет вычисляться по нескольку раз, и если однажды вычисленные значения хранить в таблице, то сложность алгоритма составит $O(n^2)$. Это следует из того, что размер таблицы не превысит $(n+1)(n+1)$. А на вычисление значения одной ячейки таблицы потребуется максимум две операции умножения и по

одной операции сложения и вычитания действительных чисел.

Текст алгоритма вычисления вероятности сбоя отказоустойчивой системы на псевдокоде приведен ниже. Его корректность следует из вышеприведенных формул и их обоснования.

```
Функция a(k, i)
        Если A[k,i] = −1, то
                Если k = 0, то
                        A[k, i] = a(0, i−1) * (1 − b[i])
                Иначе, если k = i, то
                        A[k, i] = a(k − 1, i − 1) * b[i]
                Иначе
                        A[k, i] = a(k − 1, i − 1) * b[i] +
                        + a(k, i − 1) * (1 − b[i])
        Вернуть A[k, i]
Конец функции a.
Инициализировать элементы массива A[n,n] значениями −1
A[0, 0] = 1
Надежность = a(0, n) + a(1, n) + ... + a((n−1)/3, n)
Вероятность отказа = 1 − Надежность
```

5. Оценка надежности существующих систем при использовании неоднородных по надежности компонент

Имея формулу и алгоритм для оценки надежности отказоустойчивых решений, рассмотрим несколько примеров, в которых применение существующих систем приводит к плохим результатам.

Пример 1. Рассмотрим случай, когда имеется один высоконадежный (вероятность сбоя равна 0,00001) сервер. Единственный способ повышения надежности, доступный существующим системам, заключается в увеличении параметра f – максимально допустимом количестве вычислительных узлов, выход которых из строя может выдержать система. В данном случае пусть f изменится с нуля (система не выдержит ни одного отказа, но благодаря высокой надежности сервера вероятность его отказа очень мала) на единицу. Это требует увеличения количества используемых вычислительных узлов с одного до четырех, т. е. необходимо три дополнительных вычислительных узла. Если вероятность сбоя добавляемых узлов велика (например, 0,01), то надежность получившийся системы будет меньше исходной.

Расчет надежности получившейся системы в соответствии с алгоритмом, приведенным в разд. 4, дан в таблице. Итоговая надежность равна сумме чисел из выделенных фоном ячеек – 0,99970170597. Вероятность сбоя системы целиком составит 0,00029829403. Таким образом, добавив три дополнительных узла, получили группу, которая хотя и может выдержать отказ одного из узлов, но все-таки вероятность сбоя двух или более узлов в тридцать раз выше, чем если бы использовался один высоконадежный сервер.

На приведенный пример можно взглянуть и с другой стороны. Пусть имеется система с четырьмя вычислительными узлами, вероятность сбоя каждого из которых равна 0,01. Тогда вероятность отказа всей системы равна 0,00059203. Если теперь заменить один из узлов на узел с вероятностью отказа в тысячу раз меньше, то выигрыш в вероятности отказа полученной системы окажется менее двух крат.

Пример 2. Пусть в группе примерно поровну высоконадежных и ненадежных вычислительных узлов ($\beta_{1,2,3} = 0{,}0006$, $\beta_{4,5,6,7} = 0{,}01$). Получим вероятность сбоя всей системы 0,00000506998066.

Стоит отметить следующее: если бы система состояла только из ненадежных вычислительных узлов, то вероятность отказа составила бы 0,00003396253015, а если только из надежных узлов, то 0,00000000754640. То есть вероятность сбоя уменьшается в 4500 раз. Следовательно, можно ожидать, что замена одного ненадежного узла на надежный уменьшит общую вероятность отказа примерно в 3,326 раза. Тогда можно надеяться, что общая надежность системы из трех

Пример расчета надежности группы из четырех узлов

k	i-й вычислительный узел				
	–	$\beta_1 = 0{,}00001$	$\beta_2 = 0{,}01$	$\beta_3 = 0{,}01$	$\beta_3 = 0{,}01$
	0	1	2	3	4
0	1,00000000000	0,99999000000	0,98999010000	0,98009019900	0,97028929701
1	−1,00000000000	0,00001000000	0,01000980000	0,01980960300	0,02941240896
2	−1,00000000000	−1,00000000000	0,00000010000	0,00010019700	0,00029729106
3	−1,00000000000	−1,00000000000	−1,00000000000	0,00000000100	0,00000100296
4	−1,00000000000	−1,00000000000	−1,00000000000	−1,00000000000	0,00000000001

надежных и четырех ненадежных узлов станет равной $0{,}00003396253015 \cdot 3{,}326^3 = 0{,}000000923$, что в 5,5 раза лучше, чем полученная.

Пример 3. Параметризируем пример 2. Пусть в системе $n = 10$ вычислительных узлов. Введем параметр p – количество надежных (вероятность сбоя 0,001) узлов. Остальные $n-p$ имеют вероятность отказа 0,01. Параметр p изменяется от нуля до n. Полученные результаты зависимости вероятности отказа всей системы от количества надежных вычислительных узлов представлены на рисунке.

Как видно из графика, в существующих отказоустойчивых системах при доле надежных вычислительных узлов меньше половины отдача от замены ненадежных вычислительных узлов на надежные значительно меньше (следует обратить внимание на логарифмическую шкалу), чем в случае, когда доля надежных узлов больше половины, хотя абсолютные значения вероятности сбоя в обоих случаях одинаковы.

6. Способы обеспечения отказоустойчивости в гетерогенных распределенных системах

Подробный разбор и доказательство корректности различных способов обеспечения отказоустойчивости в распределенных системах – тема для отдельной статьи. Здесь же лишь наметим общие подходы.

Способ 1. Балансировка нагрузки с учетом вероятности отказа вычислительных узлов. Если есть набор задач и дано множество вычислительных узлов, на которых эти задачи надо решить, то при создании групп для обеспечения отказоустойчивого решения каждой из задач рекомендуется следовать следующему правилу: объединять в группу вычислительные узлы с одинаковой надежностью. В этом случае, как было показано в предыдущем разделе, отдача от имеющихся ресурсов наибольшая. Если же создание групп с абсолютно одинаковыми по надежности узлами невозможно, то, используя предложенный алгоритм, можно оценить оптимальность каждого из возможных распределений и выбрать наилучшее.

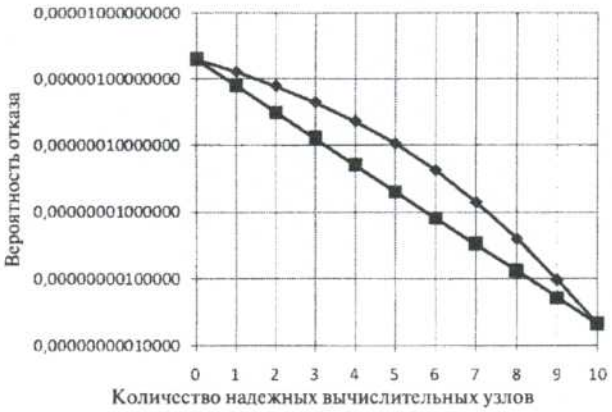


Иллюстрация зависимости общей надежности системы от количества надежных процессоров

(—◆—) – вероятность отказа, обеспечиваемая существующими отказоустойчивыми системами; (—■—) – минимально возможная вероятность отказа в системе, учитывающей неоднородную надежность компонент

Этот способ достаточно универсален и требует небольших модификаций существующих систем. Однако если перераспределение вычислительных узлов между задачами невозможно, а дана всего одна задача и фиксированный набор вычислительных узлов, которые можно использовать для ее решения, то следует использовать один из следующих способов.

Способ 2. Оценка надежности получаемого решения. Как было показано в первом примере, не всегда использование всех доступных вычислительных узлов обеспечивает максимальную отказоустойчивость. Прежде предлагается с помощью приведенного в данной статье алгоритма проверить, сможет ли какое-либо подмножество обеспечить большую надежность, чем все множество целиком.

Способ 3. Группировка вычислительных узлов. Группировать одинаковые по надежности вычислительные узлы можно и в рамках одной задачи. Так, в примере 2 можно сгруппировать четыре ненадежных узла в один и получить "виртуальный" узел с надежностью 0,00059203. Сгруппировав же его с тремя другими надежными узлами, получим вероятность отказа системы целиком 0,00000214394360, что в 2,4 раза лучше, чем полученная в том примере, хотя и она не является минимально достижимой.

Способ 4. Протокол надежной групповой рассылки, предусматривающий различную надежность вычислительных узлов. Именно протокол надежной групповой рассылки – узкое место в обеспечении отказоустойчивости (им накладывается ограничение, что доля сбойных процессов должна быть меньше 1/3 [5]). Представляется возможным создание такой реализации этого протокола, которая бы учитывала индивидуальные вероятности отказа вычислительных узлов.

Этот способ наиболее универсален, однако требует коренной переработки существующих систем.

В статье было продемонстрировано, как сбор статистических данных о надежности вычислительных узлов помогает в обеспечении более высокой отказоустойчивости создаваемых решений, особенно для гетерогенных распределенных систем. Это далеко не единственный пример работы более общего подхода, основная идея которого заключается в том, что чем более полная информация имеется о среде, в которой работает система, то тем более эффективное решение можно подобрать для этой конкретной ситуации [6–8]. Общие же решения часто оказываются гораздо менее эффективными, особенно если среда не однородна.

СПИСОК ЛИТЕРАТУРЫ

1. **Lamport L., Shostak R., Pease M.** The Byzantine Generals Problem // ACM Transactions on Programming Languages and Systems. July 1982. Vol. 4, № 3. P. 382–401.

2. **Reiter M.** The Rampart Toolkit for Building High-Integrity Services // Theory and Practice in Distributed Systems. Springer-Verlag, 1995.

3. **Kihlstrom K.P., Moser L.E., Mellar-Smith P.M.** The SecureRing group communication system // ACM Transactions on Information and System Security (TISSEC). November 2001. Vol. 4, № 4. P. 371–406.

4. **Castro M.** Practical Byzantine Fault Tolerance / Massachusetts Institute of Technology. 2001. PhD thesis.

5. **Bracha G., Toueg S.** Asynchronous Consensus and Broadcast Protocols // Journal of the ACM. October 1985. Vol. 32, № 4. P. 824–840.

6. **Фирсов А.Н.** Оптимизация на основе статистических данных асинхронной распределенной системы, устойчивой к произвольным отказам // Параллельные вычислительные технологии (ПаВТ'2009): Тр. междунар. науч. конф. Нижний Новгород, 30 марта – 3 апр. 2009 г. Челябинск: Изд-во ЮУрГУ, 2009. С. 765–771.

7. **Фирсов А.Н.** Оценка эффективности некоторых оптимизаций протоколов надежной и атомарной групповой рассылки // Вестник Пермского гос. ун-та. Математика. Механика. Информатика. 2009. С. 161–168.

8. **Фирсов А.Н.** Статистически оптимизируемая отказоустойчивая распределенная система // III конф.-конкурс грантов аспирантов и молодых ученых мех.-мат. ф-та Пермского гос. ун-та. Сб. тез. науч. докл. конф. Пермь, 2008. С. 74–78.

УДК 621.396.969.181.34

Ю.Е. Сидоров, А.В. Шумилов

ИССЛЕДОВАНИЕ ПОМЕХОУСТОЙЧИВОСТИ РАНГОВОГО ОБНАРУЖИТЕЛЯ СИГНАЛОВ ПРИ АПРИОРНОЙ НЕОПРЕДЕЛЕННОСТИ

Построение радиотехнических систем оптимального обнаружения, классификация и оценка параметров сигнала требуют наличия большого числа априорных данных, которыми разработчики систем в действительности не располагают. Отсутствие априорных данных – априорная неопределенность – может рассматриваться как специфический вид помех, и эффективная борьба с ними должна осуществляться специальными средствами [1].

В теории оптимального статистического синтеза сегодня широко используется байесовский подход, базирующийся на критерии минимума среднего риска [1–3]. Последовательная реализация этого подхода к решению задач обнаружения или различения сигналов и оценки их параметров, даже если преодолены математические трудности, обусловливает необходимость для каждого типа сигналов и помех и способа их комбинирования иметь, вообще говоря, свое специальное приемное устройство, что на практике осуществить практически невозможно. И потому приемное устройство, спроектированное оптимальным образом для аддитивных нормальных помех, может оказаться существенно менее эффективным при неаддитивных помехах, распределение которых к тому же заметно отличается от нормального.

Возникает проблема поиска структур оптимальных приемных устройств, устойчивых к изменению априорной информации о сигнале, помехе и способе их взаимодействия.

Описание алгоритма обнаружения

Рассмотрим построение правила оптимального обнаружения флуктуирующего сигнала в виде последовательности импульсов с неизвестными параметрами и законом флуктуаций в аддитивных шумах с неизвестными распределениями и дисперсией [4, с. 86]. Правило основано на сравнении колебаний, принимаемых от $m = 2$ смежных участков разрешения по дальности или по углу. Исходной выборкой служат статистически независимые отсчеты U_{1i} и U_{2i} огибающих колебаний (некогерентная обработка) с выхода детектора приемника соответственно в первом и во втором участках, $i = \overline{1,n}$, где n – число наблюдений. Просмотр дистанции осуществляется с помощью последовательного метода, причем появление сигнала допускается только в одном из участков. Присутствие сигнала в каком-либо участке наблюдения приводит к стохастическому возрастанию соответствующей величины U_j, $j = 1, 2$, по сравнению с другой.

Задачу обнаружения сформулируем как задачу проверки сложной гипотезы H о равенстве распределений величин U_1 и U_2 (сигнала нет) относительно сложной альтернативы K, состоящей в том, что одна из величин U_1 или U_2 стохастически больше другой (сигнал есть либо в первом, либо во втором участке дальности).

Проблема проверки этих гипотез инвариантна относительно всех преобразований вида $U_1' = \overline{g}_1(U_1)$, $U_2' = \overline{g}_2(U_2)$, где $\overline{g}$ – любая непрерывная и монотонная функция. Максимальным инвариантом относительно этих преобразований служит совокупность рангов величин U_1 и U_2, взятых как элементы одной выборки. Ясно, что для U_1 и U_2 можно указать только два ранговых вектора: первый ($r_1 = 1$, $r_2 = 2$) и второй ($r_1 = 2$, $r_2 = 1$), где r_1 и r_2 – ранги величин U_1 и U_2. Если над U_1 и U_2 производится n наблюдений, то среди них имеется x_1 наблюдений, дающих первый ранговый вектор (первый исход), и $x_2 = n - x_1$

наблюдений, дающих второй ранговый вектор (второй исход). Обозначим через $p = (p_1, p_2)$ вероятность этих исходов.

В отсутствие сигнала равновероятно получить первый или второй ранговый вектор и гипотеза $H: p = \pi$, где $\pi = (\pi_1, \pi_2) = \left(\frac{1}{2}, \frac{1}{2}\right)$. Если сигнал есть в одном из участков, то альтернатива $K: p \neq \pi$ (либо $p_2 > p_1$, p_1 и p_2 априорно неизвестны). Таким образом, задача обнаружения свелась к проверке таких гипотез, для которых оптимальное правило проверки удается получить лишь в предположении неограниченного объема наблюдений, т. е. при $n \to \infty$ [5]. Такое правило будет асимптотически оптимальным.

Последовательностями альтернатив $p_j^{(n)}$, стремящихся к $\pi_j = 1/2$, при $n \to \infty$, $j = \overline{1, 2}$, которые отражают наиболее интересные стороны поведения функции мощности состоятельного правила и дают полезную аппроксимацию действительной мощности при больших, но конечных n, являются те, для которых $\sqrt{n}(p_j^{(n)} - \pi_j)$ имеет нулевой предел, так что $p_j^{(n)} = \pi_j + \Delta_j/\sqrt{n} + R_n$, где $\sqrt{n}$, $R_n \to 0$ при $n \to \infty$, $\pi_j = 1/2$, $j = \overline{1, 2}$.

Пусть $y = (x - n\cdot\pi_1) / \sqrt{n} = (x_1 - n/2) / \sqrt{n}$ [5]. При $n \to \infty$ распределение $p(y)$ стремится к нормальному распределению со средними $E(y) = 0$ при гипотезе H и $E(y) = \Delta_1$ при альтернативе K и с дисперсией $\sigma^2 = \pi_1(1 - \pi_1) = 1/4$ в обоих случаях, т. е. к $p(y) = \sqrt{\frac{2}{\pi}}\exp\left\{-2(y - \Delta_1)^2\right\}$. Тогда проверяемые гипотезы H и K принимают вид: $H: \Delta_1 = 0$; $K: \Delta_1 \neq 0$. Равномерно наиболее мощное несмещенное (РНМН) ранговое правило проверки гипотезы H в этой асимптотической модели имеет форму [5]:

$$\varphi(y) = \begin{cases} 1 \text{ при } y < C_1 \text{ или } y > C_2; \\ \gamma_j \text{ при } y = C_j,\ j = \overline{1, 2}; \\ 0 \text{ при } C_1 < y < C_2, \end{cases}$$

где $C = C_1 = C_2$ (из-за симметричности $p(y)$ относительно точки $y = 0$ при гипотезе H) и γ находятся по заданной α (вероятность ложной тревоги) из условия

$$\alpha = 2\int_C^{\infty} p_H(y)dy.$$

Здесь интеграл берется по распределению y при $H: \Delta_1 = 0$. Из представленных выражений видно, что полученное правило не зависит от неизвестных распределений принимаемых сигналов и шумов и от их параметров и поэтому определяет (с учетом того, что $y = \left(x_1 - \frac{n}{2}\right)/\sqrt{n}$) структуру работоспособного в условиях непараметрической априорной неопределенности обнаружителя. Он состоит из блоков выделения величин U_1 и U_2 (линейный тракт приемника, линейный детектор); блока ранжирования, в котором происходит сравнение величин U_1 и U_2 и образование их рангов; устройства вычисления $y = \left(x_1 - \frac{n}{2}\right)/\sqrt{n}$ и блоков сравнения с нижним и верхним порогами. При превышении нижнего порога принимается решение о том, что сигнал находится в первом участке, а при превышении верхнего порога принимается решение о нахождении сигнала во втором участке.

Отметим, что правило также справедливо, если подсчитывается число исходов x_2 вместо x_1.

Функция мощности РНМН правила может быть вычислена при условии задания конкретных распределений наблюдаемых величин U_1 и U_2 при альтернативе K и будет зависеть только от параметра Δ_1. Например, пусть распределения U_1 и U_2 равны $(U_{1i}/\sigma_1^2)\exp(-U_{1i}/2\sigma_1^2)$ и $(U_{2i}/\sigma_2^2)\exp(-U_{2i}/2\sigma_2^2)$ соответственно, где $\sigma_1^2 = \sigma^2 + \varepsilon^2$ и $\sigma_2^2 = \sigma^2$, если сигнал присутствует в первом участке, и $\sigma_1^2 = \sigma^2$ $\sigma_2^2 = \sigma^2 + \varepsilon^2$, если сигнал присутствует во втором участке, а σ_1^2, σ_2^2 – неизвестные дисперсии импульсов, отраженных соответственно от первого и второго участков, на выходе линейного тракта приемника (ЛТП); σ^2 и ε^2 – неизвестные дисперсии шума и сигнала на выходе ЛТП. Тогда, если $\sigma_1^2 = \sigma^2$, $\sigma_2^2 = \sigma^2 + \varepsilon^2$, параметр $\Delta_1 = \frac{q\sqrt{n}}{2(2+q)}$, функция мощности будет описываться выражением

$$\beta(q) = 2 - \phi\left[\phi^{-1}\left(1 - \frac{\alpha}{2}\right) + \frac{q\sqrt{n}}{2+q}\right] - \phi\left[\phi^{-1}\left(1 - \frac{\alpha}{2}\right) - \frac{q\sqrt{n}}{2+q}\right],$$

где $q = \varepsilon^2/\sigma^2$ – отношение сигнал/шум; α – вероятность ложной тревоги; $\phi(z) = \frac{1}{\sqrt{2\pi}}\int_{-z}^{z}\exp\left(-\frac{t^2}{2}\right)dt$; ϕ^{-1} – функция, обратная $\phi(z)$ [6].

Итак, полученное РНМН ранговое правило обнаружения флуктуирующих сигналов в шумах с неизвестными распределением и мощностью не зависит от априорных сведений о сигналах и шумах, имеет стабильную вероятность ложной тревоги при любой мощности шума, максимальную вероятность правильного обнаружения и удобно для реализации в автоматических системах обработки гидролокационной информации.

Алгоритм моделирования обнаружителя флуктуирующего сигнала

Для моделирования обнаружителя используем новое направление в оценке эффективности и работоспособности радиотехнических систем, связанное с проведением машинных имитационных экспериментов на основе метода статистического моделирования (метода Монте-Карло). Привлекательность этого направления обусловлена уникальной возможностью получения конструктивных результатов испытания модели реальных процессов и систем (как правило, сложных) в их взаимодействии с изменяющейся окружающей средой в течение длительного времени. Эти модели могут быть развивающимися, а путем изменения их параметров можно спрогнозировать работоспособность радиотехнических систем. В предыдущей статье [7] метод статистического моделирования был успешно применен для исследования работоспособности непараметрического обнаружителя Манна – Уитни. Были определены этапы имитационного моделирования вплоть до анализа их результатов.

В данной статье обобщен опыт имитационного моделирования представленного рангового обнаружителя флуктуирующих импульсов с неизвестными параметрами и законом флуктуаций в шумах с неизвестными распределениями и дисперсией.

Алгоритм моделирования обнаружителя флуктуирующего сигнала имеет следующий вид:

1. Образование независимых случайных величин U_{1i} и U_{2i}, которые соответствуют двум элементарным участкам дальности, причем сигнал может присутствовать только в одном из этих участков.

2. Сравнение величин U_{1i} и U_{2i} и образование их рангов.

3. Определение порога $C = C_1 = C_2$ и вероятности γ_j; $0 \le \gamma_j \le 1$.

4. Вычисление $y = (x_1 - n/2) / \sqrt{n}$ и сравнение полученного значения с порогом C.

5. Вычисление вероятности событий, заключающихся в том, что $y > C$, $y < C$ или $y = C$, путем N-кратного повторения операций 1, 2, 3 и 4 при определенных n и α для каждого отношения сигнал/шум q.

Схема рангового обнаружителя приведена на рис. 1. Здесь введены обозначения: Д – детектор, ВР – вычислитель рангов, УЗОВ – устройство запоминания опорной выборки, СС – схема сравнения, СИ – счетчик инверсий, ВРС – вычислитель ранговой статистики, ЗУ – запоминающее устройство, ПУ – пороговое устройство, С – сигнал. Блок-схема моделирующего алгоритма приведена на рис. 2.

Анализ полученных характеристик

В монографии [4, с. 87] дается следующее соотношение для вычисления вероятности ложной тревоги:

$$P(|y| > C) = 2\int_{C}^{\infty} p_H(y)dy.$$

Интеграл берется по распределению y при $H: \Delta_1 = 0$.

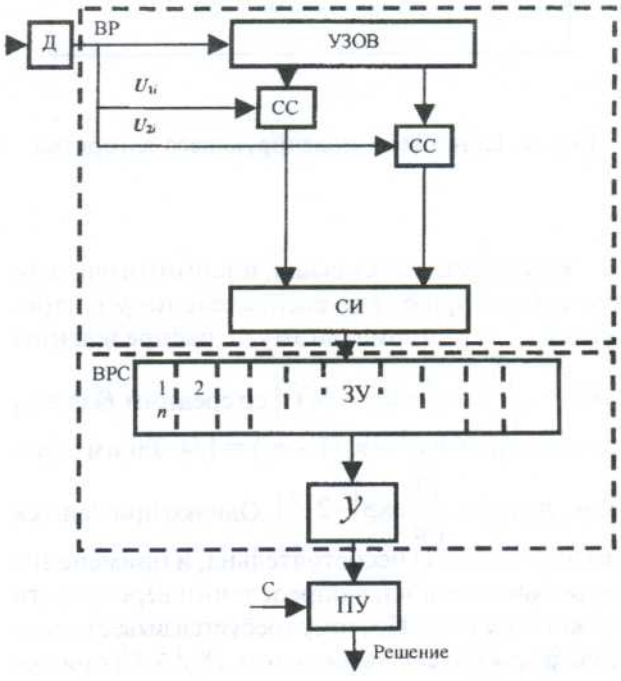


Рис. 1. Схема рангового обнаружителя



Рис. 2. Блок-схема моделирующего алгоритма

Как упоминалось выше, в асимптотическом случае, т. е. при $n \to \infty$, распределение $p(y)$ стремится к нормальному распределению $p(y)=\sqrt{\frac{2}{\pi}}\exp\{-2(y-\Delta_1)^2\}$ со средним $E(y)=\Delta_1$ и с дисперсией $\sigma^2=\pi_1(1-\pi_1)=1/4$. Таким образом, $p_H(y)=\sqrt{\frac{2}{\pi}}\exp\{-2y^2\}$. Однако при малых n данная формула несостоятельна, и применение ее вызовет ошибки в определении вероятности ложной тревоги. Поэтому требуется вывести точную формулу для вычисления $P(|y| > C)$ при конечных n. Произведем расчеты.

В случае справедливости гипотезы отсутствия сигнала проведение наблюдения фактически соответствует опыту с подбрасыванием монетки: "орел" – выпал ранговый вектор (1, 2), "решка" – вектор (2, 1). Как известно, вероятность появления "орла" или "решки" равна 1/2, а величина, определяющая количество выпавших, например, "орлов", подчиняется биномиальному распределению. Таким образом, $P(x_1) = (1/2)^n C_n^{x_1}$.

Величина y является линейным преобразованием от величины x_1. Получим

$$P(y)=(1/2)^n C_n^{\sqrt{n}\cdot y+\frac{n}{2}}.$$

Отсюда

$$P(y>C)=\sum_{C_i>C}(1/2)^n C_n^{\sqrt{n}\cdot C_i+\frac{n}{2}}.$$

Очевидно, что из-за симметричности распределения относительно нуля $P(y > C) = P(y < -C)$. Отсюда

$$P(|y|>C)=2P(y>C)=2\sum_{C_i>C}(1/2)^n C_n^{\sqrt{n}\cdot C_i+\frac{n}{2}}.$$

Следует отметить, что один из минусов дискретных статистик при конечном ограниченном n – невозможность достижения сколь угодно малой вероятности ложной тревоги α. Чем больше n, тем меньше минимальное значение α. Очевидно, чем выше порог решающего устройства, тем ниже вероятность ложной тревоги. Минимальное значение достигается при максимально возможном пороге, который соответствует случаю, когда либо $x_1 = n$, либо $x_2 = n$. Вероятность этого события в отсутствие сигнала $P = 2(1/2)^n$. Таким образом, минимально возможное n, при котором обеспечивается заданная вероятность ложной тревоги α, определяется из соотношения $2(1/2)^{n_{\min}} \le \alpha$. Логарифмируя обе части этого неравенства, можно получить

$$n_{\min} \ge \mathrm{Log}_{1/2}\left(\frac{\alpha}{2}\right).$$

Для моделирования рассмотрим два случая, в первом количество отсчетов $n = 20$, во втором $n = 50$. На основании полученных выше соотношений была составлена таблица, значения C и γ из которой используются при дальнейшем моделировании.

Данные для моделирования

n	α	C	γ
20	10^{-1}	0,894	0,793
20	10^{-2}	1,342	0,803
50	10^{-1}	0,849	0,65
50	10^{-2}	1,273	0,389

Измерения производились в приложении Simulink математического пакета Matlab R2008b. Для расчета зависимости вероятности правильного обнаружения от отношения сигнал/шум было произведено $N = 10^3$ испытаний. Зависимости вероятности правильного обнаружения β от отношения сигнал/шум $q = 10 \log \frac{\sigma_c}{\sigma_ш}$ представлены на рис. 3.

Итак, рассмотренный оптимальный ранговый обнаружитель флуктуирующих сигналов в шумах с неизвестными распределением и мощностью не зависит от априорных сведений о сигналах и шумах, имеет стабильную вероятность ложной тревоги при любой мощности шума, максимальную вероятность правильного обнаружения и удобен для реализации в автоматических системах обработки радиотехнической информации.

Опыт статистического имитационного машинного моделирования обнаружителя флуктуирующего сигнала показал, что оно является удобным и надежным средством определения эффективности (вероятности правильного обнаружения) обнаружителя при различных вариациях параметров сигнала и помех, отражающих изменения реальных условий радионаблюдений. Поэтому статистическая имитация на ЭВМ может быть рекомендована разработчикам систем обнаружения для оценки эффективности этих систем еще на стадии проектирования, что позволит сэкономить время и материально-технические ресурсы при проведении натурных испытаний этих систем. Приведенное описание машинного эксперимента с ранговым обнаружителем флуктуирующих импульсов может оказаться полезным при создании конкретных устройств.

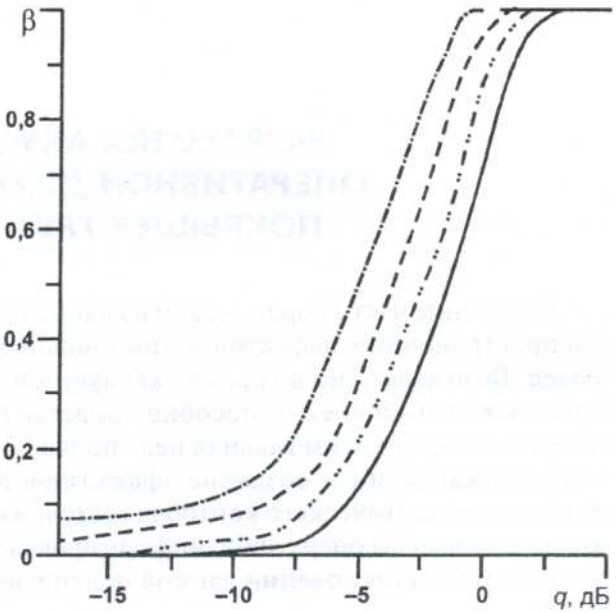


Рис 3. Вероятность правильного обнаружения
(—·—) – $\alpha = 10^{-1}$, $n = 50$;
(———) – $\alpha = 10^{-2}$, $n = 20$;
(— —) – $\alpha = 10^{-1}$, $n = 20$;
(— ·· —) – $\alpha = 10^{-2}$, $n = 50$

СПИСОК ЛИТЕРАТУРЫ

1. **Левин Б.Р.** Оптимальные алгоритмы обнаружения сигналов, устойчивые к изменению априорных данных // Изв. вузов СССР. Радиоэлектроника. 1970. Т. 13, № 2. С. 109–121.

2. **Миддлтон Д.** Введение в статистическую теорию связи: В 2 т. М.: Сов. радио, 1962. Т. 2. 831 с.

3. **Цикин И.А.** Оптимальная обработка сигналов в радиотехнических системах: Учеб. пособие / ЛПИ им. М.И. Калинина. Л., 1986. 77 с.

4. **Сидоров Ю.Е.** Статистический синтез автоматизированных решающих систем при априорной неопределенности. М.: Воениздат, 1993. 231 с.

5. **Леман Э.** Проверка статистических гипотез. М.: Наука, 1964. 498 с.

6. **Прокофьев В.Н.** Асимптотический оптимальный обнаружитель флуктуирующих сигналов в шумах неизвестной интенсивности // Изв. вузов СССР. Радиоэлектроника. 1971. Т. 14, № 6. С. 626–632.

7. **Сидоров Ю.Е., Шумилов А.В.** Исследование непараметрического обнаружителя радиосигналов // Научно-технические ведомости СПбГПУ. Информатика, телекоммуникации, управление. 2008. № 6. С. 78–85.

УДК 62-254

В.А. Грудинин, Т.П. Белякова

РАЗРАБОТКА АКУСТИЧЕСКОГО МЕТОДА ОПЕРАТИВНОЙ ДИАГНОСТИКИ СОСТОЯНИЯ ПОКРЫШЕК ГРУЗОВЫХ АВТОМОБИЛЕЙ

Отслоение части корда – одно из наиболее распространенных дефектов автомобильных колес. Возникает оно во время эксплуатации грузовых автомобилей и способно привести к аварии. В связи с этим главная цель проводимых исследований – создание эффективной системы автоматического контроля состояния колес, которая бы оперативно информировала водителя о возникновении данной неисправности.

Движение исправных и неисправных (с отслоившейся частью корда) колес сопровождается сильно отличающимися даже на слух акустическими сигналами [1, 2]. Всестороннее исследование данных сигналов позволяет найти и сформулировать конкретные признаки (критерии) возникновения дефекта, на основе которых выстраивается система оперативного контроля состояния колес во время движения автомобиля.

Изучение процесса возникновения упомянутых акустических сигналов позволяет сделать заключение, что их появление обусловлено взаимодействием протектора шины и дорожного покрытия во время движения автомобильного колеса. Малейшее изменение данного взаимодействия (любой дефект покрышки колеса) неминуемо приводит к соответствующему изменению акустических сигналов [3].

Таким образом, отслоение части корда, проявляющееся в виде квазипериодических ударов отслоившейся части о дорожное покрытие, должно сопровождаться акустическими сигналами, в которых присутствуют составляющие, обусловленные резонансными явлениями, а также подверженные амплитудной модуляции (модулированные квазипериодическими ударами отслоившейся части корда) [3–5].

С помощью экспериментальной установки (грузовик, оснащенный системой записи акустических сигналов, сопровождающих движение его колес) были записаны акустические сигналы, сопровождающие движение исправных и неисправных колес. Поиску и сравнению отдельных параметров (критериев) полученных сигналов и посвящены описываемые ниже исследования.

Сравнение статистических характеристик исходных сигналов, полученных для исправных и неисправных колес, не показало явных признаков их различия. Причиной этого является наличие в регистрируемых сигналах большого количества внешних шумовых составляющих, не обусловленных взаимодействием протектора шины и дорожного покрытия (шум мотора, других транспортных средств и т. п.).

Спектральный анализ сигналов позволил найти полосу частот (1930–2290 Гц), в которой находятся составляющие, обусловленные резонансными явлениями. Уровень энергии данных составляющих для неисправного колеса на 8–16 дБ превышает аналогичный уровень, полученный для исправного колеса (рис. 1); здесь и далее приводятся уровни сигналов, рассчитанные по отношению к максимально возможному. Наиболее устойчив данный критерий (уровень энергии сигнала в полосе частот 1930–2290 Гц), если интервал времени, для которого оценивается энергия полосового сигна-

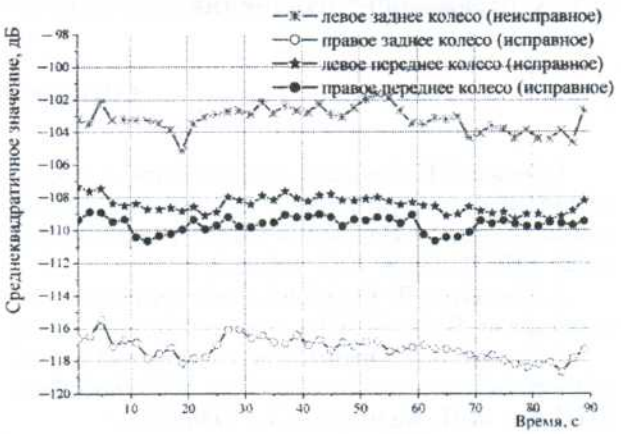


Рис. 1. Уровнеграммы полосовых сигналов исправных и неисправных колес

ла, составляет 11 периодов вращения колеса. Анализ временной структуры данного полосового сигнала свидетельствует об отсутствии в нем всплесков, обусловленных ударами отслоившейся части корда, т. е. об отсутствии амплитудной модуляции. Необходимо также отметить, что уровни соответствующих полосовых сигналов, полученные для различных (передних и задних) исправных колес, различаются на 8–9 дБ (рис. 1), в то время как аналогичные уровни, полученные для исправных колес, расположенных на одной оси, отличаются не более чем на 2 дБ (рис. 1). Данный факт связан с различной нагрузкой, прикладываемой на переднюю и заднюю оси автомобиля, т. е. на передние и задние колеса. В связи с этим было принято решение осуществлять сравнение уровней энергии полосовых сигналов для колес, расположенных на одной оси, т. е. отдельно для задних и передних колес. При этом подразумевается, что вероятность отслоения части корда одновременно для двух колес на одной оси пренебрежимо мала.

Исследование остальных спектральных составляющих сигналов неисправных колес в области частот 3–18 кГц позволило выявить наличие у шумоподобных составляющих квазипериодических всплесков уровня, обусловленных ударами отслоившейся части корда, т. е. амплитудную модуляцию [5]. Наибольшая относительная амплитуда данных всплесков (пик-фактор) соответствует полосовым шумоподобным сигналам, приходящимся на область частот 5050–5430 Гц. Уменьшение относительной амплитуды всплесков в сторону более высоких частот обусловлено сильным ослаблением высокочастотных акустических колебаний в воздушной среде при сохранении неизменного уровня собственных шумов системы записи сигналов. В свою очередь, уменьшение относительной амплитуды всплесков в сторону более низких частот обусловлено увеличением уровня внешних шумов.

Дальнейшие исследования были посвящены анализу огибающей (закона модуляции) именно этого субполосного сигнала. Для него были получены и проанализированы следующие параметры: средний уровень, пик-фактор, максимальные и минимальные значения, автокорреляционная функция и т. п. [6, 7]. Кроме перечисленных критериев подробному анализу подвергся также спектр сигнала огибающей.

Классический способ регистрации подобного рода дефектов заключается в оценке изменения величины пик-фактора [4]. Основной недостаток данного метода в том, что он позволяет регистрировать наличие дефекта только на последней стадии его развития (предаварийное состояние), а это, по понятным причинам, неприемлемо для нашего случая.

Другой не менее известный критерий регистрации подобных дефектов, относящийся уже к частотной области, – сумма энергий первых 8–12 гармоник сигнала огибающей, кратных частоте вращения колеса [5]. Однако для его использования в исходном виде необходимо, чтобы сигналы были установившиеся, а также имелась достаточно точная информация о частоте вращения колеса, т. е. о скорости движения автомобиля и диаметре колеса. Добиться выполнения данных условий можно только на специальных испытательных стендах [5]. Исследуемые сигналы в нашем случае можно условно считать установившимися лишь на весьма коротких временных отрезках (единицы секунд). А косвенная оценка частоты вращения колеса на основе измерения его диаметра и скорости движения автомобиля характеризуется большой погрешностью (единицы герц).

С учетом вышеизложенного на основе анализа возможных значений ускорения автомобиля, ожидаемых значений частот вращения колеса и требуемого разрешения спектрального анализа по частоте было выбрано компромиссное значение временного интервала, используемого для получения спектральных оценок (две секунды). Проблема косвенного измерения частоты вращения колеса (частоты первой гармоники в спектре сигнала огибающей) была решена посредством поиска максимальной частотной составляющей в области её ожидаемого расположения. Границы данной области оценивались посредством косвенного измерения (на основе приблизительно измеренных значений скорости автомобиля и диаметра колеса) значения частоты вращения колеса.

Особенности временной структуры сигнала огибающей (периодические всплески, обусловленные ударами отслоившейся части корда) позволяют предположить, что ее спектр будет

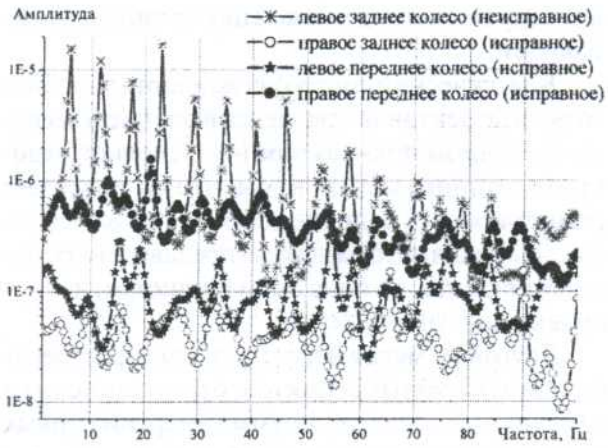


Рис. 2. Спектральные оценки сигнала огибающей

состоять преимущественно из гармоник, кратных частоте вращения колеса (рис. 2). В качестве метода получения спектральных оценок был выбран наиболее подходящий для данного случая модифицированный ковариационный метод (МКМ), параметры которого подбирались эмпирически [8].

Полученные (для отдельных выборок сигналов неисправных колес) значения суммы энергии первых 11 гармоник, кратных частоте вращения колеса, на 15–20 дБ (см. таблицу) превышают аналогичные значения, полученные для исправных колес (сравниваются сигналы колес, расположенных на одной оси). Однако проверка данного критерия на интервале времени длительностью в несколько минут выявила наличие отдельных сбоев, которые происходили практически одновременно для передних и задних колес, точнее с некоторым запаздыванием для задних. Детальное рассмотрение временной структуры сигнала огибающей (как для исправных, так и для неисправных колес) позволило выявить наличие в нем отдельных всплесков, амплитуда которых на несколько порядков превышает остальные. Было установлено, что данные всплески не периодические, а их относительное расположение на оси времени для передних и задних колес позволяет сделать предположение о том, что они обусловлены дефектами не колеса, а дорожного покрытия (например, выбоинами). Данное предположение было подтверждено экспериментально. Для этого на дороге в определенном порядке были размещены искусственно созданные дефекты дорожного покрытия. В записанных при движении автомобиля по данной дороге сигналах сильные всплески огибающей исправных и неисправных колес имели ожидаемый (соответствующий размещению дефектов на дорожном покрытии) порядок следования.

Пороговые значения критериев для исправных и неисправных колес

Критерий	Максимальное значение для исправных колес	Минимальное значение для неисправных колес
Разница уровней энергии полосовых сигналов, дБ	4,44	8,32
Разница уровней энергии первых 11 гармоник сигнала огибающей, дБ	2,56	17,05

Наличие в сигнале огибающей сильных всплесков, обусловленных дефектами дорожного покрытия, делает практически невозможным какой-либо анализ данных сигналов с целью регистрации отслоения части корда. Другими словами, невозможно отличить дефект дорожного покрытия от дефекта колеса, поскольку они практически одинаково влияют на процессы взаимодействия протектора шины колеса и дорожного покрытия и, следовательно, на временную и частотную структуру исследуемых сигналов.

Так возникает весьма сложная проблема локализации всплесков, обусловленных неровностью дорожного покрытия, и устранения их влияния на процедуру регистрации дефекта колеса, т. е. на процедуру оценки энергии первых 11 гармоник сигнала огибающей, кратных частоте вращения колеса.

При более детальном рассмотрении акустических сигналов, обусловленных дефектами колес и дороги, можно обнаружить их существенное отличие, которое было положено в основу метода устранения или уменьшения влияния неровности дорожного покрытия на получаемые спектральные оценки. Оно заключается в случайном характере и большой амплитуде всплесков, обусловленных дефектами дорожного покрытия. Именно большая энергия данных случайных всплесков приводит к значительным искажениям последующих спектральных оценок.

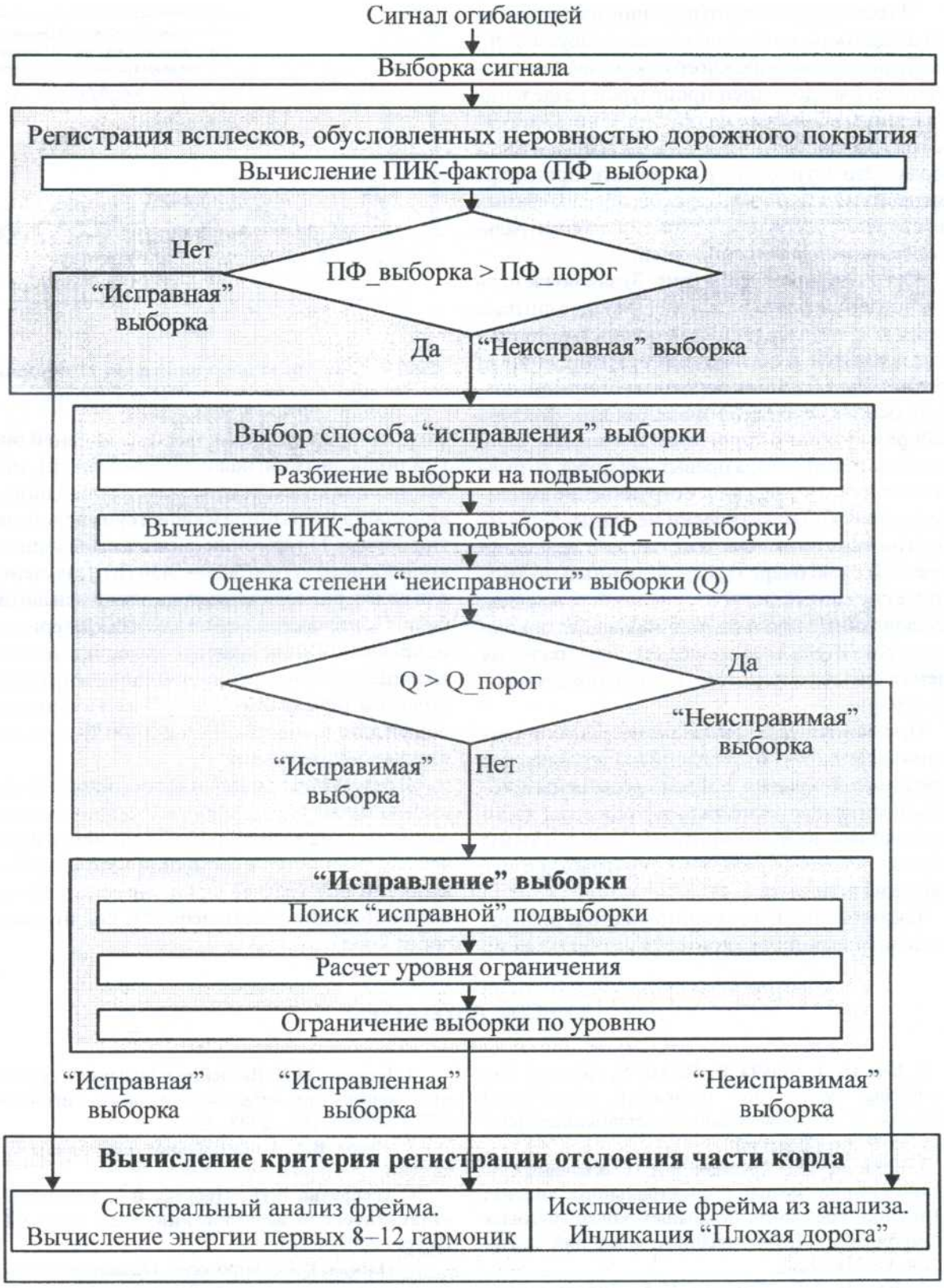


Рис. 3. Алгоритм регистрации и удаления всплесков, обусловленных неровностью дорожного покрытия

На основе учета этого различия был разработан алгоритм регистрации всплесков, обусловленных неровностью дорожного покрытия. Впоследствии он был дополнен процедурой разделения (селекции) всплесков на обусловленные неровностью дорожного покрытия и отслоением части корда. Это позволило разработать процедуру уменьшения влияния всплесков, обусловленных неровностью дорожного покрытия, на спектральные оценки сигнала огибающей.

Суть этой процедуры (рис. 3) заключается в установлении факта наличия в выборке сигнала всплесков, обусловленных неровностью дорожного покрытия, и последующем ограничении их уровня. Факт наличия всплеска устанавливается на основе сравнения значения пик-фактора выборки сигнала с пороговым значением. Если значение пик-фактора превышает порог, то имеется всплеск. Последующее ограничение амплитуды всплесков практически не сказывается на энергии гармоник, кратных частоте вращения колеса. В свою очередь, ограниченный по амплитуде случайный (не повторяющийся) всплеск, обусловленный неровностью дорожного покрытия, практически не искажает спектральные оценки сигналов исправных и неисправных колес (см. рис. 2).

Применение алгоритма регистрации и удаления всплесков, обусловленных неровностью дорожного покрытия, сделало возможным использование для регистрации отслоения части корда колеса второго критерия в виде суммы энергий первых 11 гармоник, кратных частоте вращения колеса (рис. 4).

Таким образом, в созданном методе диагностики оцениваются два критерия возникновения

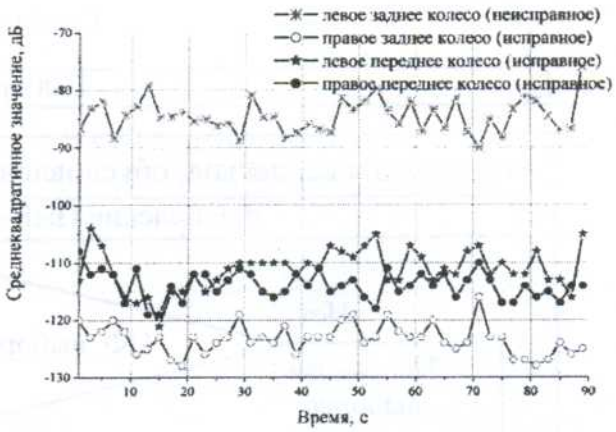


Рис. 4. Суммарная энергия первых 11 гармоник, кратных частоте вращения колеса

отслоения части корда: разница уровней энергии полосовых сигналов (1930–2290 Гц), полученных для колес, расположенных на одной оси автомобиля, и разница уровней суммарной энергии первых 11 гармоник сигнала огибающей узкополосного шума (5050–5430 Гц), полученных для колес, расположенных на одной оси автомобиля. Пороговые значения данных критериев для исправных и неисправных колес приведены в таблице, где даны результаты, полученные с использованием записей акустических сигналов общей длительностью 20 ч, для 100 %-ного доверительного интервала.

В результате проведенных исследований акустических сигналов, сопровождающих движение автомобильных колес, разработан метод акустической диагностики состоянии автомобильных колес, позволяющий оперативно регистрировать возникновение такого дефекта, как отслоение части корда.

СПИСОК ЛИТЕРАТУРЫ

1. **Leppala J., Rasila T., Ripatti S.** Acoustic tyre monitoring system for decreasing wheel-based accidents // http://www.taipaletelematics.com/tele/ATMOS_White_Paper.pdf

2. **Белякова Т.П., Грудинин В.А., Пшеничный С.А.** Использование методов спектрального анализа для оценки состояния покрышек колес грузовых автомобилей // Труды учебных заведений связи. № 174. С. 218–226.

3. **Барков А.В., Баркова Н.А.** Вибрационная диагностика машин и оборудования. Анализ вибрации / СПбГМТУ. СПб., 2004. 156 с.

4. **Баркова Н.А.** Введение в диагностику роторных машин по виброакустическим сигналам / СПбГМТУ. СПб., 2002. 156 с.

5. **Русов В.А.** Спектральная вибродиагностика. Вып. 1. Пермь, 1996. 424 с.

6. **Говорухин В.Н., Цибулин В.Г.** Компьютер в математическом исследовании: Учеб. курс. СПб.: Питер, 2001. 624 с.

7. **Назаров Н.Г.** Метрология. Основные понятия и математические модели. М.: Высш. школа, 2002. 245 с.

8. **Марпл С.Л.** Цифровой спектральный анализ и его приложение. М.: Мир, 1990. 332 с.

УДК 519.688

Б.Б. Казаков, С.А. Ивановский

ЭФФЕКТИВНЫЕ МЕТОДЫ ВЫДЕЛЕНИЯ ДВИЖЕНИЯ ОБЪЕКТОВ В ПОСЛЕДОВАТЕЛЬНОСТИ ИЗОБРАЖЕНИЙ

Выделение движения – важная часть задачи слежения за объектами в последовательности изображений. Основная трудность заключается в сопоставлении объектов, особенно когда они быстро движутся относительно частоты смены кадров. В этом случае необходимо иметь априорную информацию о модели и начальной скорости движения объектов или использовать вычислительно-сложные алгоритмы анализа последовательности кадров, например алгоритмы вычисления оптического потока.

При осуществлении этих задач могут быть использованы как следящие, так и обзорные видеокамеры. Разработанные алгоритмы позволяют в большой степени автоматизировать процесс слежения за объектами путем автоматической идентификации и выделения движущихся объектов в видеофрагментах. Использование параметров движения в качестве оцениваемых признаков помогает выделить движущиеся объекты даже в тех случаях, когда анализ отдельных кадров не позволяет этого добиться из-за небольших отличий наблюдаемых объектов от фона.

Процесс разработки и анализа алгоритмов потребовал проведения множества вычислительных экспериментов. Чтобы упростить эксперименты, был создан программный комплекс визуального исследования алгоритмов. Основная цель его создания – возможность визуального проектирования, исследования и модификации алгоритмов в интерактивном режиме.

Оптический поток. *Оптический поток* [1] – это поле двумерных векторов, соответствующее смещению пикселов изображений объектов при переходе к следующему кадру. Оптический поток (ОП) позволяет выделить движущиеся объекты в тех случаях, когда классические методы сегментации не позволяют этого сделать, и определить траекторию движущихся объектов на изображениях наблюдаемой сцены. Также существуют методы, позволяющие классифицировать тип деятельности человека по наблюдаемому в видеокадрах полю оптического потока.

Основное уравнение оптического потока имеет вид

$$I(x, y, t) = I(x + u, y + v, t + 1), \qquad (1)$$

где (u, v) – вектор оптического потока в точке с координатами (x, y) в момент времени t, который необходимо определить; $I(x, y, t)$, $I(x, y, t + 1)$ – значения интенсивностей точек изображения с координатами (x, y) в момент времени t и $t + 1$.

Различные методы решения этого уравнения отличаются друг от друга точностью, объемом вычислений и областью применения [5].

Локальные методы вычисления оптического потока. Для решения уравнения (1) построим разложения функции $I_1(x + u, y + v)$ в ряд Тейлора первого порядка [1]:

$$I_1 + \frac{\partial I_1}{\partial x}u + \frac{\partial I_1}{\partial y}v - I_0 = 0, \qquad (2)$$

где $I_0 = I_0(x, y) = I_0(x, y, t)$, $I_1 = I_1(x, y) = I_1(x, y, t + 1)$, или

$$I_x(x, y)u + I_y(x, y)v = -I_t, \qquad (3)$$

где $I_x = \frac{\partial I_1}{\partial x}$, $I_y = \frac{\partial I_1}{\partial y}$, $I_t = I_1 - I_0$.

Уравнение, рассматриваемое в каждой отдельной точке, недоопределено. Считая оптический поток в локальной области постоянным, найдем решение, используя метод наименьших квадратов (МНК). Чтобы найти вектор оптического потока в точке (x, y), необходимо знать минимум функционала

$$\sum_{j \in \omega} O_c(j)(I_x(j)u + I_y(j)v + I_t(j))^2, \qquad (4)$$

в котором ω – локальная окрестность вокруг точки, представленная в виде регулярной сетки точек; j – индекс точки с координатами (x_j, y_j) в локальной окрестности; $I_x(j)$, $I_y(j)$ – частные производные изображения в момент времени $(t + 1)$ в точке j; $O_c(j)$ – коэффициент взвешивания в точке j, который для получения корректного результата может быть выбран как коэффициент

двумерного биномиального фильтра с радиусом, равным радиусу окрестности.

Чтобы найти вектор ОП в точке (x, y), минимизирующий функционал (4), необходимо решить систему линейных уравнений

$$\begin{pmatrix} \sum_{j\in\omega} O_c(j)I_x^2(j) & \sum_{j\in\omega} O_c(j)I_x(j)I_y(j) \\ \sum_{j\in\omega} O_c(j)I_x(j)I_y(j) & \sum_{j\in\omega} O_c(j)I_y^2(j) \end{pmatrix} \begin{pmatrix} u \\ v \end{pmatrix} = \begin{pmatrix} \sum_{j\in\omega} O_c(j)I_x(j)I_t(j) \\ \sum_{j\in\omega} O_c(j)I_y(j)I_t(j) \end{pmatrix}. \quad (5)$$

Если нужно получить результаты на реальных изображениях, следует использовать мультимасштабное представление изображения [4]; для получения устойчивых оценок оптического потока надо применить какую-либо методику регуляризации.

Глобальные вариационные методы вычисления потока. Вариационные методы вычисления ОП основаны на минимизации функционала

$$\min_u \{E_{HS} = E_{HSr} + \lambda E_{HSi}\}, \quad (6)$$

в котором E_{HSr} – регуляризующий член; E_{HSi} – член подобия.

Классический метод предложен в работе [2]:

$$E_{HSr} = \int_\Omega |\nabla u|^2 + |\nabla v|^2 d\Omega;$$
$$E_{HSi} = \int_\Omega (I_1(x+u, y+v) - I_0(x,y))^2 d\Omega, \quad (7)$$

где область Ω соответствует всему изображению.

Одна из современных модификаций этого метода TVL1 представлена в работе [3], в которой описан итерационный метод минимизации функционала:

$$\min_u \{E_{TVL1} = E_{TVr} + \lambda E_{L1i}\}. \quad (8)$$

Здесь

$$E_{TVr} = \int_\Omega |\nabla u| + |\nabla v| d\Omega; E_{L1i} = \int_\Omega |I_1(x+u, y+v) - I_0(x,y)| d\Omega \quad (9)$$

и регуляризующий член E_{TVr} допускает разрывы в поле оптического потока (в функционале [2] накладывалось ограничение на пространственную непрерывность оптического потока), а информационный член E_{L1} за счет использования нормы $L1$ является робастным функционалом, что делает результаты менее чувствительными к шуму.

Модифицированный локальный метод вычисления оптического потока. Вариационный метод вычисления оптического потока позволяет добиться более точного результата, чем локальные методы, однако требует больших вычислительных затрат, поскольку необходимо применять итерационный метод нахождения глобального минимума в задаче большой размерности. Локальный метод позволяет вычислить оптический поток гораздо быстрее, ибо позволяет найти решение для каждой точки кадра в замкнутом виде. При этом точность вычисляемого потока ниже: в локальном методе не учитывается определенная гладкость функции оптического потока.

Рассмотрим две модификации стандартного метода, которые позволяют регуляризовать решение локального метода, сохраняя при этом возможность нахождения решения в замкнутом виде.

Первая модификация заключается в использовании медианной фильтрации вычисленного оптического потока в мультимасштабном методе. При переходе от крупного масштаба к более мелкому векторы оптического потока фильтруются медианным фильтром. Этот фильтр – своего рода аппроксимация глобальной сглаживающей компоненты в функционале вариационного метода и обеспечивает свойство робастности.

Вторая модификация состоит в добавлении регуляризирующей составляющей в минимизируемый функционал. Он принимает вид

$$E = \sum_{j\in\Omega} O_c(j)(I_x(j)u + I_y(j)v + I_t(j))^2 + \sum_{j\in\Omega} \beta(j)O_c(j)(u^2 + v^2), \quad (10)$$

где $\beta(j)$ – параметр в соответствующей точке окрестности, значение которого может варьироваться в зависимости от анализа характеристик достоверности получаемых в этой точке оценок.

Внедрение этих двух модификаций, с одной стороны, ненамного увеличивает вычислительную сложность алгоритма, с другой – позволяет получить результаты, мало уступающие глобальному вариационному методу, а в некоторых ситуациях превосходящие его.

Слежение за объектами с использованием оптического потока

На основе модуля вычисления оптического потока разработан программный комплекс слежения за объектами. Разработанный метод слежения за объектами состоит из трех основных частей: сегментации изображений, выделения объектов, сопоставления объектов при смене кадров.

С целью сегментации изображений в зависимости от сложности сцены использовались алгоритм пороговой сегментации или алгоритм водораздела. Также в зависимости от типа сцены в качестве исходных данных для алгоритма сегментации может выступать или само изображение, или модуль вычисленного оптического потока.

На основе оптического потока определяется соответствие между объектами при смене кадров. Для каждого объекта в первом кадре определяется: среднее значение вектора оптического потока (этот вектор используется в роли вектора перемещения объекта); расчетное положение объекта в следующем кадре; соответствующий объект следующего кадра.

Структурная схема метода слежения за объектами представлена на рис. 1.

Реализация на графическом процессоре. В течение последних лет появились новые средства ускорения вычислений с использование графических процессоров (GPU). Графический процессор – это параллельный конвейер. В современных ускорителях такого типа присутствует около 30–150 процессорных модулей. За один такт графический процессор может параллельно выполнить 30–150 операций. Это и делается в случае, когда алгоритм требует выполнить одновременно множество независимых операций с плавающей точкой.

Для решения подобных задач можно использовать язык параллельных вычислений CUDA [6], разработанный фирмой NVIDIA. Программа состоит из двух основных частей: вызывающего и инициализирующего модуля, работающего на центральном процессоре, и масштабируемого ядра, работающего на GPU. Ядро организовано таким образом, чтобы его можно было выполнить на различном количестве процессорных модулей, что позволяет повышать скорость выполнения программы пропорционально количеству модулей.

Итерационный метод решения уравнения (8) организован таким образом, что его легко перенести на архитектуру GPU. Как показали результаты эксперимента, реализация TVL1-алгоритма на GPU позволила значительно повысить скорость вычислений и вычислять оптический поток в реальном времени для изображений среднего размера (512×512). При этом время вычислений в пакете matlab, например, составляет 7 с, а при использовании ускорителя GPU – 0,03 с, т. е. за одну секунду можно обработать 33 кадра.

Сравнение применимости методов вычисления оптического потока для слежения за объектами. Чтобы сравнить возможность использования различных методов вычисления оптического потока при выделении объектов, определим результаты работы методов на четырех сюжетах. Оценим точность выделения ОП, применяя три метода: нерегуляризованный локальный метод, глобальный вариационный метод, регуляризованный локальный метод.

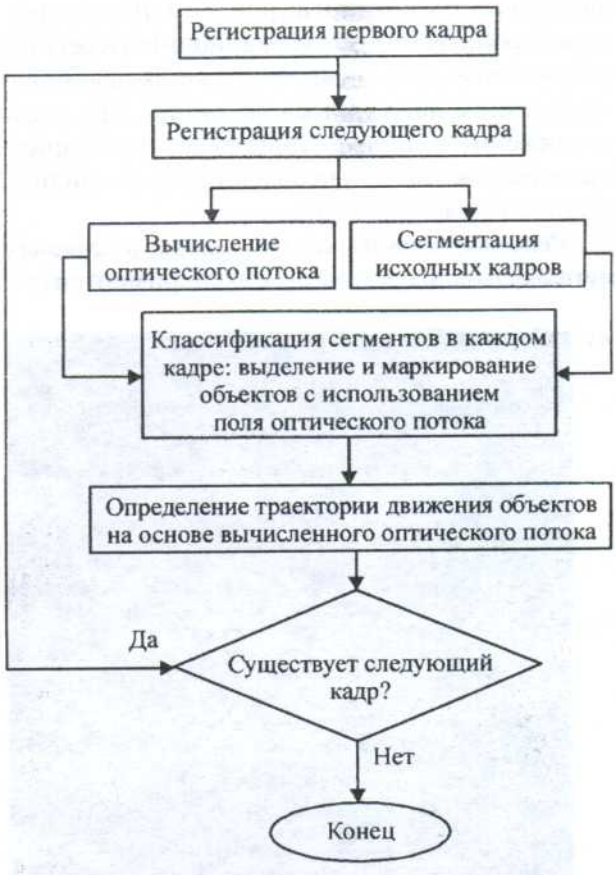


Рис. 1. Схема слежения за объектами на основе оптического потока

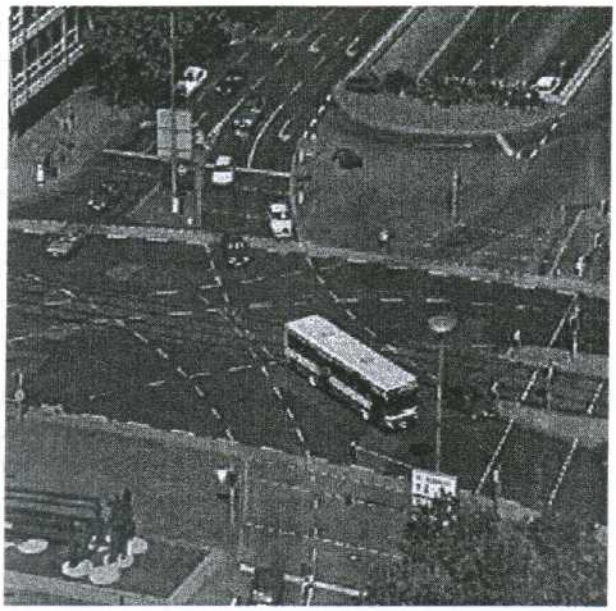


Рис. 2. Кадр из исходной последовательности (движение на перекрестке)

Результат обработки видеозаписи движения на перекрестке показан на рис. 2 и 3. В этом примере хорошо видно, что вариационный и регуляризованный локальный методы справляются с выделением движущихся объектов. При этом протяженный автобус лучше выделяется вариационным методом, а небольшие автомобили – регуляризованным МНК.

Рекомендации по использованию различных методов. Анализ результатов экспериментов позволяет сформировать определенные выводы по области применения рассмотренных методов.

Глобальные вариационные методы в общем случае позволяют получить устойчивый и предсказуемый результат. Из проведенных экспериментов только в одном случае выделение объектов было затруднено при использовании этих методов. В то же время эти методы и наиболее трудоемки, поскольку единственные из всех представленных используют итерационную модель решения задачи большой размерности (обычно скорость сходимости достаточно велика). При расчете на центральном процессоре для больших и средних изображений эти методы можно применять только в режиме постобработки. Использование для расчета современного графического сопроцессора позволяет значительно уменьшить время вычислений и использовать методы для слежения за объектами в реальном времени на средних по размеру изображениях (512×512).

Локальные методы не позволяют стабильно выделить объекты и пригодны для слежения только в редких случаях.

Регуляризованные локальные методы чаще всего дают результаты, достаточно близкие к результатам глобального вариационного метода, а в определенных случаях выделяют объекты точнее. Существенное преимущество регуляризованных локальных методов по сравнению с глобальными – намного меньший объем вычислений. Оценка производительности показала, что вре-

а) *б*)

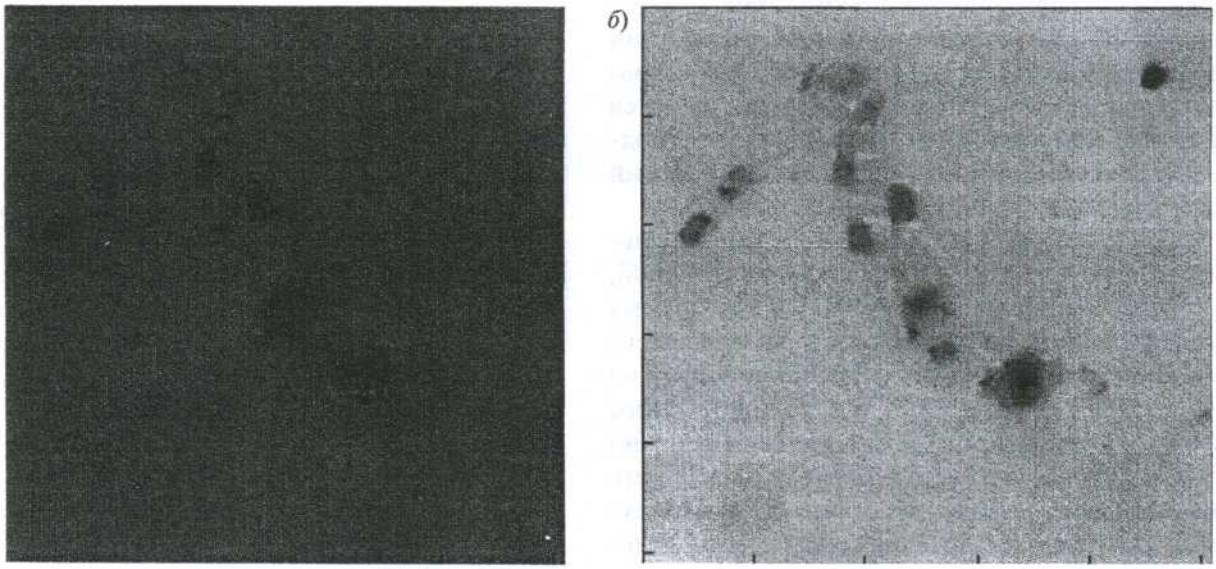


Рис. 3. Вычисленный оптический поток: вариационный метод (*а*) и регуляризованный локальный метод (*б*)

мя вычисления между парой кадров приблизительно лишь в восемь раз больше времени выполнения стандартной операции свертки одного кадра. Это делает возможным их применение во встраиваемых системах, в которых недоступны большие вычислительные мощности. Перенесение этих методов на графические сопроцессоры позволит следить за объектами в реальном времени в видеофрагментах с High Definition (1920×1800) и более высоким разрешением.

Сеточные методы используют несколько отличную схему вычислений, в которой решается система уравнений с размерностью, пропорциональной количеству точек в сетке. Для успешного применения сеточных методов необходимо построить хорошую начальную сетку, а также следить за тем, чтобы ее узлы не выходили за границы объекта. Последнее может привести к нарушению процесса слежение и непредсказуемому поведению узлов сетки. Эксперименты показали, что приемлемых результатов слежения за небольшими объектами можно добиться при приблизительно 17 узлах сетки. Это обусловливает необходимость решения системы линейных уравнений с 34 неизвестными, что приводит к наименьшему количеству операций из всех рассмотренных методов и делает их очень привлекательными для использования во встраиваемых системах.

Программный комплекс визуального исследования алгоритмов. Необходимость проведения большего количества экспериментов привела к созданию программного комплекса экспериментального исследования алгоритмов. Основной целью данного программного комплекса является возможность визуального проектирования, исследования и модификации алгоритмов в интерактивном режиме.

Алгоритм описывается в виде направленного графа. Для построения графа программа, реализующая алгоритм, разбивается на смысловые блоки. Отдельный блок представляется в виде черного ящика с набором входных и выходных данных. Блоки соответствуют узлам графа. Поток данных между блоками определяет ребра графа. Чтобы определить семантику ребер, каждому из них приписывается имя выходного и входного параметра. Два узла графа могут соединяться больше чем одним ребром при условии, что они (ребра) соответствуют разным входным параметрам. Естественное ограничение – то, что к одному узлу не могут присоединяться два ребра с одинаковым входным параметром. Значения входных параметров блока могут задаваться или с помощью ребер (в этом случае значение входного параметра равно значению соответствующего выходного параметра связанного блока и может изменяться при перевычислении блока), или в виде константных значений. Для задания константных значений в программном комплексе существуют стандартные элементы пользовательского интерфейса, позволяющие задавать атомарные значения (числа, строки, булевские значения), списки, массивы произвольной структуры.

Для задания семантики блоков выбран язык python, его основные достоинства:

интерпретируемость – это позволяет менять семантику блоков "на лету" без необходимости перекомпиляции;

наличие обширной библиотеки функций (для языка python существуют библиотеки математических функций, отображения произвольных графиков, пользовательского интерфейса, работы с графами и т. д.) [3];

наличие привычного синтаксиса и возможность использования нетипизируемых переменных (нетипизируемые переменные облегчают передачу данных между блоками; синтаксис языка напоминает синтаксис C++, однако имеются существенные отличия);

способность на лету встраивать подпрограммы на языке CUDA для эффективного вычисления «узких мест» вычислительных алгоритмов.

Любое изменение описания алгоритма или тестовых данных приводит к моментальному перевычислению выходных данных без необходимости перекомпиляции, перезапуска, переконфигурации программы. Результаты промежуточных вычислений буферизуются, что дает возможность сократить количество вычислений. При изменении блока анализ графа позволяет выполнять не всю программу целиком, а только последовательность зависимых операторов (подграф графа).

Используя рассмотренные выше методы, можно с высокой точностью выделять различные классы движущихся объектов в видеопоследовательностях. В зависимости от класса решаемой задачи и имеющихся технических средств можно использовать различные методы слежения. В большинстве случаев существует возможность выделить движение объекта в реальном времени.

Дальнейшее развитие методов слежения заключается в соединении алгоритмов вычисления

оптического потока с традиционными схемами рекуррентной (последовательной) обработки (типа фильтра Калмана).

Другим существенным ограничением рассмотренных локальных методов является предположение о постоянстве оптического потока в локальной области, которое нарушается на границах объекта. В рассмотренных вариационных методах влияние этого ограничения уменьшается за счет использования сглаживающих весовых функций. Внедрение каких-либо методов уменьшения такого влияния на границах объекта в локальные методы может существенно повысить точность выделения объектов и увеличить точность вычисления потока в целом.

СПИСОК ЛИТЕРАТУРЫ

1. **Lucas B., Kanade T.** An iterative image registration technique with an application to stereo vision // Proc. of Imaging understanding workshop. 1981. P. 121–130.

2. **Horn B.K.P., Schunck B.G.** Determing optical flow // Artificial Intelligence. 1981. Vol. 17. P. 185–203.

3. **Zach C., Pock T., Bischof H.** A duality based approach for realtime tv-l1 optical flow // Ann. Symp. German Association Patt. Recogn, 2007. P. 214–223.

4. **Luettgen Mark R., Clem Karl W., Willsky Alan S.** Efficient Multiscale Regularization with Applications to the Computation of Optical Flow // IEEE Trans. on image processing. 1993. P. 41–64.

5. **Barron J.L., Fleet D.J., Beauchemin S.S.** Performance of optical flow techniques // Intern. Journal of Computer Vision. 1994. P. 43–77.

6. **Ryoo S., Rodrigues C., Stone S.** et al. Program optimization study on a 128-core GPU // The First Workshop on General Purpose Processing on Graphics Processing Units. 2007.

УДК: 004.258

А.В. Томилин, М.С. Стуков, А.А. Стенин

ПРИМЕНЕНИЕ ДИВЕРСИФИКАЦИИ В ЗАДАЧАХ ДИНАМИЧЕСКОЙ ЗАЩИТЫ ПАМЯТИ

Современные методы несанкционированного доступа (НСД) к компьютерным системам (КС), использующие различного рода уязвимости и ошибки работы с памятью, позволяют менее привилегированному пользователю КС, взаимодействующему с более привилегированной программой, воспользоваться ее правами доступа к системным ресурсам с целью изменения хода ее выполнения. Примерами могут служить несанкционированный запуск командной оболочки (shell) с правами администратора системы, выполнение вредоносного кода или же обход функций программной защиты. Следствием успешной попытки НСД к КС может стать:

несанкционированный доступ к конфиденциальной информации;

нарушение целостности конфиденциальной информации;

подмена или повышение привилегий пользователя КС;

отказ в обслуживании сервисов целевой КС.

Как свидетельствует статистика CVE NVD NIST за 2008 год, в коммерческих операционных системах (ОС) семейства Microsoft Windows, проприетарных ОС семейства UNIX, свободных ОС семейства GNU/Linux было зарегистрировано 5844 уязвимости. При этом 862[1] из них (14,75 %)

[1] Статистика взята с учетом технических отчетов и уведомлений US-CERT и OVAL из CVE NVD NIST – Common Vulnerabilities and Exposures (CVE) National Vulnerability Database (NVD) National Institute of Standards and Technology (NIST) – Общие уязвимости и воздействия (CVE) в Национальном хранилище стандартов, основные уязвимости управления данными (NVD) при Национальном институте стандартов и технологий (NIST) США (см.: электр. ресурс http://nvd.nist.gov/).

пришлись на долю "классической" уязвимости "переполнение буфера" [1]. Статистика CVE NVD NIST, обзор популярного списка рассылки BugTraq, случаи массового распространения вирусов в сети Интернет, эксплуатирующих критические уязвимости в компонентах ОС, такие, как "эпидемии" LoveSun, Sasser, Code Red, показывают, что методы НСД (атак) на основе техник переполнения представляют наиболее серьезную угрозу информационной безопасности КС.

В статье рассматривается подход к защите КС, основанный на диверсификации программ. Предлагаемая технология, названная *стохастической адресной обфускацией*, представляет собой комбинацию трех различных техник – рандомизации размещения базовых адресов сегментов данных, стека, "кучи" (нераспределенной памяти), кода и библиотек, рандомизации порядка размещения буферов и функций (*пермутация*) и рандомизации дистанций (относительных адресов) между буферами (переменными) и функциями.

Анализ атак переполнения. Изучение механизмов атак, использующих техники переполнения [2–7], показало, что для успешного проведения любой из них нарушитель должен детально понимать внутреннее устройство целевой программы, а именно знать (предсказать) точное расположение в памяти уязвимых данных, указателей, кода, библиотек; расположение и размеры используемых буферов (переменных) и порядок их следования. Алгоритмы атак в общем случае выполняют подмену адреса возврата, хранимого в стеке, подмену указателя на код или данные либо выполняют переполнение уязвимого буфера с целью модификации указателя или адреса возврата под буфером. В современных ОС, как правило, сегмент кода доступен только для чтения, для остальных сегментов памяти C/C++ программ можно выделить следующие комбинации цели атаки и метода ее достижения:

1. Модификация указателя на код (функции, виртуальные функции, функции динамической компоновки).

2. Модификация указателя на данные (сохраненного в стеке, в куче, в области неинициализированных данных (BSS), в таблицах GOT, в таблицах динамического связывания, VTBL, в таблицах переходов).

3. Модификация данных (переменные, динамические буферы, буферы longjmp).

В табл. 1 приводится классификация "вектора атаки" согласно комбинациям цели атаки и объекта модификации.

Для осуществления комбинаций 1 и 2 нарушителю требуется знать абсолютный адрес цели модификации, поэтому подобный класс атак можно назвать *атаками, зависящими от абсолютных адресов* (а); для осуществления комбинации 3 ему необходимо вычислить относительное расстояние (дистанцию) между перезаписываемым буфером и данными (адресами), подлежащими модификации, поэтому такой класс атаки можно назвать *атаками, зависящими от относительных адресов* (б). На рисунке представлена схема атак согласно их целям, объединяющая (а) и (б).

Анализ методов динамической защиты памяти. Современные методы динамической защиты памяти C/C++ программ от атак на основе техник переполнения можно разделить на пять категорий:

Таблица 1

Классификация вектора атак

Объект атаки / Цель модификации		Стек	Указатель в стеке	Куча	Указатель в куче	BSS	Указатель в BSS
Адрес возврата		+	+	–	+	–	+
Сохраненный базовый указатель		+	+	–	+	–	+
Указатель функции	Локальная переменная	+	+	+	+	+	+
	Параметр функции	+	+	+	+	+	+
	Служебные таблицы	–	+	–	+	–	+
Буфер longjmp	Локальная переменная	+	+	+	+	+	+
	Параметр функции	+	+	+	+	+	+
Указатель кадра		+	+	–	+	–	+



Схема атак согласно целям

методы проверочных значений (StackGuard, Stack Shield, Stack smashing protector);

методы проверки объектов (Jones & Kelly, CRED, Pointguard);

неисполняемая память (PaX, MemGuard);

методы на основе эталонных функций с защитой адреса возврата (LibSafe);

методы диверсификации (PaX ASLR).

Проверочные значения эффективны против атак разрушения стека и их аналогов, но не защищают в ситуациях, когда объектом атаки является куча или BSS. Применение неисполняемой памяти приводит к существенному снижению производительности ОС, подобные методы сложны в реализации и не поддерживаются некоторыми аппаратными архитектурами. Проверки объектов эффективны только против атак, нацеленных на указатели, т. е. когда объект атаки – куча. Методы на основе эталонных функций с защитой адреса возврата заменяют ограниченный набор системных функций (getc( ), strcpy( ), printf( ) и т. д.) на их доверенные аналоги и не отслеживают ситуации переполнения пользовательских функций.

Задачи динамической защиты памяти C/C++ программ могут быть эффективно решены методами, подобными диверсификации, предложенной Форрестом, Сомаяджи и Экли в [8]. Их подход заключается в искусственном увеличении стекового кадра добавлением случайного числа пробелов. При каждой операции размещения в стеке объекта размером больше 16 байт в стек случайным образом добавляется от 8 до 64 (с дискретным увеличением, кратным восьми) "пустых" байт. Таким образом, каждая копия уязвимой программы становится уникальной, т. е. отлаженная атака против одной копии программы будет неудачной против других. Описанный метод применяется в модификации GCC v2.7.2.1.

Стохастическая адресная обфускация. Случайность в порядке размещения объектов программы существенно усложняет злоумышленнику задачу сбора информации, необходимой для проведения атаки, значительно повышает трудоемкость разработки атак, снижает не только вероятность успеха атаки, но и степень повторяемости атак на разных экземплярах одной и той же программы. Кроме того, для достижения успеха нарушитель вынужден выполнить серию одинаковых атак, что, вероятнее всего, приведет к разрушению целевой программы и, соответственно, к обнаружению попытки НСД (например, внезапный останов сервиса ОС MS Windows или крах демона ОС UNIX).

На наш взгляд, подход диверсификации программ рандомизацией содержимого стековых кадров, предложенный Форрестом, Сомаяджи и Экли, целесообразно расширить до рандомизации адресов размещения объектов внутри сегментов и рандомизации абсолютных адресов размещения сегментов программ в памяти. Данные задачи могут быть решены за счет комбинации следующих трансформаций объектов защищаемых программ (табл. 2).

В отличие от метода PaX ASLR[2], где используются две техники трансформации (ран-

[2] PaX – заплатка (patch) к ядру ОС GNU/Linux 2,4 и 2,6, которая предоставляет возможность настроить минимальные права доступа приложений к страницам памяти и технологию диверсификации ASLR (Address space layout randomization).

Таблица 2

Методы и объекты трансформации

Метод	Объекты
Рандомизация базовых адресов	Сегмент кода. Сегмент данных. Область неинициализированных данных (BSS). Сегмент стека. Куча
Рандомизация адресов	Разделяемые библиотеки: а) при использовании mmap( ); б) на этапе связывания. Пользовательские функции. Локальные переменные в стеке
Пермутация	Статические переменные. Функции (в динамических разделяемых библиотеках или в исполняемом файле). Стековый кадр
Рандомизация дистанций	Успешные запросы выделения памяти malloc( ). Стековый кадр: а) между базовым указателем и локальной переменной; б) перед параметрами функций. Функции (в динамических разделяемых библиотеках или в исполняемом файле)

домизация базового адреса mmap( ) разделяемых библиотек и рандомизация базового адреса стека кучи и адресного пространства процесса ОС), стохастическая адресная обфускация использует четыре общие техники трансформации и затрагивает практически все аспекты функционирования защищаемой программы.

В общем случае к трансформации программ предъявляются следующие требования:

трансформация должна быть "прозрачной", т. е. не может требовать перекомпиляции защищаемых программ и библиотек;

трансформация должна обеспечивать универсальную защиту от всех методов НСД на основе техник переполнения;

трансформация не должна вносить существенных задержек в работу защищаемых программ.

Предлагаемые методы рандомизации абсолютных адресов объектов программ затруднят осуществление атак (а), рандомизация дистанций между буферами (переменными), функциями и их пермутация, соответственно, затруднит осуществление атак (б). Предлагаемую технологию – стохастической адресной обфускации – можно отнести к семейству методов динамической защиты памяти. Направления дальнейших исследований заключаются в разработке прототипа защиты, осуществляющего описанную трансформацию, и экспериментальном исследовании его свойств.

СПИСОК ЛИТЕРАТУРЫ

1. **Fayolle P.-A., Glaume V.** A buffer overflow study. Attacks & Defenses, ENSEIEB. Networks and Distributed Systems. 2002. P. 18–28.

2. **Levy E.** Smashing the stack for fun and profit // Phrack Magazine. 1996. Vol. 49. Article 14.

3. **Matt Conover.** Security Development, Heap overflow tutorial. 1999.

4. **Dil Dog.** The Tao of Windows Buffer Overflow. 1998.

5. **Riq A., Gera S.** Advances in format string exploitation // Phrack Magazine. 2002. Vol. 59. Article 12.

6. **Rix D.** Smashing C++ VPTR // Phrack Magazine. 2000. Vol. 56. Article 8.

7. **Blexim F.** Basic Integer Overflows // Phrack Magazine. 2002. Vol. 60. Article 10.

8. **Forrest S., Somayaji A., Ackley D.** Building diverse computer systems // 6th Workshop on Hot Topics in Operating Systems. 1997. P. 67–72.

УДК 621.391:004.421

В.С. Тутыгин, С.В. Шедов

НОВЫЕ АЛГОРИТМЫ ОБНАРУЖЕНИЯ И ОПРЕДЕЛЕНИЯ ПАРАМЕТРОВ ЗАШУМЛЕННЫХ СИГНАЛОВ

Одна из важнейших задач цифровой обработки сигналов – обнаружение информативного сигнала в потоке данных, искаженных шумами и помехами, и определение его параметров. Для этого применяются различные методы, такие, как временная фильтрация (накопление), частотная фильтрация, прямое и обратное преобразование Фурье, вейвлет-преобразование, корреляционный анализ [1]. Часто задача обнаружения и определения параметров информативных сигналов решается в условиях присутствия структурно-детерминированных помех, т. е. помех, свойства которых известны или могут быть определены, и флуктуационного шума. При оптимальном приеме сигналов на фоне структурно-детерминированной помехи и флуктуационного шума основными процедурами обработки становятся компенсация структурно-детерминированной помехи и фильтрация шума [2].

Весьма часто используемым средством цифровой обработки сигналов выступает быстрое преобразование Фурье (БПФ), а информативным параметром сигнала – значение частоты. В силу того, что во многих случаях изучаемый сигнал значительно зашумлен, значение частоты, получаемое с помощью БПФ, недостаточно точно. В данной статье обсуждаются причины, вызывающие погрешности определения параметров зашумленных сигналов, описаны предложенные авторами новые алгоритмы (DT-, FSA- и MFSA-алгоритмы) дополнительной цифровой обработки сигналов, производимой после выполнения БПФ, позволяющие значительно повысить точность определения значения частоты.

Задача вычисления значения частоты зашумленного дискретизированного сигнала актуальна, в частности, при спектральном анализе эхосигналов спектрометров ядерного магнитного резонанса (ЯМР). Сигнал ядерного спинового эха в магнитоупорядоченных веществах может быть обнаружен с помощью БПФ даже в условиях, когда амплитуда шума превышает амплитуду эхо-сигнала. Однако с учетом особенности формы эхо-сигнала (рис. 1) значение частоты будет определено со значительной погрешностью. Значение же частоты является информативным параметром эхо-сигнала.

Второй пример задачи, связанной с частотным анализом сигналов, – анализ спутниковых линейно-частотно-модулированных сигналов (ЛЧМ-сигналов). Используются они для топографической съемки поверхности Земли. В случае, когда факт проведения топографической съемки требуется контролировать, необходимо решить задачу обнаружения спутниковых ЛЧМ-сигналов и определения их параметров, в частности диапазона частот. Известен и широко используется такой способ об-

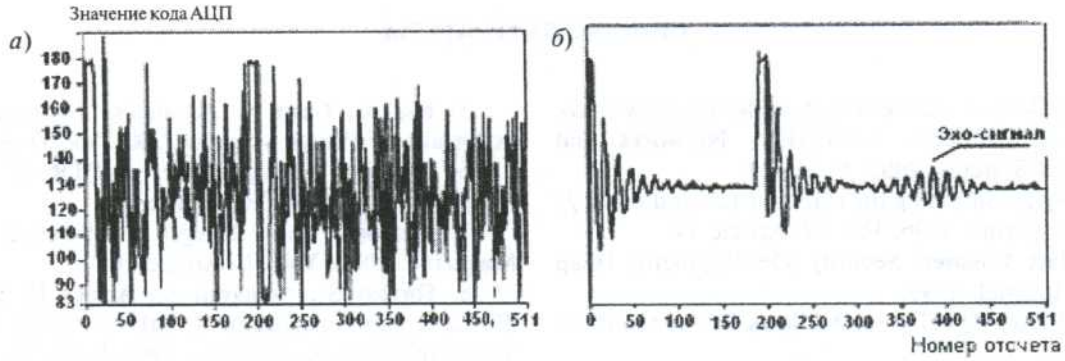


Рис. 1. Исходный (*а*) и накопленный (*б*) сигнал в борате железа

наружения спутниковых ЛЧМ-сигналов, как получение мгновенных частотных спектров спутниковых сигналов с помощью быстрого преобразования Фурье (БПФ); последующая цифровая обработка потока частотных спектров применяется для обнаружения ЛЧМ-сигналов в потоке спектров и определения их параметров.

Рассмотрим причины возникновения сложностей обнаружения информативных сигналов в спектрах и погрешностей определения параметров спектров, а также предлагаемые DT-, FSA-, MFSA-алгоритмы дополнительной цифровой обработки спектров, полученных с помощью БПФ.

Результатом БПФ дискретизированного эхо-сигнала спектрометра ЯМР определенной частоты становится количество периодов сигнала во временном окне. Если частота отсчетов или интервал дискретности по времени при измерении сигнала известен, то по количеству периодов во временном окне можно установить и частоту измеряемого сигнала. Точность определения частоты в спектре входного сигнала вполне определенна и зависит от количества периодов p сигнала. Если количество периодов целое, то частота с помощью БПФ находится абсолютно точно (при отсутствии зашумленности сигнала). Если же количество периодов не целое, то появляется погрешность определения частоты. Максимальное значение погрешности равняется $1/p$. В некоторых практически важных случаях, например при обработке эхо-сигналов (см. рис. 1) импульсных спектрометров ЯМР, количество периодов анализируемого сигнала во временном окне принципиально ограничено и составляет приблизительно 10 [3]. В этом случае погрешность определения частоты с помощью БПФ достигает 1/10, т. е. 10 %.

Влияние шума в регистрируемом сигнале во временной области можно значительно уменьшать за счет многократного повторения эксперимента и синхронного накопления эхо-сигналов (см. рис. 1).

Однако, как показано в [2], увеличение количества накоплений позволяет улучшать отношение сигнал/шум без искажения формы и уменьшения амплитуды накопленного эхо-сигнала лишь до некоторого предела. После этого предела накопление уже не приносит ощутимого улучшения качества. В частности, при большом времени накопления на эхо-сигнал начинают влиять постепенные изменения параметров приборов, входящих в состав импульсного спектрометра ЯМР. При ограничении времени проведения анализа веществ количество возможных накоплений сигнала должно быть либо ограничено, либо отсутствовать. Проблема обработки эхо-сигналов в условиях значительных шумов возникает и в том случае, когда имеется необходимость исследования веществ малой концентрации. Поэтому задача выполнения частотного анализа зашумленных эхо-сигналов актуальна.

Непосредственное использование БПФ для зашумленного сигнала не позволяет получить точное значение количества периодов во временном окне и частоты измеряемого сигнала в том случае, когда количество периодов не целое, а анализируемый сигнал занимает только часть временной области, модулирован по амплитуде и зашумлен. В практически важных случаях частотного анализа сигналов спектрометров ЯМР форма и начальная фаза эхо-сигналов известны. Это дает возможность создать эталонные сигналы, соответствующие ожидаемому эхо-сигналу по форме и начальной фазе, и произвести их корреляционное сравнение. Коэффициент корреляции эхо-сигнала с эталонным сигналом будет равен единице, если частоты эхо-сигнала и эталонного сигнала равны и эхо-сигнал не зашумлен. Поэтому при отсутствии шумов найти частоту эхо-сигнала можно, производя корреляционное сравнение с эталонными сигналами, частоту эталонных сигналов подбирать до выполнения условия, когда коэффициент корреляции станет равен единице. Однако коэффициент корреляции уменьшается как при разнице частот эхо- и эталонного сигналов, так и при совпадении частот, но из-за наличия шума. Поэтому данным способом определить частоту зашумленного сигнала невозможно.

Усовершенствование алгоритма определения частоты сигнала достигается за счет сочетания положительных качеств корреляционного подхода и БПФ с целью повышения точности определения частоты дискретизированного сигнала в условиях, когда обрабатываемый сигнал во временной области зашумлен, количество периодов не целое, сигнал модулирован по амплитуде, количество периодов сигнала мало.

Идея предлагаемого алгоритма, описанного в [4–6], названного нами DT-алгоритмом (Debelova-Tutygin-алгоритмом), заключается в том, что в небольшой окрестности от предполагаемой частоты сигнала (приближенное значение частоты сигнала может быть найдено с помощью БПФ) вычисляются коэффициенты корреляции с несколькими эталонными сигналами в некоторой окрестности от приближенного значения частоты. Затем с помощью сплайн-интерполяции и передискретизации строится функция, выражающая зависимость коэффициента корреляции от частоты эталонов и находится максимум этой функции, по положению максимума определяется частота эталонного сигнала. Функция, построенная таким образом, имеет вид параболы с явно выраженным максимумом (рис. 2) в случае как незашумленного, так и зашумленного сигнала, что и позволяет определить частоту эхо-сигнала более точно, чем это позволяет сделать БПФ. При наличии шума форма функции сохраняется, уменьшается лишь абсолютное значение максимума.

Точность определения частоты сигнала тем выше, чем ближе начальное приближение к истинной частоте. Поэтому DT-алгоритм применяет итерационное вычисление, на каждом этапе итерации в качестве начального приближения используется уточненное значение частоты, полученное на предыдущем этапе. В качестве первого приближения берется частота, определенная с помощью БПФ.

Краткое описание DT-алгоритма.

1. Выполнение БПФ рабочего числового массива исходных данных $X[i]$, $i = 0, \ldots, 2^n$ и получение числового массива частотного спектра $Y[j]$.

2. Нахождение значения j_{max}, соответствующего максимальному значению Y. (Найденная величина j соответствует приближенному значению периодов в числовом массиве $X[i]$.)

3. Создание $2k + 1$ эталонных числовых массивов исходных данных $X_s[i]$, $s = 0, 1, 2, \ldots, 2k$, с числом периодов $j_{max} - 1 + s/k$.

4. Вычисление коэффициентов корреляции рабочего числового массива со всеми эталонными и формирование числового массива коэффициентов корреляции $KK[m]$, $m = 0, 1, 2, \ldots, 2k$.

5. Выполнение сплайн-интерполяции для массива $KK[m]$ (нахождение непрерывной функциональной зависимости $F(m)$, соответствующей массиву $KK[m]$).

6. Выполнение передискретизации на основе найденной функциональной зависимости $F(m)$ для массива $KK[m]$ с увеличением количества элементов массива в r раз, т. е. формирование массива $KK1[m1]$, где $m1 = m \cdot r$.

7. Нахождение значения $m1_{max}$, которое соответствует максимальному значению $KK1$ (значение $m1_{max}/r$ будет представлять вещественное число, соответствующее уточненному – в общем случае нецелому – значению количества периодов в рабочем числовом массиве исходных данных $X[i]$).

8. Вычисление разности $err = m1_{max}/r - j_{max}$.

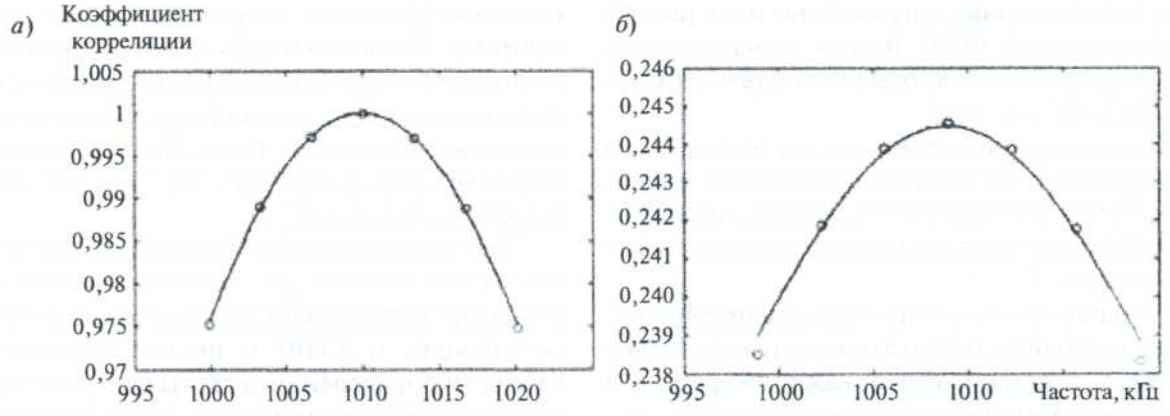


Рис. 2. Зависимость коэффициента корреляции от частоты при отсутствии шума (*а*) и при отношении сигнал/шум 1/3 (*б*)

Точное значение частоты равно 1010. Аппроксимирующие кривые построены с помощью функции сплайн-аппроксимации spaps в MATLAB

9. Если $err < err_{доп}$, то $j_{max} = m1_{max}/r$ и переход к п. 3, иначе – переход к п. 10.

10. Вывод найденного точного значения количества периодов $m1_{max}/r$ (в общем случае не целого) и частоты $f = (m1_{max}/r)/(dt \cdot 2n)$, где dt – шаг дискретности по времени при измерении сигнала $X(t)$.

Точность определения основной частоты при использовании предложенного алгоритма зависит от значений k и r и тем выше, чем они больше, однако если анализируемый сигнал зашумлен, существенное ее увеличение происходит при росте k и r лишь до некоторого предела. В частности, при соотношении сигнал/шум 1/2–1/3, $k = 3$ и $r = 10$, как показали проведенные исследования, оказывается наилучшим выбором по критерию точность/время анализа.

Количество итераций для расчета частоты с заданной точностью с помощью описанного алгоритма зависит от того, насколько близко к искомой частоте находится начальное приближение. Эффективность же БПФ в качестве средства для определения начального приближения частоты с увеличением уровня зашумленности сигнала снижается, погрешность определения частоты возрастает. Это связано с тем, что числовой массив, представляющий сигнал в частотной области, при зашумленности сигнала во временной области также будет зашумлен, и когда отношение сигнал/шум в частотной области станет меньше единицы, правильно установить (хотя бы приблизительно) частоту с помощью БПФ окажется невозможно. В связи с этим для определения начального приближения частоты при обработке значительно зашумленных сигналов нами был предложен новый FSA-алгоритм (Fourier-Spline-interpolation-Approximation) и программа [7], описанные ниже.

На рис. 3 приведена экспериментально полученная в режиме компьютерного моделирования зависимость погрешности определения частоты зашумленных эхо-сигналов в диапазоне частот от 1 до 5 МГц при частоте отсчетов 100 МГц. Генерация нормально распределенного шума выполнена с помощью функции randn с СКО, равным 0,5, при амплитуде эхо-сигнала, равной единице. Время затухания 1,3 мкс. Усреднение произведено по 200 реализациям, значение частоты изменялось с шагом 10 кГц.

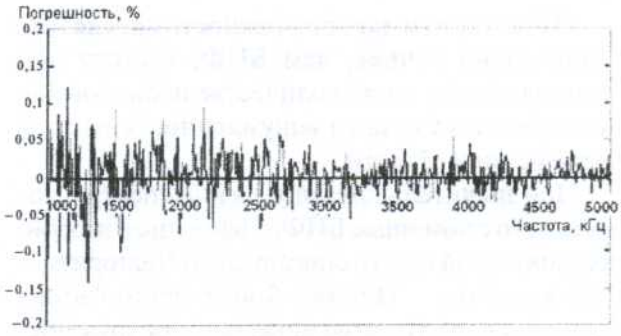


Рис. 3. Среднее значение погрешности определения частоты при использовании DT-алгоритма, полученное путем компьютерного моделирования. Частота отсчетов 100 МГц

Погрешность определения частоты после усреднения при использовании DT-алгоритма, как следует из рис. 3, в этих условиях не превышает 0,15 %. При использовании БПФ погрешность была 3 %. Коэффициент вариации частоты при использовании DT-алгоритма составил 1 %, при использовании БПФ – 4,5 %.

DT-алгоритм исследован нами в режиме компьютерного моделирования в среде MATLAB с нормальным, равномерным и белым шумом и при обработке эхо-сигналов импульсного спектрометра ЯМР. Для регистрации сигналов применялся быстродействующий АЦП с памятью LAn 10M6PCI (50 МГц, два канала, 256 Кбайт памяти). Программа регистрации и обработки данных была реализована в среде LabWindows/CVI 8.0 [7].

Эффективность предлагаемых авторами алгоритма и программы частотного анализа исследовалась в отношении сигналов, подобных получаемым с импульсных спектрометров ЯМР и приведенных на рис. 1, с одной частотной составляющей и экспоненциальной огибающей.

Результаты проведенных исследований возможностей нового метода позволили сделать следующие заключения:

DT-алгоритм дает возможность определять частоту точнее, чем БПФ, даже при большом количестве периодов сигнала во временной области;

DT-алгоритм позволяет получить точное значение частоты при обработке незашумленных сигналов;

DT-алгоритм по сравнению с БПФ обеспечивает более высокую точность при обработке зашумленных сигналов;

DT-алгоритм дает возможность определить значительно точнее, чем БПФ, частоту при сверхмалом (от двух) количестве периодов как незашумленного, так и зашумленного сигналов во временной области.

При значительном зашумлении спектра, полученного с помощью БПФ, эффективен второй, предложенный нами и описанный в [7] алгоритм – FSA-алгоритм (Fourier-Spline-interpolation-Approximation). В нем возможность обнаружения зашумленного сигнала и определения его параметров по спектру, полученному с помощью БПФ, достигается за счет сплайн-интерполяции, передискретизации с увеличением количества элементов дискретизированного сигнала, последующей аппроксимации массива частотного спектра аналитической функцией

$$f(x) = ae^{-\left(\frac{|x-b|}{c}\right)^d},$$

значения параметров a, b, c, d которой находятся методом оптимизации.

Найденное таким образом значение параметра b указывает на уточненное значение количества периодов в дискретизированном сигнале, что позволяет определить с высокой точностью значение частоты в спектре как при целом, так и при нецелом количестве периодов сигнала. Функция именно такого вида хорошо соответствует частотным спектрам разнообразной формы. Параметр a определяет амплитуду, b – положение линии в спектральном диапазоне, c – крутизну фронта и спада, d – степень островершинности (плосковершинности).

Краткое описание FSA-алгоритма.

1. Выполнение БПФ рабочего числового массива исходных данных $X[i]$, $i = 0, ..., M$, $M = 2n$, n – целое число, и получение числового массива частотного спектра $Y[j]$, $j = 1, ..., s$, $s <= 2n$. Величину s целесообразно ограничить максимально возможным ожидаемым количеством периодов сигнала $X[i]$.

2. Выполнение сплайн-интерполяции и передискретизации для массива $Y[j]$, $j = 1, ..., s$, с увеличением количества точек массива в K раз. Получаем массив частотного спектра $Z(p)$, $p = 1, ..., S$, $S = s \cdot K$. Количество точек в этом массиве больше в K раз, чем в массиве $Y[j]$.

3. Аппроксимация числового массива $Z(p)$ аналитической зависимостью

$$f(p) = ae^{-\left(\frac{|p-b|}{c}\right)^d}$$

путем нахождения численных значений коэффициентов a, b, c, d методом оптимизации по алгоритму Гаусса–Зейделя с фиксированным количеством N шагов по принципу минимизации суммы квадратов отклонений функций $z(p)$ и $f(p)$:

$$\sum_{p=1}^{S} \varepsilon_p^2 = \sum_{p=1}^{S} (z(p) - f(p))^2.$$

4. Вычисление уточненного (в общем случае нецелого) значения количества периодов в числовом массиве исходных данных $X[i]$)

$$Q = b/K$$

и частоты

$$f = Q/(dt\, M),$$

где dt – шаг дискретности по времени при измерении сигнала $X(t)$; M – количество значений дискретизированного сигнала в массиве $X[i]$).

Пример результатов обработки сигнала с помощью данного алгоритма приведен на рис. 4.

Точность определения основной частоты при использовании предложенного FSA-алгоритма зависит от значений K и N и тем выше, чем больше значения K и N, однако если анализи-

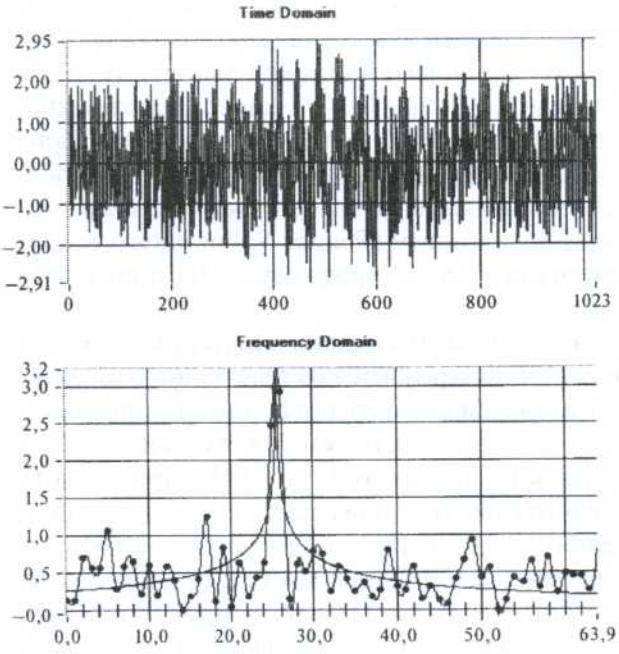


Рис. 4. Пример результатов обработки спектра БПФ с помощью FSA-алгоритма

руемый сигнал зашумлен, существенное увеличение точности определения частоты происходит при увеличении K и N лишь до некоторого предела. В частности, при соотношении сигнал/шум 1/2 – 1/3, $K = 10$ и $N = 200$, как показали наши исследования, оказывается наилучшим выбором по критерию точность/время анализа.

Высокая эффективность предлагаемого нового метода частотного анализа обеспечивается при условии, если количество периодов сигнала во временной области будет не менее двух, а количество отсчетов за период – не менее 20.

При решении второй задачи – обнаружении и определении параметров спутникового ЛЧМ-сигнала – мы имеем дело с потоком мгновенных спектров, полученных также с помощью БПФ. В каждом спектре (кадре), содержащем N (например, 128) дискретных значений, содержится шумовой фон и может содержаться одно или два дискретных информативных значения, принадлежащих ЛЧМ-сигналу (рис. 5).

Если мгновенное значение частоты, принадлежащей ЛЧМ-сигналу, точно совпадает с одним из N дискретных значений частоты в кадре, информативное значение в кадре может быть только одно, если же нет – информативных значений будет два. Один ЛЧМ-сигнал занимает некоторую полосу частот, поэтому он порождает M последовательных кадров, содержащих информативные значения. Значение M зависит как от полосы частот H ЛЧМ-сигнала, так и от шага дискретности по частоте h в спектре, полученном с помощью БПФ, а именно:

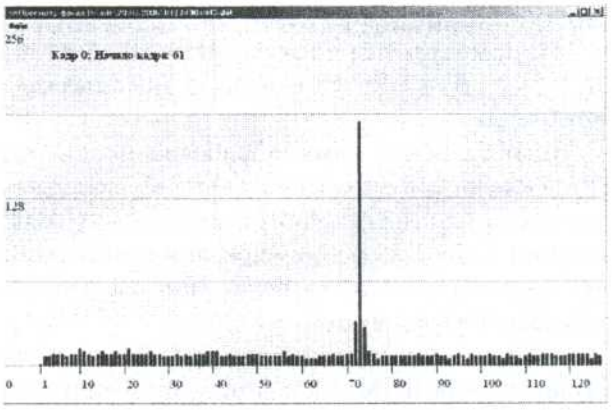


Рис. 5. Мгновенный частотный спектр, содержащий фрагмент ЛЧМ-сигнала

$M = H/h$. ЛЧМ-сигналы излучаются с некоторой периодичностью во времени, поэтому после каждых M кадров, содержащих информативные значения, следует P кадров, содержащих только шумовой фон.

Сложность задачи связана еще и с тем, что кроме фонового случайного шума принимаемый сигнал может содержать на отдельных участках сектора наблюдения узкополосный сигнал помехи, который может быть ошибочно принят за ЛЧМ-сигнал.

Известный способ обнаружения информативных сигналов основан на детектировании превышения амплитудой дискретных значений порогового уровня – уровня шума. Известный способ реконструкции ЛЧМ-сигнала по полученным мгновенным M спектрам – суммирование этих M кадров (спектров). Идентификация параметров ЛЧМ-сигнала производится по реконструированному спектру.

Недостаток известного способа обнаружения заключается в том, что кроме информационных значений, принадлежащих ЛЧМ-сигналу, в кадре могут содержаться отдельные одиночные дискретные значения узкополосного сигнала помехи большой амплитуды, такой же или большей, чем информативные. Поскольку информативная линия в одном кадре может быть только одна (или две), дискретные значения, принадлежащие ЛЧМ-сигналу и узкополосной помехе, неотличимы. Это приводит к ошибкам обнаружения и определения параметров ЛЧМ-сигнала. Кроме того, при разбиении потока спектров на следующие одна за другой группы, содержащие по M спектров в группе, возможен пропуск ЛЧМ-сигнала за счет того, что в состав M спектров группы попадет лишь часть ЛЧМ-сигнала.

В предлагаемом нами модифицированном FSA-алгоритме (MFSA-алгоритме) упомянутые недостатки устранены за счет того, что поток спектров (кадров) разбивается на группы по $S = 2M$ кадров в каждой группе; производится суммирование (накопление) S кадров каждой группы; номер кадра P_{i+1} начала каждой последующей группы вычисляется как $P_i + S - M$, а также за счет последующей аппроксимации накопленных спектров (кадров) аналитической функцией, предложенной нами в [7] для решения задачи частотного анализа сигналов спектрометров [3]:

$$f(x) = ae^{-\left(\frac{|x-b|}{c}\right)^d},$$

значения параметров a, b, c, d которой находятся методом оптимизации. Найденное таким образом значение параметра b определяет центральную частоту ЛЧМ-сигнала.

Функция именно такого вида хорошо соответствует ЛЧМ-сигналам. Параметр a определяет амплитуду, b – положение линии в спектральном диапазоне, c – крутизну фронта и спада, d – степень островершинности (плосковершинности). Отсутствие ложного обнаружения, т. е. принятия узкополосного сигнала помехи за ЛЧМ-сигнал, достигается за счет того, что при $c^d < H_1$ или $c^d > H_2$ (величины H_1 и H_2 связаны с полосой H обнаруживаемого ЛЧМ-сигнала) в качестве результата аппроксимации принимается $f(x) = 0$.

Краткое описание MFSA-алгоритма:

1. Начальные установки значений S, M, H, H_1, H_2, $i = 0$, $P_0 = -M$; $M >= (H/h)$, где H – полоса частот ЛЧМ-сигнала, h – шаг дискретности по частоте в спектре БПФ.

2. Вычисление начального номера кадра i-й группы кадров $P_i + 1 = P_i + S - M$.

3. Накопление $S = 2M$ последовательных мгновенных спектров (кадров) из последовательного потока спектров, $M > H/h$, где H – полоса частот ЛЧМ-сигнала, h – шаг дискретности по частоте в спектре БПФ.

4. Выполнение сплайн-интерполяции и передискретизации для накопленного спектра $Y[j]$, $j = 1, ..., M$ с увеличением количества точек массива в K раз. В результате получаем массив частотного спектра $Z(p)$, $p = 1..N$, $S = NK$. Количество точек в этом массиве больше в K раз, чем в массиве $Y[j]$.

5. Аппроксимация центрированного числового массива $Z(p)$ аналитической зависимостью

$$f(p) = ae^{-\left(\frac{|p-b|}{c}\right)^d}$$

путем нахождения численных значений коэффициентов a, b, c, d методом оптимизации по алгоритму Гаусса–Зейделя с фиксированным количеством шагов по принципу минимизации суммы квадратов отклонений функций $z(p)$ и $f(p)$:

$$\sum_{p=1}^{S} \varepsilon_p^2 = \sum_{p=1}^{S} ((z(p) - \bar{z}) - f(p))^2.$$

6. Вычисление значения f центральной частоты в ЛЧМ-сигнале:

$$f = F_0 + h \cdot b/K, \text{ если } H_2 > c^d > H_1,$$

где F_0 – начальная частота спектрального диапазона в спектре БПФ; h – шаг дискретности по частоте в спектре БПФ; H_1, H_2 – константы, значение которых связано с полосой обнаруживаемого ЛЧМ-сигнала.

7. Инкремент i и переход к п. 2, если $c^d <= H_1$ или $c^d >= H_2$.

Точность определения основной частоты при использовании предложенного MFSA-алгоритма зависит от значений K и тем выше, чем больше значения K, однако если анализируемый сигнал зашумлен, существенное увеличение точности определения частоты происходит при увеличении K лишь до некоторого предела. В частности, при соотношении сигнал/шум 1/1 $K = 10$, как показали наши исследования, оказывается наилучшим выбором по критерию точность/время анализа.

На рис. 6 приведены примеры результатов обработки потока мгновенных спектров, полученных в результате быстрого преобразования Фурье сигналов спутника RADARSAT. Пунктирной линией показан накопленный спектр, сплошной – аппроксимирующая функция, с помощью которой производится обнаружение и определение центральной частоты ЛЧМ-сигнала, созданные программой обнаружения и определения параметров дискретизированного ЛЧМ-сигнала [8], реализующей MFSA-алгоритм.

Параметры обработки: $M = 10$, $S = 20$, $H_1 = 5 \cdot 10^6$, $H_2 = 2 \cdot 10^8$. Объем анализируемых кадров 64 000.

Выполненная нами обработка реального ЛЧМ-сигнала спутника RADARSAT, содержащего весьма значительный объем данных (64 000 кадров), с использованием предложенного нами MFSA-алгоритма позволяет сделать следующие выводы:

алгоритм может успешно использоваться для обнаружения ЛЧМ-сигналов, если амплитуда ЛЧМ-сигнала превышает уровень шума;

алгоритм полностью исключает ошибочное принятие сигнала помехи, имеющего большую

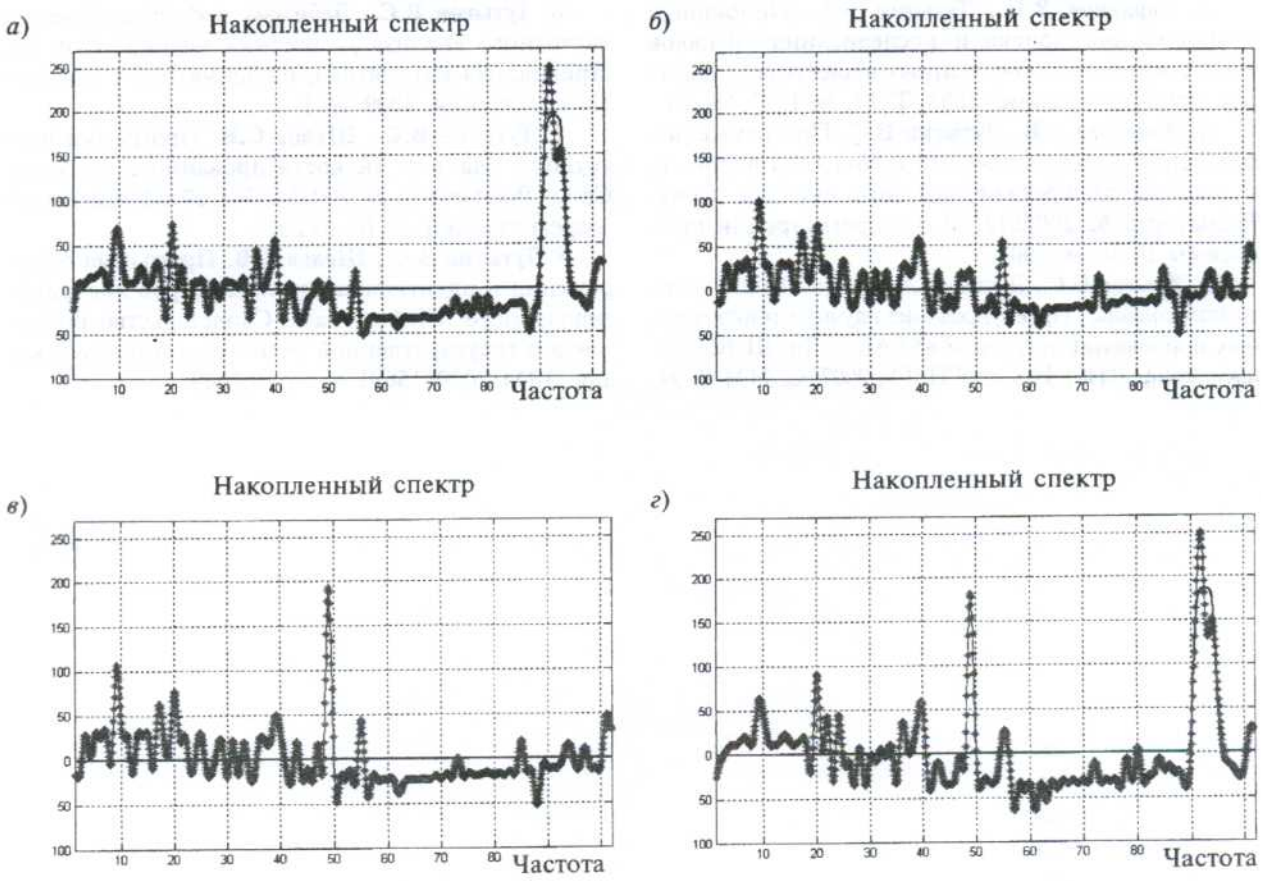


Рис. 6. Примеры обнаружения ЛЧМ-сигнала при использовании MFSA-алгоритма. В накопленных спектрах (кадрах) присутствует зашумленный ЛЧМ-сигнал (*а*), только шум (*б*), только шум и узкополосный сигнал помехи большой амплитуды (*в*), зашумленный ЛЧМ-сигнал и узкополосный сигнал помехи большой амплитуды (*г*)

амплитуду, но отличающегося по ширине полосы от ЛЧМ-сигнала, за ЛЧМ-сигнал;

алгоритм обеспечивает надежное обнаружение ЛЧМ-сигнала и в том случае, когда одновременно с ЛЧМ-сигналом присутствует помеха, сравнимая по амплитуде с ЛЧМ-сигналом, но отличающаяся по ширине полосы частот;

хотя алгоритм разрабатывался и был проверен для решения конкретной задачи, он может быть применен путем настройки параметров для обнаружения и идентификации параметров других спутниковых широкополосных сигналов по их частотным спектрам.

Полученные нами как путем моделирования, так и при обработке реальных зашумленных сигналов импульсного спектрометра ЯМР и реальных зашумленных сигналов спутника RADARSAT, имеющих весьма большой объем, результаты показали, что при использовании предложенных нами новых DT, FSA и MFSA-алгоритмов дополнительной цифровой обработки спектров упомянутых сигналов, полученных с помощью БПФ, могут быть значительно улучшены возможности обнаружения зашумленных эхо-сигналов спектрометров ЯМР и ЛЧМ спутниковых сигналов и повышена точность определения их параметров.

СПИСОК ЛИТЕРАТУРЫ

1. **Макс Ж.** Методы и техника обработки сигналов при физических измерениях: В 2 т. М.: Мир, 1983.

2. **Котоусов А.С., Морозов А.К.** Оптимальная фильтрация сигналов и компенсация помех. М.: Горячая линия – Телеком, 2008.

3. **Тарханов В.И., Тутыгин В.С.** Приборный комплекс для поиска и исследования сигналов ЯМР в магнитоупорядоченных веществах // Научное приборостроение. 2003. Т. 13, № 1. С. 58–63.

4. **Дебелова А.В., Тутыгин В.С.** Программа определения точных значений количества периодов и частоты дискретизированного сигнала. Свид. Роспатента № 2007612320 о госрегистрации программы от 01.06.2007.

5. **Тутыгин В.С., Дебелова А.В.** Новый метод частотного анализа: Проектирование научных и инженерных приложений в среде MATLAB // Тр. III Всерос. науч. конф. СПб.: Изд-во СПбГУ, 2007. С. 1494–1499.

6. **Тутыгин В.С., Дебелова А.В.** Новый метод частотного анализа // Вестник СПбГУ. Сер. 10. Прикладная математика, информатика и процессы управления. 2009. № 1.

7. **Тутыгин В.С., Шедов С.В.** Программа частотного анализа дискретизированного сигнала. Свид. Роспатента № 2007613363 об официальной госрегистрации от 09.10.2007.

8. **Тутыгин В.С., Шедов С.В.** Программа обнаружения и идентификации параметров дискретизированного ЛЧМ-сигнала. Свидетельство Роспатента о государственной регистрации программы для ЭВМ 2009615691 от 13.10.2009.

УДК 621.37

А.В. Ненашев, В.А. Охотников

РАЗРАБОТКА МЕТОДА АНАЛИЗА И МОДЕЛИРОВАНИЯ НЕЛИНЕЙНЫХ РАДИОТЕХНИЧЕСКИХ УСТРОЙСТВ

При создании современных средств радиосвязи широко используются системы автоматизированного проектирования (САПР). Они позволяют моделировать работу разрабатываемых устройств и оптимизировать их параметры. Увеличение интенсивности использования радиочастотных диапазонов и повышение требований к электромагнитной совместимости заставляют уделять значительное внимание нелинейным свойствам радиотехнических устройств. Задача создания методов анализа и моделирования в этом случае делится на две части: построение адекватных математических моделей нелинейных элементов, в частности полупроводниковых приборов, и разработка методов исследования указанных устройств с использованием таких моделей. Способы решения первой части задачи постоянно развиваются вместе с совершенствованием полупроводниковых технологий. В то же время вторая часть задачи, т. е. моделирование всего устройства с учётом нелинейных свойств, решается методами, основы которых разработаны несколько десятилетий назад.

Широко применявшийся в течение длительного времени метод численного решения дифференциальных уравнений, реализованный в пакете программ PSpice и других, подобных ему, выходит из употребления. Это связано с тем, что для устройств, работающих в диапазонах сверхвысоких частот, имеющих большой разброс постоянных времени, сложно получить решение в установившемся режиме. В популярном сегодня пакете Microwave Office применяется два метода нелинейного анализа: метод рядов Вольтерра (РВ) и метод гармонического баланса (ГБ). Первый является математически строгим, но применим лишь при малой нелинейности, когда нелинейные зависимости достаточно точно описываются степенным рядом не выше 3–5-го порядков. Метод ГБ – приближенный, так как основан на укороченных рядах Фурье с небольшим числом гармоник и строится в виде итерационной процедуры, сходимость которой не всегда можно оценить заранее. Кроме того, считается, что метод ГБ из-за необходимости многократных обращений матриц большого размера вносит ошибки, не позволяющие использовать его в режиме малой нелинейности. В [1] был предложен новый метод нелинейного моделирования, названный методом параметрических характеристик (ПХ), при котором нелинейные элементы рассматриваются как параметрические, их параметры изменяются под действием напряжений и токов. Описание метода приведено на примере типовой схемы диодного преобразователя частоты с одним нелинейным элементом – диодом, причём основные звенья схемы соединены в замкнутую цепь.

Постановка задачи: разработать метод, аналогичный методу ПХ для радиотехнических устройств произвольной конфигурации и степени сложности. Задача будет решаться на примере моделирования усилительного каскада СВЧ, выполненного на полевом транзисторе с барьером Шоттки (ПТШ). Такие усилительные каскады широко применяются в современных радиорелейных и спутниковых системах связи, например в качестве входных малошумящих усилителей. Как упоминалось выше, весьма важен вопрос выбора адекватной математической модели нелинейного элемента, в данном случае ПТШ. Ввиду сложности физических процессов, протекающих в структуре такого транзистора в диапазоне СВЧ, существует достаточно большое

число схем замещения, используемых в качестве математических моделей. Так, в пакете Microwave Office имеется четыре-пять таких моделей. Все они характеризуются наборами числовых параметров и могут быть использованы при моделировании. Среди разработчиков популярна модель Materka [2], поскольку она во многих случаях обеспечивает хорошее совпадение результатов расчёта и эксперимента. Схема замещения ПТШ, соответствующая этой модели, изображена на рис. 1.

Нелинейные свойства транзистора здесь отражены диодом, отображающим барьер Шоттки между затвором и каналом, нелинейными ёмкостями C_{gs} и C_{gd}, объёмным сопротивлением Ri, зависимыми источниками токов i_{ds} и i_{gd}. Также присутствуют активные сопротивления истока RS и стока RD, которые считаются линейными. Активная составляющая i_{gs} тока затвор–исток зависит от "внутреннего" напряжения v_{gsi} затвор–исток. Ток стока i_{ds} зависит от двух напряжений: v_{gsi} и внутреннего напряжения v_{dsi} сток–исток. Активная составляющая тока промежутка затвор–сток i_{gd} зависит от напряжения на нём v_{gdi}, равного разности напряжений $v1$ и v_{dsi}. Свойства модели Materka задаются с помощью более двадцати численных параметров, которые программными средствами Microwave Office могут быть найдены по экспериментальным характеристикам.

В настоящей статье использованы параметры транзистора NE713 фирмы NEC, приведенные в [2]. Выражения, описывающие нелинейные зависимости, имеются в материалах пакета Microwave Office и здесь не приводятся. Эквивалентная схема каскада, необходимая для расчётов, получается из рис. 1 подключением источника сигнала и нагрузки. Их сопротивления

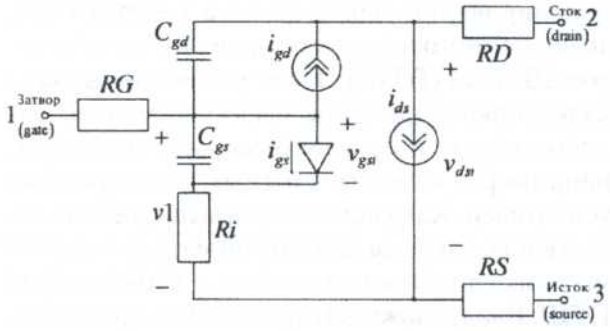


Рис. 1. Схема замещения ПТШ в соответствии с моделью Materka

будем считать в общем случае комплексными. При рассмотрении конкретных реализаций каскада это позволит учитывать в их составе блокировочные и развязывающие ёмкости и элементы цепей питания и смещения.

В большинстве случаев интересен анализ работы радиотехнических устройств в установившемся режиме. При этом используются методы анализа и моделирования, относящиеся к спектральным, например метод ГБ. Эквивалентная схема устройства представляется в виде одной из универсальных нелинейных моделей [3], в которых разделены линейная и нелинейная части. Для анализа линейной части применяют обычные методы теории цепей, а способ определения токов и напряжений в нелинейных элементах зависит от применяемого метода моделирования. В методе ГБ их записывают в виде укороченных рядов Фурье и решают систему алгебраических уравнений для коэффициентов. Решение получают путём итерационной процедуры, т. е. в виде последовательных приближений. В методе РВ нелинейные зависимости записывают в виде степенного ряда, при этом решение представляется как сумма нескольких членов функционального ряда Вольтерра. Ядра ряда при увеличении их порядка быстро становятся громоздкими, что ограничивает применимость метода случаями малой нелинейности, хотя принципиально ничто не препятствует применению ряда из сколь угодно большого числа слагаемых.

В [4] показано, что возможно решение уравнения нелинейной системы при произвольном функциональном представлении нелинейных зависимостей. Решению способствует метод последовательных приближений в виде функционального ряда. Условием сходимости такого ряда к точному решению является выполнение принципа сжимающих отображений, о котором будет сказано ниже. Подобный подход принципиально реализован в методе параметрических характеристик [1]. Для схемы, включающей только один нелинейный элемент, метод реализуется путём составления уравнения относительно напряжения на этом элементе, а решение получается методом последовательных приближений. Если схема более сложная, с несколькими нелинейными элементами, требуется составить систему уравнений относительно соответствующих напряжений. Для этого следует выработать до-

статочно общий алгоритм и использовать его при анализе устройств произвольной сложности. В методе ПХ нелинейные элементы представляются как параметрические, у которых параметры изменяются во времени под действием приложенных напряжений и протекающих токов. На каждом шаге устройство считается параметрическим, т. е. линейным. Поэтому в более общем случае логично перейти от исходной схемы, содержащей нелинейные резисторы, ёмкости и индуктивности, к линейной схеме, в которой нелинейные двухполюсники заменены согласно известной теореме компенсации источниками тока или ЭДС. В частности, нелинейные резисторы и конденсаторы заменяются источниками тока.

Применим указанный подход к моделированию усилителя на ПТШ. Подставив вместо резистивных и емкостных нелинейных элементов источники тока и добавив к схеме рис. 1 источник сигнала и нагрузку, получим эквивалентную схему рис. 2. На схеме использованы следующие обозначения, которых не было на рис. 1: $e_г(t)$ – ЭДС источника сигнала, $Z_г$ – сопротивление источника, $Z_н$ – сопротивление нагрузки, $i_{gs\Sigma}$ и $i_{gd\Sigma}$ – полные токи промежутков затвор–исток и затвор–сток, каждый из которых равен сумме активной и емкостной составляющих. На каждом шаге вычислений они, как и ток i_{ds}, считаются известными.

С целью сокращения выражений введём обозначения: $Z_{г\Sigma} = Z_г + RG$, $Z_{н\Sigma} = Z_н + RD$. Для анализа схемы можно использовать любой из методов теории линейных цепей. Часто отдают предпочтение методу контурных токов. Опустив для краткости уравнения, составленные по этому методу, приведём полученные из них выражения напряжений в операторной форме в области преобразования Лапласа (ПЛ):

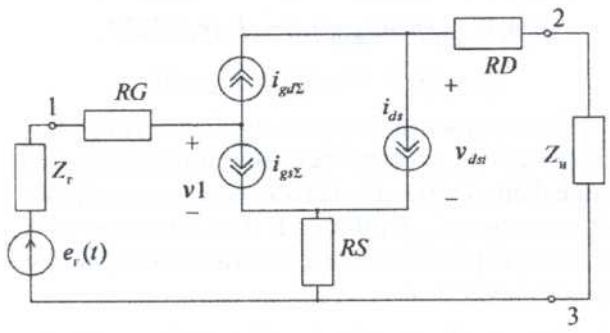


Рис. 2. Эквивалентная схема усилительного каскада на ПТШ

$$V1(p) = E_г(p) - [I_{gs\Sigma}(p) + I_{gd\Sigma}(p)]Z_{г\Sigma}(p) - [I_{gsS}(p) + I_{ds}(p)]RS; \quad (1)$$

$$V_{dsi}(p) = [I_{gd\Sigma}(p) - I_{ds}(p)]Z_{н\Sigma}(p) - [I_{gsS}(p) + I_{ds}(p)]RS; \quad (2)$$

$$V_{gdi}(p) = V1(p) - V_{dsi}(p). \quad (3)$$

Чтобы найти напряжения $V_{gsi}(p)$, нужно составить ещё одно уравнение, выразив его через напряжение $V1(p)$, равное сумме $V_{gsi}(p)$ и падения напряжения на нелинейном резисторе Ri. Здесь нельзя применить теорему компенсации, поэтому возьмем методику, использованную в [1] для случая последовательного включения элементов схемы. Уравнение – основные правила действий с такими уравнениями описаны в [1] – записывается в области многомерного преобразования Лапласа (МПЛ):

$$V_{gsi}(p_1, p_2, p_3, ...)\{1 + [G_{gs}(p_2)Ri(p_3) + p_1 C_{gs}(p_2)Ri(p_3)]\}, \quad (4)$$

где $G_{gs}(p)$ – изображение активной проводимости $g_{gs}(t) = i_{gs}(t)/v_{gsi}(t)$ промежутка затвор–сток, которая при нулевом напряжении равна

$$\mathrm{d}\, i_{gs}(v_{gsi}) \big/ \mathrm{d}\, v_{gsi}\Big|_{v_{gsi}=0}.$$

Взаимное соответствие оригинала и изображения обозначим символом $g_{gs}(t) \overset{L}{\longleftrightarrow} G_{gs}(p)$. Ёмкость $C_{gs}(t) = C_{gs}[v_{gsi}(t)]$ и резистор $Ri(t) = Ri[v_{gsi}(t)]$ также зависят от времени, их изображения

$$C_{gs}(t) \overset{L}{\longleftrightarrow} C_{gs}(p),\ Ri(t) \overset{L}{\longleftrightarrow} Ri(p).$$

Произведения изображений некоторых величин в (4), записанных с различными аргументами, соответствуют, по правилам действий с МПЛ, перемножению временных функций. Для удобства введём обозначения

$$a(t) = g_{gs}(t)Ri(t),\ b(t) = C_{gs}(t)Ri(t)$$

и представим каждую из этих функций в виде начального значения и переменной части

$$a(t) = g_{gs0}Ri_0 + a_{\sim}(t),\ b(t) = C_{gs0}Ri_0 + b_{\sim}(t).$$

Их изображения

$$A(p) = g_{gs0}Ri_0 + A_{\sim}(p),\ B(p) = C_{gs0}Ri_0 + B_{\sim}(p).$$

Заметим, что слагаемые $g_{gs0}Ri_0$ и $C_{gs0}Ri_0$ играют роль постоянных коэффициентов, поэтому в изображениях их не нужно записывать с множителем $1/p$. Тогда уравнение (4) можно переписать:

$$V_{gsi}(p_1, p_2, p_3, ...)\{1 + (g_{gs0} + s_1 C_{gs0})Ri_0 + + [A_{\sim}(p_2) + + p_1 B_{\sim}(p_2)]\} = V1(p_1). \quad (5)$$

На основе уравнений (1)–(5) можно построить процедуру последовательных приближений. В условиях, когда имеется неопределённость выбора начального приближения, целесообразно начинать с линеаризованной схемы. Нужно задать постоянные напряжения сток–исток v_{ds0} и внешнее смещение на затворе v_{gs0}, затем, пользуясь любым доступным методом, определить "внутренние" напряжения в состоянии покоя v_{gsi0}, v_{dsi0}, v_{gdi0}.

Далее следует найти значения нелинейных параметров и токов в точке покоя:

$$C_{gs0} = C_{gs}(v_{gsi0}),\ C_{gd0} = C_{gd}(v_{gdi0}),$$
$$i_{gs0} = i_{gs}(v_{gsi0}),\ i_{gd0} = i_{gd}(v_{gdi0}),$$
$$i_{ds0} = i_{ds}(v_{gsi0}, v_{dsi0}),\ Ri_0 = Ri(v_{gsi0}),$$

а также проводимости

$$G_{gs0} = \mathrm{d}\, i_{gs}(v_{gsi}) / \mathrm{d}\, v_{gsi}\Big|_{v_{gsi} = v_{gsi0}},$$

$$G_{gd0} = \mathrm{d}\, i_{gd}(v_{gdi}) / \mathrm{d}\, v_{gdi}\Big|_{v_{gdi} = v_{gdi0}}.$$

Зависимости $C_{gs}(v_{gsi})$, $C_{gd}(v_{gdi})$, $i_{gs}(v_{gsi})$, $i_{gd}(v_{gdi})$, $i_{ds}(v_{gsi}, v_{dsi})$, $Ri(v_{gsi})$ берутся из описания модели Materka.

Переменная составляющая тока стока в линеаризованном варианте выражается через крутизну и выходную проводимость:

$$S_0 = \partial i_{ds}(v_{gsi}, v_{dsi}) / \partial v_{gsi}\Big|_{v_{gsi} = v_{gsi0}, v_{dsi} = v_{dsi0}};$$

$$G22_0 = \partial i_{ds}(v_{gsi}, v_{dsi}) / \partial v_{dsi}\Big|_{v_{gsi} = v_{gsi0}, v_{dsi} = v_{dsi0}};$$

$$i_{ds\sim}(t) = \left[v_{gsi}(t) - v_{gsi0}\right] S_0 + \left[v_{dsi}(t) - v_{dsi0}\right] G22_0.$$

Методы анализа линейных цепей хорошо известны, поэтому не конкретизируем способ нахождения начального (нулевого) приближения, которое получается для линеаризованной схемы. Полагаем, что оно получено в виде выражений основных напряжений во временной области и их изображений в области ПЛ:

$$v1_{(0)}(t) \xleftrightarrow{L} V1_{(0)}(p),\ v_{dsi(0)}(t) \xleftrightarrow{L} V_{dsi(0)}(p),$$

$$v_{gdi(0)}(t) \xleftrightarrow{L} V_{gdi(0)}(p).$$

Нижний индекс в скобках означает номер приближения к напряжениям или номер версии законов, по которым изменяются нелинейные параметры схемы и зависимые токи. Напряжение v_{gsi0} получается из (5) приравниванием слагаемых первого порядка в правой и левой частях:

$$v_{gsi(0)}(t) \xleftrightarrow{L} V_{gsi(0)}(p) = V1_{(0)}(p)K(p),$$

где $K(p) = 1/(g_{gs0}Ri_0 + pC_{gs0}Ri_0)$.

По напряжениям находим начальные версии изменения параметров нелинейных элементов и зависимых токов схемы рис. 2:

$$C_{gs(0)}(t) = C_{gs}[v_{gsi(0)}(t)],\ C_{gd(0)}(t) = C_{gd}[v_{gdi(0)}(t)],$$
$$Ri_{(0)}(t) = Ri[v_{gsi(0)}(t)];$$
$$i_{gs\Sigma(0)}(t) = i_{gs(0)}(t) + C_{gs(0)}(t)\mathrm{d}[v_{gsi(0)}(t)]/\mathrm{d}t,$$

где $i_{gs(0)}(t) = i_{gs}[v_{gsi(0)}(t)]$;

$$i_{gd\Sigma(0)}(t) = i_{gd(0)}(t) + C_{gd(0)}(t)\mathrm{d}[v_{gdi(0)}(t)]/\mathrm{d}t,$$

где $i_{gd(0)}(t) = i_{gd}[v_{gdi(0)}(t)]$;

$$i_{ds(0)}(t) = i_{ds}[v_{gsi(0)}(t), v_{dsi(0)}(t)].$$

Переменным параметрам и токам ставятся в соответствие изображения $C_{gs(0)}(p)$, $C_{gd(0)}(p)$, $Ri_{(0)}(p)$, $I_{ds(0)}(p)$, а для первых двух токов они записываются в области МПЛ:

$$I_{gs\Sigma(0)}(p) = \{I_{gs(0)}(p_1) + p_1 V_{gsi(0)}(p_1) C_{gs(0)}(p_2)\}^*;$$
$$I_{gd\Sigma(0)}(p) = \{I_{gd(0)}(p_1) + p_1 V_{gdi(0)}(p_1) C_{gd(0)}(p_2)\}^*.$$

Символом $\{\circ\}^*$ обозначен переход от МПЛ к одномерному ПЛ. Затем на основе (1)–(3) могут быть найдены первые приближения к напряжениям, состоящие из слагаемых первого и второго порядков:

$$V1_{(1)}(p) = \{E_{\mathrm{T}}(p_1) - [I_{gs(0)}(p_1) + p_1 V_{gsi(0)}(p_1) C_{gs(0)}(p_2) + + I_{gd(0)}(p_1) + p_1 V_{gdi(0)}(p_1) C_{gd(0)}(p_2)] Z_{\mathrm{T}\Sigma}(p_1 + p_2) - - [I_{gs(0)}(p_1) + p_1 V_{gsi(0)}(p_1) C_{gd(0)} p_2 + I_{ds(0)}(p_1)] RS\}^*;$$

$$V_{dsi(1)}(p) = \{[I_{gd(0)}(p_1) + p_1 V_{gdi(0)}(p_1) C_{gd(0)}(p_2) - - I_{ds}(p_1)] Z_{\mathrm{H}\Sigma}(p_1 + p_2) - [I_{gs(0)}(p_1) + + p_1 V_{gsi(0)}(p_1) C_{gs(0)}(p_2) + I_{ds}(p_1)] RS\}^*;$$

$$V_{gdi(1)}(p) = V1_{(1)}(p) - V_{dsi(1)}(p).$$

Для получения первого приближения к напряжению v_{gsi} в соответствии с приведёнными выше формулами находятся начальные версии переменных $a_{\sim(0)}(t)$ и $b_{\sim(0)}(t)$ и их изображения. Затем из (5) берётся сумма слагаемых первого и второго порядков:

$$V_{gsi(1)}(\mathrm{p}) = \{V1_{(1)}(p_1)K(p_1) - V_{gsi(0)}(p_1)[(A_{\sim(0)}(p_2) + + p_1 B_{\sim(0)}(p_2))]K(p_1 + p_2)\}^*.$$

Далее, перейдя во временную область, на основе полученных напряжений вычисляются первые версии законов изменения параметров нелинейных элементов и зависимых токов, после чего аналогично первому получается второе приближение к напряжениям. При этом следует иметь в виду, что анализ ведётся в установившемся режиме, поэтому переход из временной области в частотную и обратно должен производиться с помощью пары преобразований Фурье, а не Лапласа.

Производя дальнейшие операции подобным образом, можно прийти к выражениям для произвольного k-го приближения:

$$V1_{(k)}(p) = \{E_r(p_1) - [I_{gs(k-1)}(p_1) + p_1 V_{gsi(k-1)} \times \times (p_1) C_{gs(k-1)}(p_2) + + I_{gd(k-1)}(p_1) + p_1 V_{gdi(k-1)}(p_1) C_{gd(k-1)} \times \times (p_2)] Z_{r\Sigma}(p_1 + p_2) - [I_{gs(k-1)}(p_1) + + p_1 V_{gsi(k-1)}(p_1) C_{gs(k-1)}(p_2) + I_{ds(k-1)}(p_1)] RS\}^*; \quad (6)$$

$$V_{dsi(k)}(p) = \{[I_{gd(k-1)}(p_1) + p_1 V_{gdi(k-1)}(p_1) C_{gd(k-1)} \times \times (p_2) - I_{ds}(p_1)] \times Z_{n\Sigma}(p_1 + p_2) - [I_{gs(k-1)}(p_1) + + p_1 V_{gsi(k-1)}(p_1) C_{gs(k-1)}(p_2) + I_{ds}(p_1)] RS\}^*; \quad (7)$$

$$V_{gdi(k)}(p) = V1_{(k)}(p) - V_{dsi(k)}(p); \quad (8)$$

$$V_{dsi(k)}(p) = \{V1_{(1)}(p_1) K(p_1) - V_{gsi(k-1)}(p_1) \times \times [(A_{-(k-1)}(p_2) + p_1 B_{-(k-1)}(p_2))] K(p_1 + p_2)\}^*. \quad (9)$$

Для наглядности перейдём во временную область:

$$v1_{(k)}(t) = e_r(t) - \int_0^t \Big\{ i_{gs}\left[v_{gsi(k-1)}(\tau)\right] + \mathrm{d}v_{gsi(k-1)}(\tau)/\mathrm{d}\tau \cdot C_{gs}\left[v_{gsi(k-1)}(\tau)\right] + + i_{gd}\left[v_{gdi(k-1)}(\tau)\right] + \mathrm{d}v_{gdi(k-1)}(\tau)/\mathrm{d}\tau \cdot C_{gd}\left[v_{gdi(k-1)}(\tau)\right]\Big\} \times \times z_{r\Sigma}(t-\tau)\mathrm{d}\tau - \Big\{ i_{gs}\left[v_{gsi(k-1)}(t)\right] + \mathrm{d}v_{gsi(k-1)}(t)/\mathrm{d}t \cdot C_{gs}\left[v_{gsi(k-1)}(t)\right] + + i_{ds}\left[v_{gsi(k-1)}(t), v_{dsi(k-1)}(t)\right]\Big\} RS; \quad (10)$$

$$v_{dsi(k)}(t) = \int_0^t \Big\{ i_{gd}\left[v_{gdi(k-1)}(\tau)\right] + \mathrm{d}v_{gdi(k-1)}(\tau)/\mathrm{d}\tau \times \times C_{gd}\left[v_{gdi(k-1)}(\tau)\right] - - i_{ds}\left[v_{gsi(k-1)}(\tau), v_{dsi(k-1)}(\tau)\right]\Big\} z_{n\Sigma}(t-\tau)\mathrm{d}\tau - \Big\{ i_{gs}\left[v_{gsi(k-1)}(t)\right] + + \mathrm{d}v_{gsi(k-1)}(t)/\mathrm{d}t \cdot C_{gs}\left[v_{gsi(k-1)}(t)\right] + + i_{ds}\left[v_{gsi(k-1)}(t), v_{dsi(k-1)}(t)\right]\Big\} RS, \quad (11)$$

где $z_{r\Sigma}(t)$ и $z_{n\Sigma}(t)$ – импульсные характеристики комплексных сопротивлений $Z_{r\Sigma}$ и $Z_{n\Sigma}$;

$$v_{gdi(k)}(t) = v1_{(k)}(t) - v_{dsi(k)}(t); \quad (12)$$

$$v_{gsi(k)}(t) = \int_0^t \Big\{ v1_{(k)}(\tau) - \left[v_{gsi(k-1)}(\tau) a_{-(k-1)}(\tau) + + \mathrm{d}v_{gsi(k-1)}(\tau)/\mathrm{d}\tau \cdot b_{-(k-1)}(\tau)\right] h(t-\tau)\mathrm{d}\tau, \quad (13)$$

где $h(t)$ – импульсная характеристика, являющаяся оригиналом $K(p)$.

Если произвести подстановку (10) в (12) и (13), а также (11) в (13), получим три уравнения, которые обобщенно можно представить следующим образом:

$$v_{gsi(k)}(t) = \Phi 1[v_{gsi(k-1)}(t), v_{dsi(k-1)}(t), v_{gdi(k-1)}(t)]; \quad (14)$$

$$v_{dsi(k)}(t) = \Phi 2[v_{gsi(k-1)}(t), v_{dsi(k-1)}(t), v_{gdi(k-1)}(t)]; \quad (15)$$

$$v_{gdi(k)}(t) = \Phi 3[v_{gsi(k-1)}(t), v_{dsi(k-1)}(t), v_{gdi(k-1)}(t)]. \quad (16)$$

Левые части равенств представляют собой основные напряжения в эквивалентной схеме рис. 2, а правые части имеют вид функционалов Ф1, Ф2, Ф3, в составе которых имеются как линейные, так и нелинейные операции над этими напряжениями. И левые и правые части можно представить как векторы, состоящие из трёх компонент каждый. Уравнение примет вид $\mathbf{V} = \Phi[\mathbf{V}]$. Согласно теории функционального анализа [4] вектор функционалов отображает вектор напряжений в то же пространство, в котором он определен. Решение уравнения может быть найдено методом последовательных приближений, если выполняется принцип сжимающих отображений [4]. Упрощенно этот принцип можно толковать следующим образом: он выполняется, если малому приращению вектора напряжений $\mathbf{V}$ соответствует меньшее (по абсолютному значению) приращение вектора функционалов $\Phi[\mathbf{V}]$. Практически последовательные приближения компонент вектора $\mathbf{V}$ могут быть вычислены в соответствии с выражениями (14)–(16).

Вопрос проверки выполнения принципа сжимающих отображений, а следовательно, сходимости алгоритма последовательных приближений достаточно сложен и требует отдельного рассмотрения в каждом случае. Однако существует простой приём, известный из теории итерационных вычислений, позволяющий обеспечить сходимость. Он сводится к искусственному уменьшению приращения отображения. Для этого выражения (14)–(16) следует изменить следующим образом:

$$v_{gsi(k)}(t) = \Delta \cdot \Phi 1[v_{gsi(k-1)}(t), v_{dsi(k-1)}(t), v_{gdi(k-1)}(t)] + \\ + (1 - D)v_{gsi(k-1)}(t); \quad (17)$$

$$v_{dsi(k)}(t) = D \cdot \Phi 2[v_{gsi(k-1)}(t), v_{dsi(k-1)}(t), v_{gdi(k-1)}(t)] + \\ + (1 - D)v_{dsi(k-1)}(t); \quad (18)$$

$$v_{gdi(k)}(t) = \Delta \cdot \Phi 3[v_{gsi(k-1)}(t), v_{dsi(k-1)}(t), v_{gdi(k-1)}(t)] + \\ + (1 - D)v_{gdi(k-1)}(t), \quad (19)$$

где, выбирая коэффициент $\Delta < 1$, можно обеспечить сходимость в большинстве практических случаев. Выбор величины Δ может производиться различными методами. В данной статье эта задача не рассматривалась, а выбор Δ производился посредством проб.

Были сделаны расчёты характеристик усилительного каскада на ПТШ предложенным методом. Для сравнения велись вычисления в пакете Microwave Office методами ГБ и РВ. Расчёты проводились для частоты сигнала 10 ГГц при напряжении питания $v_{ds0} = 3$ В и смещении затвор–исток $v_{gs0} = 0{,}4$ В. На рис. 3 показана передаточная характеристика, т. е. зависимость уровня выходной мощности от уровня мощности входного сигнала. На рис. 4 изображена зависимость уровня продукта интермодуляции третьего порядка от уровня сигнала на входе.

После расчётов можно сделать следующие выводы:

результаты расчётов передаточной характеристики (см. рис. 3) методами ГБ и РВ практически повторяют приведённые в [2]. Отличие наблюдается для метода РВ в области входной мощности больше 9 дБм, что объясняется применением для этого метода в [2] другой модели;

результаты расчётов передаточной характеристики предложенным методом практически совпадают до уровня входной мощности 9 дБм с результатами метода ГБ. При большей входной мощности имеется отличие результатов на (0,5–0,6) дБ;

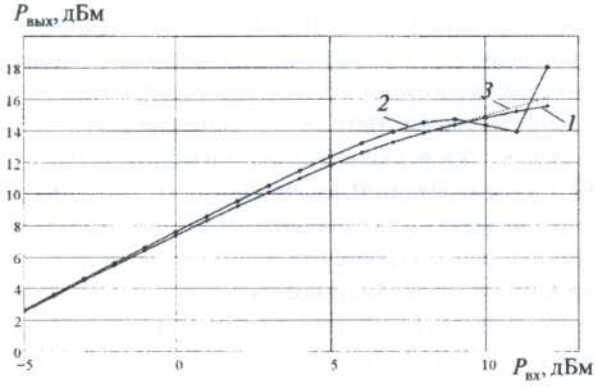


Рис. 3. Передаточная характеристика по мощности
Методы: ГБ (*1*), РВ (*2*) и предложенный (*3*)

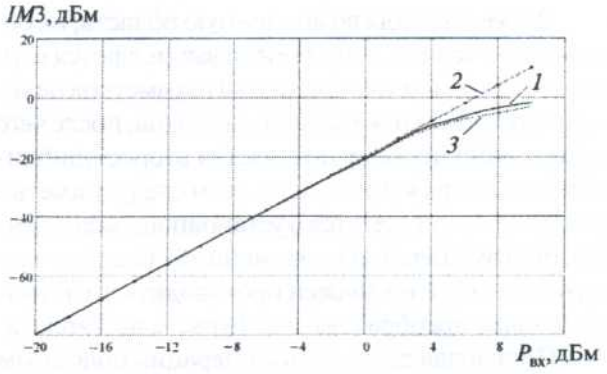


Рис. 4. Уровень интермодуляции третьего порядка
Методы: ГБ (*1*), РВ (*2*) и предложенный (*3*)

результаты расчётов уровня интермодуляционной помехи третьего порядка ИМ3 (см. рис. 4) всеми тремя методами очень близки до уровня входной мощности примерно –6 дБм. При большей мощности метод РВ даёт явно нереальные результаты, характеризующиеся линейной зависимостью;

результаты расчётов помехи ИМ3 предложенным методом и методом ГБ практически совпадают до уровня входной мощности 0 дБм, после чего появляется отличие до (1,5–1,6) дБ.

Таким образом, результаты, полученные предложенным методом моделирования, практически совпадают с результатами метода ГБ при входной мощности, соответствующей небольшой и средней степени нелинейности. Различие появляется в области значительной нелинейности. При этом есть основания полагать, что менее точные результаты даёт метод ГБ, поскольку он построен на заведомо приближенном алгоритме. Следует заметить, что в рассмотренном примере не проявилось приписываемое методу ГБ увеличение погрешности вычисления нелинейных продуктов при малых уровнях входной мощности. Так, результаты вычисления продукта интермодуляции третьего порядка всеми тремя рассмотренными методами практически совпадают для входных уровней, не превышающих –6 дБм.

Итак, предложен метод анализа и моделирования нелинейных устройств, основанный на решении уравнения системы по принципу последовательных приближений. Для разделения действий над линейной и нелинейной частями схемы применена теорема компенсации из теории цепей. Приведённые результаты расчётов для схемы усилительного каскада на ПТШ показывают хорошее совпадение с результатами существующих методов.

СПИСОК ЛИТЕРАТУРЫ

1. **Ненашев А.В.** Разработка метода параметрических характеристик для моделирования нелинейных радиотехнических устройств // Научно-технические ведомости СПбГПУ. Информатика, телекоммуникации, управление. 2008. № 6 (69). С. 156–163.

2. **Аверина Л.И., Усков Г.К.** Определение параметров нелинейных моделей полевого транзистора в системе схемотехнического СВЧ проектирования Microwave Office // Радиолокация, навигация, связь: Труды IX междунар. науч.-техн. конф. / ООО "Саквоее". Т. 1. Воронеж, 2003. С. 457–465.

3. **Алексеев О.В., Асович П.Л., Соловьёв А.А.** Спектральные методы анализа нелинейных радиоустройств с помощью ЭВМ. М.: Радио и связь, 1985. 152 с.

4. **Пупков К.А., Капалин В.И., Ющенко А.С.** Функциональные ряды в теории нелинейных систем. М.: Наука, 1976. 448 с.

УДК 004.421.4

В.А. Карелин, И.В. Карелин

ПОЛУЧЕНИЕ МАКСИМАЛЬНО ПРАВДОПОДОБНЫХ ОЦЕНОК МАЛЫХ ОТКЛОНЕНИЙ ЧАСТОТ ВЫСОКОСТАБИЛЬНЫХ ГАРМОНИЧЕСКИХ СИГНАЛОВ

Показатели качества различных радиотехнических систем, таких, как точность действия, разрешающая и пропускная способность, помехозащищенность, в значительной степени определяются метрологическими характеристиками задающих генераторов (хронизаторов), роль которых выполняют меры частоты [1]. Метрологические параметры частоты сигналов задающих генераторов определяются путем сравнения с частотой сигнала эталонной меры. В связи с достаточно малой относительной нестабильностью частоты мер ($\sim 10^{-10}$–10^{-14}) предъявляются чрезвычайно высокие требования к измерительным устройствам, которые должны быть прецизионными, с характеристиками, близкими к потенциальным пределам, зависящими от эффективности алгоритмов обработки сигналов и достигаемой минимизации погрешностей оценок разности фаз и частот сличаемых сигналов. Теоретически оптимальными являются оценки, полученные на основе марковской теории нелинейной фильтрации с помощью ММП, но их реализация приводит к сложным корреляционным и многоканальным устройствам [2, 3]. Поэтому на практике широкое распространение получили упрощенные алгоритмы, использующие малоэффективные среднеинтегральные оценки. При этом для повышения разрешающей способности осуществляется аналоговое умножение разности сличаемых частот и перенос ее с помощью гетеродинирования на низкую частоту – подставку [1], что искажает информацию об измеряемом параметре.

Для синтеза измерителей, обеспечивающих оптимальные характеристики в установившемся режиме при стационарных помехах, возможен выбор дискриминатора и синтез весовой функции усредняющего фильтра, обеспечивающего максимально правдоподобную оценку частоты, методами линейной фильтрации Винера. Правомерность такого подхода подтверждается еще и тем, что нелинейная зависимость входного сигнала от оцениваемого параметра разрушается именно в дискриминаторе. При этом возможен синтез алгоритма сличения частот – несложного для реализации и обладающего максимальным правдоподобием.

Цель исследования – решение задачи синтеза оптимальной по критерию максимального правдоподобия весовой функции усредняющего фильтра для оценки средней разности частот сигналов задающих генераторов при квадратичной функции потерь.

В качестве модели сигнала исследуемой меры частоты рассмотрим аддитивную смесь гармонического сигнала $s(t)$ и случайного узкополосного процесса $\xi(t)$.

$$X(t) = s(t) + \xi(t) = U_x\cos[(\omega_0 + \Delta\omega)t + \varphi] + \\ + A(t)\cos[(\omega_0 + \Delta\omega)t + \theta(t)] = \\ = U(t)\cos[(\omega_0 + \Delta\omega)t + \varphi(t)] = U(t)\cos\Phi(t), \quad (1)$$

где ω_0, $\Delta\omega$, φ – угловая частота сигнала опорной меры, отклонение частоты и начальная фаза измеряемого сигнала; $A(t)$, $\theta(t)$ – огибающая и фаза случайного процесса $\xi(t)$; $U(t)$, $\varphi(t)$ и $\Phi(t)$ – огибающая, случайная фаза и полная фаза аддитивной смеси.

Модель сигнала $S_0(t)$ опорной меры частоты можно выразить в виде

$$S_0(t) = U_0\cos\omega_0 t. \quad (2)$$

Пусть в качестве дискриминатора выбран измеритель разности фаз $\Delta\Phi(t)$, работающий в соответствии с известным алгоритмом [4]:

$$\Delta\Phi(t) = \Phi(t) - \omega_0 t = \operatorname{arctg}\frac{\int_0^T X(t)\cdot\sin\omega_0 t\,dt}{\int_0^T X(t)\cdot\cos\omega_0 t\,dt}, \quad (3)$$

где T – интервал времени усреднения.

Мгновенная разность частот $\Delta\omega(t)$, среднее значение которой подлежит оценке, связана с $\Delta\Phi(t)$ соотношением

$$\Delta\omega(t) = \frac{d\Delta\omega(t)}{dt} = \Delta\omega + \frac{d\Delta\Phi(t)}{dt} = \Delta\omega + \upsilon(t). \quad (4)$$

Случайная разность частот $\upsilon(t)$ определяет шумовую составляющую погрешности измерения среднего значения $\Delta\omega(t)$, а характер изменения ее спектрально-корреляционных характеристик существенно влияет на дисперсию результата усреднения.

Оценка математического ожидания $\Delta\omega(t)$ получается с помощью алгоритма [5]

$$\overline{\omega}(t) = \int_0^T \Delta\omega(t)\,g(t)\,dt, \quad (5)$$

где $g(t)$ – весовая функция усредняющего фильтра, зависящая от корреляционных характеристик случайного процесса $\Delta\omega(t)$ и удовлетворяющая условию несмещенности оценки

$$\int_0^T g(t)\,dt = 1. \quad (6)$$

Определение весовой функции $g(t)$, обеспечивающей минимальную дисперсию оценки (5), связано с решением интегрального уравнения [5], точное решение которого возможно только для ограниченного числа частных случаев бесконечного времени наблюдения и модели СПМ помехи в виде дробно-рациональной функции [5].

В настоящей статье предлагается алгоритм определения весовой функции усредняющего фильтра, позволяющий избежать указанных трудностей.

Предположим, что искомая функция $g(t)$ дифференцируема как минимум один раз.

Учитывая (4) и (5), можно выразить погрешность оценки

$$\Delta = \int_0^T \upsilon(t)\,g(t)\,dt. \quad (7)$$

Тогда оператор (7) может быть переписан следующим образом:

$$\Delta = \int_0^T g(t)\,d[\varphi(t)] = g(T)\varphi(T) - g(0)\varphi(0) - \int_0^T \dot{g}(t)\varphi(t)\,dt. \quad (8)$$

Чтобы учесть условия несмещенности (6), представим искомую функцию $g(t)$ в таком виде:

$$g(t) = \begin{cases} \dfrac{g_1(t)}{\int_0^T g(t)\,dt} = c\,g_1(t) & 0 \le t \le T; \\ 0 & t < 0; \quad t > T. \end{cases} \quad (9)$$

Подставив (9) в выражение (8), получим

$$\Delta = c\int_0^T [\delta(t-T)g_1(t) - \delta(t)g_1(t) - \dot{g}_1(t)]\,\varphi(t)\,dt, \quad (10)$$

где $\delta(t)$ – дельта-функция Дирака.

Запишем выражение для дисперсии $\sigma^2(\Delta)$ оценки:

$$\sigma^2(\Delta) = c^2\left\langle \begin{array}{c} \int_0^T\int_0^T [\delta(r-T)\,g_1(r) - \\ -\delta(r)\,g_1(r) - \dot{g}(r)]\times \\ \times[\delta(s-T)\,g_1(s) - \delta(s)\,g_1(s) - \\ -\dot{g}_1(s)]\varphi(r)\,\varphi(s)\,dr\,ds \end{array} \right\rangle, \quad (11)$$

где угловые скобки означают операцию статистического усреднения.

Преобразование (11) с учетом свойства некоррелированности стационарного случайного

процесса и его производной в совпадающие моменты времени, а также фильтрующего свойства дельта-функции позволяет записать

$$\sigma^2(\Delta)=\left\{\begin{array}{c}\sigma_0^2[g_1^2(T)-2g_1(T)\,g_1(0)R(0,T)+g_1^2(0)]+\\ +\left\langle\int\limits_0^T\int\limits_0^T\dot g_1(r)\dot g_1(s)\,\varphi(r)\varphi(s)\,dr\,ds\right\rangle\end{array}\right\}=$$

$$=c^2\left\{\begin{array}{c}\sigma_0^2[\dot g_1^2(T)-2g_1(T)g_1(0)R(0,T)+g_1^2(0)]+\\ +\int\limits_0^T\int\limits_0^T\dot g_1(r)\dot g_1(s)\langle\varphi(r)\varphi(s)\rangle\,dr\,ds\end{array}\right\}= \quad (12)$$

$$=c^2\sigma_0^2[g_1(T)-2g_1(T)g_1(0)R(0,T)+g_1^2(0)+$$

$$+\int\limits_0^T\int\limits_0^T\dot g_1(r)\dot g_1(s)R(r,s)\,dr\,ds],$$

где σ_0^2 и $R(r, s)$ – дисперсия и коэффициент корреляции случайного процесса $\varphi(t)$.

В силу стационарности процесса $\varphi(t)$ $R(r, s) = R(r-s)$, следовательно,

$$\int\limits_0^T\int\limits_0^T\dot g_1(r)\dot g(s)R(r-s)\,dr\,ds=$$

$$=2\int\limits_0^T R(p)\int\limits_0^T\dot g_1(t)\dot g_1(t+p)\,dt\,dp.$$

Поэтому выражение (12) может быть переписано:

$$\sigma^2(\Delta)=c^2\sigma_0^2\left[\begin{array}{c}g_1^2(T)-2g_1(T)g_1(0)R(T)+\\ +g_1^2(0)+2\int\limits_0^T\dot g_1(t)\dot g_1(t+p)\,dt\,dp\end{array}\right]. \quad (13)$$

Для преобразования (13) получим следующие очевидные соотношения:

$$g_1(T)=g_1(0)+\int\limits_0^T\dot g_1(t)\,dt; \quad (14)$$

$$c=\frac{1}{\int\limits_0^T g_1(t)\,dt}=\frac{1}{\left[g_1(0)T+\int\limits_0^t\dot g(t)(T-t)\,dt\right]}. \quad (15)$$

Подставив (14) и (15) в (13), получим следующее выражение дисперсии оценки:

$$\sigma^2(\Delta)=\sigma_0^2\frac{2g_1^2(0)[1-R(T)]+2g_1(0)}{\left[g_1(0)T+\int\limits_0^T\dot g_1(t)[T-t]\,dt\right]^2}\times$$

$$\times\frac{[1-R(T)]\int\limits_0^T g_1(t)+\left[\int\limits_0^T\dot g_1(t)\,dt\right]^2}{\left[g_1(0)T+\int\limits_0^T\dot g_1(t)[T-t]\,dt\right]^2}+$$

$$+\frac{2\int\limits_0^T R(p)\int\limits_0^{T-p}\dot g_1(t)g_1(t+p)\,dt\,dp}{\left[g_1(0)T+\int\limits_0^T\dot g_1(t)[T-t]\,dt\right]^2}. \quad (16)$$

Выражение (16) представляет из себя функционал, зависящий от $g_1(0)$ и $\dot g_1(t)$ (время наблюдения T считаем фиксированным). С целью определения минимума функционала (16) продифференцируем его по $\dot g_1(t)$ и $g_1(0)$ и, приравняв частные производные к нулю, получим систему из двух уравнений:

$$2g_1(0)[1-R(T)]=aT-[1-R(\tau)]\int\limits_0^T\dot g_1(t)\,dt; \quad (17)$$

$$\int\limits_0^T R(p)\dot g_1(t+p)=a[T-t]-g_1(0)[1-R(\tau)]-\int\limits_0^T\dot g_1(t)\,dt; \quad (18)$$

где $a=\dfrac{2g_1^2(0)[1-R(\tau)]+2g_1(0)}{g_1(0)+\int\limits_0^T\dot g_1(t)[T-t]\,dt}\times$

$$\times\frac{[1-R(T)]\int\limits_o^T\dot g_1(t)\,dt\left[\int\limits_0^T\dot g_1(t)\,dt\right]^2}{g_1(0)+\int\limits_0^T\dot g_1(t)[T-t]\,dt}. \quad (19)$$

Объединив уравнения (17) и (18), имеем:

$$\int\limits_0^T R(p)\dot g_1(t+p)\,dp+a\left(\frac{T}{2}-t\right)-\frac{1}{2}[1+R(T)]\int\limits_0^T\dot g_1(t)\,dt. \quad (20)$$

Заметив, что коэффициент a (см. (19)) не зависит от t, продифференцируем обе части выражения (20):

$$\int\limits_0^T R(p)\ddot g_1(t+p)\,dp=-a. \quad (21)$$

Левая часть (21) не будет зависеть от t только при условии: $\ddot g_1(t)=\text{const}$.

Поэтому $\ddot g_1\int\limits_0^T R(p)\,dp=-a$, или

$$\ddot{g}_1 = -\frac{a}{\int_0^T R(p)\,dp}. \qquad (22)$$

Таким образом, оптимальное значение производной весовой функции $g(t)$ можно найти, проинтегрировав дифференциальное уравнение (22):

$$\dot{g}_1(t) = C + \frac{at}{\int_0^T R(p)dp}, \qquad (23)$$

где C – постоянная интегрирования, которую определим, подставив (23) в (20):

$$C = 2a\left\{\frac{\dot{T}}{2} + \frac{T^2[1+R(t)]}{4^T\int_0^T R(p)dp} + \frac{\int_0^T pR(p)dp}{\int_0^T R(p)dp}\right\} \Big/ \Big/ \left\{\int_0^T R(p)dp + T[1+R(T)]\right\}. \qquad (24)$$

Выражения (23) и (24) определяют параметры оптимальной весовой функции в зависимости от коэффициента корреляции и времени наблюдения T.

Для сравнения эффективности полученного алгоритма с алгоритмом, синтезированным методом теории нелинейной фильтрации, рассмотрим некоторые практически важные случаи.

1. Модель СПМ $\varphi(t)$ описывается выражением

$$S_\varphi(\omega) = \frac{\omega_0\,\alpha}{\omega^2 + \alpha^2}.$$

Например, при фазовых флуктуациях смеси сигнала и аддитивного шума, прошедших линейную цепь, $R(p) = e^{-\alpha(P)}$. Полагая $1/\alpha \le T$, из (24) получаем $C = b/2$, где $b = T/\tau_k$, τ_k – время корреляции случайного процесса $\varphi(t)$.

Из (9) с учетом (22) получим выражение для оптимальной весовой функции:

$$g(t) = \begin{cases} \dfrac{[T + bt(1 - t/T)]}{T(1 + b/6)}, 0 \le t \le T; \\ 0, t < 0; t > T. \end{cases} \qquad (25)$$

Из (16) определим дисперсию оценки $\sigma^2(\Delta)$ разности частот $\Delta\omega$:

$$\sigma^2(\Delta) = \frac{2\sigma_0^2}{T^2(1 + b/6)}. \qquad (26)$$

Сравнение (26) с дисперсией широко распространенной среднеинтегральной оценки $\sigma^2(\Delta) = \dfrac{2\sigma_0^2}{T^2}$ дает выигрыш в точности в $(1 + b/6)$ раз, т. е. при $b >> 1$ достигает порядка и более.

2. Модель СПМ $\varphi(t)$ равномерная (белый шум). В этом случае имеем:

$$g(t) = \begin{cases} \dfrac{T + t(T - t)}{T(1 + T/6)}, 0 \le t \le T; \\ 0, t < 0; t > T; \end{cases}$$

$$\sigma^2(\Delta) = \frac{2\sigma_0^2}{T^2(1 + T/6)}. \qquad (27)$$

Для сравнения эффективности синтезированного алгоритма с алгоритмом нелинейной фильтрации выразим (27) через отношение сигнал/шум. Известно [4], что $\sigma_0^2 = \dfrac{N_0}{2E_c}$, где N_0 – энергия аддитивной помехи, E_c – энергия сигнала за время наблюдения. Тогда

$$\sigma^2(\Delta) = \frac{N_0}{E_c T^2(1 + T/6)},$$

или при $T/6 >> 1$

$$\sigma^2(\Delta) = \frac{6N_0}{E_c T^3},$$

т. е. совпадает с дисперсией оценки алгоритма теории нелинейной фильтрации [3] для дискретного времени при предположении некоррелированности отсчетов смеси сигнала и шума и симметричности спектральной модели СПМ помехи.

3. Необходимо заметить, что при выводе выражения для оптимальной весовой функции усредняющего фильтра единственным допущением относительно случайного процесса $\varphi(t)$ была его стационарность. Но можно показать, что с помощью синтезированного алгоритма можно получить конструктивные результаты на классе случайных процессов $\varphi(t)$, спектр мощности которых обладает бесконечной интенсивностью в области низких частот, например винеровский процесс.

Для определения модели $R(p)$ в этом случае воспользуемся представлением процесса $\varphi(t)$ в виде квазистационарного случайного процесса с бесконечной памятью. Согласно [6] коэффициент корреляции такого процесса может быть записан в виде $R(p) = 1 - kp$, где k – const, и для конечных T справедливо

$$kT << 1. \qquad (28)$$

Из (9) и (22) с учетом (28) получим

$$g(t) = \begin{cases} 1/T,\ 0 \le t \le T; \\ 0,\ t < 0; t > T. \end{cases}$$

То есть получим среднеинтегральный алгоритм усреднения, приводящий, как известно [7], к оптимальному результату в случае "белого шума", моделью которого в данном примере описывается сглаживаемый процесс $\upsilon(t)$, получаемый дифференцированием $\varphi(t)$.

Заметим, что полученный алгоритм синтеза оптимальной весовой функции усредняющего фильтра инвариантен относительно типа дискриминатора, один из примеров которого приведен в данной статье.

На основе проведенного исследования разработано цифровое устройство [8], позволяющее получать высокоточные оценки малых отклонений частоты, не прибегая к преобразователям, которые повышают разрешающую способность, но существенно искажают информацию об оцениваемом параметре исследуемого сигнала.

С использованием вариационного подхода получен алгоритм синтеза оптимальной (по критерию минимума среднего квадрата ошибки) весовой функции усредняющего фильтра для оценки малых отклонений частоты сигналов задающих высокостабильных генераторов. Показано, что оценки на выходе синтезированных фильтров имеют наивысшую точность, совпадающую с точностью оценок, получаемых с помощью классической теории нелинейной фильтрации, но требуют для своей реализации гораздо меньше аппаратных и программных средств.

Рассмотренная методика синтеза усредняющих фильтров с целью получения конструктивных результатов не требует предположения о некоррелированности помехи, так как в выражения для параметров оптимальной весовой функции в явном виде входит коэффициент корреляции паразитных фазовых флуктуаций.

СПИСОК ЛИТЕРАТУРЫ

1. **Виленкин С.Я.** Статистические методы исследования стационарных поцессов и систем автоматического регулирования. М.: Советское радио, 1967. 120 с.

2. Аппаратура для частотных и временных измерений / Под ред. А.П. Горшкова. М.: Советское радио, 1971.

3. **Тихонов В.И.** Оптимальный прием сигналов. М.: Радио и связь, 1983. 320 с.

4. Цифровые радиоприемные системы: Справочник / М.И. Жоддишский и др.; Под. ред. М.И. Жоддишского. М.: Радио и связь, 1990. 208 с.

5. **Пестряков В.Б.** Фазовые радиотехнические системы (основы статистической теории) М.: Советское радио, 1968. 468 с.

6. **Кеннер М.С.** Шум типа 1/f // ТИИЭР. 1982. Т. 70, № 2.

7. **Тихонов В.И.** Статистическая радиотехника. 2-е изд. М.: Радио и связь, 1982.

УДК 519.714.23

В.Н. Коновалов, И.В. Коновалов

ПОДХОД К МНОГОУРОВНЕВОЙ ОПТИМИЗАЦИИ КОМБИНАЦИОННЫХ ЛОГИЧЕСКИХ СХЕМ

Один из завершающих этапов логического проектирования цифровых схем – покрытие или технологическое отображение (technology mapping), т. е. преобразование технологически независимой схемы, полученной в результате многоуровневой декомпозиции и детализации и состоящей из различных двухвходовых логических блоков, в эквивалентную схему в целевом технологическом базисе [3, 5, 6]. В современных интегральных схемах таким технологическим базисом могут служить двухвходовые элементы И-НЕ либо ИЛИ-НЕ, реализованные в виде базовых матричных кристаллов или программируемых пользователем вентильных матриц (ППВМ

или FPGA – Field Programmable Gate Arrays). Один из вариантов конфигурируемых логических блоков ППВМ представляет собой регулярные топологические структуры на основе последовательно соединённых МОП-транзисторов с каналами *n*- и *p*-типа (SLC – Simple Logic Cells). Между цепочками таких транзисторов расположены трассировочные каналы, в которых реализуются необходимые межсоединения транзисторов [8].

Основными критериями оптимизации при многоуровневом синтезе являются:

1) длина критического пути при прохождении сигнала от входа к выходу схемы (максимальная задержка сигнала);

2) площадь кристалла, занимаемая логической сетью. От указанной площади в конечном итоге зависит стоимость проектируемой схемы. Причём при определении площади сети часто учитывается не только площадь логических элементов, но и площадь межсоединений [1, 2]. Это условие ставит не только задачу минимизации количества узлов логической сети (equivalent gate), но и количества связей между ними (slices).

Как правило, оптимизацию комбинационных схем связывают с минимизацией систем булевых функций. Методы решения указанной задачи классические и заключаются в нахождении ДНФ, минимальной по числу литералов (минимальной ДНФ), либо по числу конъюнкций (кратчайшей ДНФ) [2, 4, 7].

Рассматриваемый ниже алгоритм оптимизации ориентирован на многоуровневые логические схемы и инвариантен по отношению к логическому базису. Процесс оптимизации включает выполнение двух этапов:

1) исключение из схемы очевидных избыточных элементов;

2) анализ схемы с учётом логической недоопределённости.

Рассмотрим более подробно выполнение указанных этапов.

Предварительная оптимизация схемы

На первом (предварительном) этапе оптимизации производится исключение из схемы цепочек последовательно соединенных инверторов и элементов, выполняющих одинаковые логические функции.

Цепочки инверторов могут образовываться в схемах при покрытии элементами 2И-НЕ либо 2ИЛИ-НЕ, поскольку для совпадения числовых последовательностей при формальном замещении покрываемых блоков требуется инвертирование входных или выходных последовательностей покрывающих элементов [6].

Элементы, выполняющие одинаковые логические функции, как правило, оказываются подключенными к одноимённым входам схемы. В таком случае на их выходах формируются одинаковые логические последовательности, и один из рассматриваемых элементов может быть удалён, а выходной сигнал оставшегося элемента используется для подачи на оба узла синтезируемой схемы.

Указанные алгоритмы первого этапа оптимизации достаточно очевидны и не требуют строгих доказательств и пояснений.

Анализ схемы с учётом логической недоопределённости

В начале второго этапа оптимизации проводится полный анализ схемы. Он позволяет получить числовые последовательности на выходах всех элементов и ещё раз убедиться в правильности предшествующих процедур синтеза и первого этапа оптимизации. Далее от выхода к входу выполняется анализ отдельных элементов с учётом возможной логической недоопределённости их входных числовых последовательностей. При таком анализе в поступающие на входы логических элементов числовые последовательности искусственно вводится неопределённость (звёздочки). Если, например, на входы элемента 2ИЛИ поступают две какие-либо последовательности, то одну из них можно фиксировать, а во вторую ввести звёздочки на тех местах, на которых в первой последовательности стоят логические единицы. Ведь если на один из входов элемента 2ИЛИ подаётся единица, то на его выходе будет присутствовать единица независимо от состояния на втором входе.

Далее полученная таким образом последовательность со звёздочками сравнивается со всеми числовыми последовательностями, реализуемыми на выходах других логических элементов схемы. Может оказаться так, что рассматриваемая недоопределённая последовательность не противоречит какой-либо из них. В этом случае выход найденного элемента соединяется с рассматриваемым входом анали-

зируемого элемента. Кроме того, при проведении сравнения можно осуществлять поиск и непротиворечивых инверсных последовательностей. В данном случае на соответствующий вход анализируемого элемента найденная последовательность будет подаваться через инвертор.

Оказавшиеся теперь лишними логические элементы, которые раньше участвовали в формировании анализируемой входной последовательности, могут быть исключены из логической схемы.

Если подходящая числовая последовательность не найдена, то последовательность проанализированного входа фиксируется и неопределённость вводится в последовательность другого входа.

При анализе входов элемента 2И всё делается точно так же, но звёздочки в анализируемой последовательности вводятся на тех местах, на которых фиксированная последовательность принимает нулевое значение. Аналогичные правила формирования недоопределённых входных последовательностей можно сформулировать и для других логических элементов (2И-НЕ, 2ИЛИ-НЕ и т. д.).

Однако следует заметить, что при выполнении рассмотренных выше операций, связанных с изменением топологии схемы, нельзя допускать образования петель обратной связи, которые могут превратить комбинационное устройство в конечный автомат с памятью. То есть если найдена непротиворечивая последовательность у далее стоящего элемента и он подчинён исследуемому, то эти элементы соединять нельзя.

Основные идеи описанной выше методики оптимизации комбинационных логических схем были сформулированы более 30 лет назад доцентом кафедры ЭВМ Калужского филиала МВТУ им. Н.Э. Баумана Юрием Павловичем Голубевым. После смерти автора в 1990 году значительная часть материалов его исследований оказалась утраченной и неопубликованной. Предлагаемый алгоритм многоуровневой оптимизации представляет собой дальнейшее развитие и усовершенствование идей Ю.П. Голубева.

Многоуровневая оптимизация логических схем

Во многих случаях, особенно при анализе сложных разветвлённых схем, целесообразно проводить многоуровневую оптимизацию, которая позволяет увеличить количество возможных недоопределённых элементов в анализируемых последовательностях. Это в конечном итоге значительно расширяет возможности подбора непротиворечивых последовательностей. При этом анализируемые последовательности можно сравнивать ещё и с входными переменными, их инверсиями или логическими константами.

Процедуру многоуровневой оптимизации удобно рассмотреть на примере схемы мультиплексора 2/1, характеризуемого числовой последовательностью 0011 0101 и представляющего собой наиболее общий вариант трёхвходового логического блока, выделяемого из схемы в процессе детализации [3]. Один из вариантов схемы рассматриваемого мультиплексора представлен на рис. 1.

Выполнения первого этапа оптимизации для рассматриваемой схемы не требуется, поэтому перейдём сразу ко второму этапу и проведём её полный анализ [5]:

4_0	0000	1111	4_0	0000	1111
1_0	0101	0101	1_0	0101	0101
A	0000	0101	B	1111	0101
2_0	0011	0011	A	0000	0101
B	1111	0101	C	0011	0001
C	0011	0001	D	0011	0101

Последовательность на выходе блока D совпадает с заданной последовательностью (0011 0101), что говорит о правильности работы синтезированной схемы. Проведём повторный анализ блока D, но на этот раз зафиксируем последовательность, поступающую с выхода блока A,

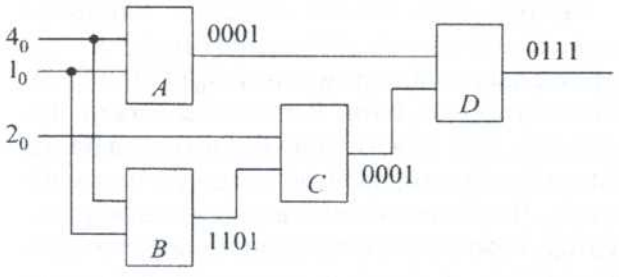


Рис. 1. Вариант схемы мультиплексора 2/1

поскольку она формируется непосредственно из входных сигналов, и построим последовательность C^*, которая не нарушит выходную последовательность схемы.

Поскольку блок D выполняет операцию ИЛИ, то наличие двух логических единиц в последовательности A позволяет соответствующие элементы последовательности C^* заменить на звёздочки. Выходная последовательность схемы при этом не изменится:

$$\begin{array}{ll} A & 0000\ 0101 \\ C^{*\vee} & 0011\ 0*0* \\ \hline D & 0011\ 0101 \end{array}$$

Обнаружить C^* среди последовательностей, формируемых на выходах элементов рассматриваемой схемы, не удалось, поэтому по методике Ю.П. Голубева оптимизация схемы невозможна. Однако проведём повторный анализ, теперь уже для блока C, с учётом введённых в схему изменений. Для этого зафиксируем поступающую на его вход последовательность 2_0 и попытаемся подобрать последовательность B^{**}, которая не нарушит построенной выше последовательности C^*. При выполнении указанной процедуры следует учесть, что блок C выполняет операцию И, и наличие логических нулей в последовательности 2_0 позволяет соответствующие элементы последовательности B^{**} заменить на звёздочки. Кроме того, звёздочкой можно заменить и восьмой элемент рассматриваемой последовательности, поскольку на этом месте в последовательности C^* также находится символ "*":

$$\begin{array}{ll} 2_0 & 0011\ 0011 \\ B^{**\&} & **11\ **0* \\ \hline C^* & 0011\ 0*0* \end{array}$$

Полученная таким образом числовая последовательность B^{**} не противоречит инверсии входной переменной 4_0 (1111 0000). Следовательно, блок B из схемы можно исключить, а на нижний (младший) вход блока C подать инвертированную входную переменную 4_0. Чтобы не вводить в схему лишний инвертор, можно соответствующим образом преобразовать числовую последовательность блока C – 0010. Такое преобразование соответствует инвертированию его младшей входной переменной [6]. Оптимизированная схема мультиплексора 2/1 представлена на рис. 2. Повторный анализ полученной схемы позволяет сделать вывод о том, что процедура двухуровневой оптимизации не нарушила логику работы синтезируемой схемы.

На рис. 3 приведён пример покрытия детализированной и оптимизированной схемы рис. 2 элементами 2И-НЕ. Указанная процедура заключается в формальном замещении каждого из блоков его реализацией на заданных логических элементах [6].

Дальнейшая оптимизация заключается в удалении из схемы цепочек последовательно соединённых инверторов. Окончательная оптимизированная схема мультиплексора 2/1, реализованная на элементах 2И-НЕ, представлена на рис. 4.

Рассмотренные выше алгоритмы для более сложных схем можно распространить на большее количество уровней оптимизации с целью увеличения недоопределённости рассматриваемых числовых последовательностей.

Многоуровневая оптимизация логических схем может выполняться в несколько проходов и проводиться на различных эта-

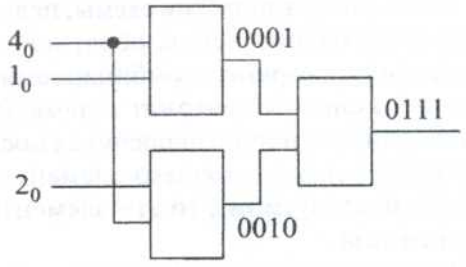


Рис. 2. Оптимизированная схема мультиплексора 2/1

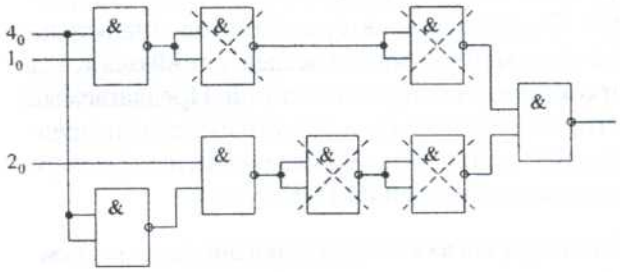


Рис. 3. Пример покрытия схемы рис. 2 элементами 2И-НЕ

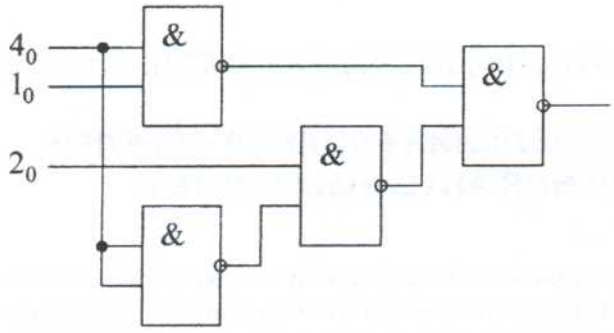


Рис. 4. Окончательно оптимизированная схема мультиплексора 2/1

пах синтеза (например, до и после этапа покрытия), что позволяет исключить из схемы значительное количество двухвходовых блоков и упростить дальнейшую реализацию процедур синтеза.

В [3] было показано, что любую комбинационную схему можно разделить на блоки, содержащие в общем случае не более трёх входов и относящиеся к одному из пяти типов. Можно предварительно провести оптимизацию указанных блоков и выполнять покрытие уже оптимизированными схемами. Такую процедуру можно назвать локальной оптимизацией, в отличие от процедуры глобальной оптимизации, рассмотренной выше.

Предлагаемый алгоритм многоуровневой оптимизации имеет несколько преимуществ по сравнению с существующими. Опишем вкратце особенности рассматриваемого алгоритма, а также отметим все его преимущества:

алгоритм не привязан к конкретному логическому базису. Это даёт возможность проводить процедуру оптимизации на различных этапах синтеза: на этапе детализации (предварительная оптимизация), а затем на этапе покрытия (окончательная оптимизация);

процедура оптимизации – многоуровневая, что значительно расширяет возможности поиска непротиворечивых последовательностей;

удаление элементов с одинаковыми входами и цепочек последовательно соединённых инверторов может рассматриваться как частный случай многоуровневой оптимизации, что позволяет говорить об универсальности рассматриваемого алгоритма;

процедура оптимизации на каждом этапе синтеза может проводиться неоднократно с учётом вносимых изменений до тех пор, пока в схеме находятся непротиворечивые последовательности;

при многоуровневой оптимизации необходимо избегать появления петель обратной связи внутри схемы, которые могут привести к нарушению логики функционирования схемы.

СПИСОК ЛИТЕРАТУРЫ

1. **Бибило П.Н., Леончик П.В.** Исследование эффективности логической минимизации в процессе синтеза комбинационных схем // Информатика. 2007. № 1. С. 22–29.

2. **Закревский А.Д., Торопов Н.Р.** Минимизация булевых функций многих переменных в классе ДНФ – итеративный метод и программная реализация // Прикладная дискретная математика. 2009. № 1. С. 5–14.

3. **Коновалов В.Н., Белов А.А., Коновалов И.В.** и др. Детализация комбинационных логических схем для покрытия простыми логическими элементами // Труды МГТУ 594. Методы исследования и проектирования сложных технических систем: Сб. ст. М.: Изд-во МГТУ им. Н.Э. Баумана, 2007. С. 103–113.

4. **Леончик П.В.** Минимизация систем булевых функций в классе дизъюнктивных нормальных форм // Информатика. 2006. № 1. С. 88–96.

5. **Коновалов В.Н., Коновалов И.В., Белов А.А.** и др. Логический синтез комбинационных схем на основе методов декомпозиции // Изв. ТулГУ. Сер. Вычислительная техника. Информационные технологии. Системы управления. Тула: Изд-во ТулГУ, 2006. Вып. 3. Системы управления. Т. 1. С. 231–233.

6. **Коновалов В.Н., Коновалов И.В., Белов А.А.** и др. Подход к покрытию комбинационных схем двухвходовыми логическими элементами // Труды МГТУ 594. Методы исследования и проектирования сложных технических систем: Сб. ст. М.: Изд-во МГТУ им. Н.Э. Баумана, 2007. С. 114–122.

7. **Поттосин Ю.В., Торопов Н.Р., Шестаков Е.А.** Метод минимизации системы полностью определённых булевых функций // Информатика. 2008. № 2. С. 102–110.

8. **Угрюмов Е.П.** Цифровая схемотехника: Учеб. пособие для вузов. 2-е изд., перераб. и доп. СПб.: БХВ-Петербург, 2004. 800 с.

УДК 621.396.67

Ю.В. Серединский, В.П. Горбунов, Ю.Н. Новиков

ПРИМЕНЕНИЕ ФЕРРИТОВ ДЛЯ ОСЛАБЛЕНИЯ ВЛИЯНИЯ КАБЕЛЕЙ НА ХАРАКТЕРИСТИКИ ИЗЛУЧЕНИЯ АНТЕННЫХ СИСТЕМ

Статья посвящена решению проблемы уменьшения паразитного влияния кабелей на характеристики антенных систем (АС), возникающего при построении многоярусных антенных комплексов (АК).

Ярусная конструкция антенной системы

При создании антенных систем многодиапазонных антенных комплексов (например, используемых в радиопеленгаторах) разработчики стремятся повышать их эффективность и широкополосность, укладываясь в компактные габаритные размеры. Часто для охвата широкого диапазона частот рабочую полосу разбивают на поддиапазоны, в каждом из которых работает отдельная АС. В этом случае возникает проблема электромагнитной совместимости – необходимо обеспечивать требуемые характеристики каждой АС при плотной компоновке всего АК. Одно из решений – разнесение антенн на ярусы. Пример ярусной конструкции АК представлен на рис. 1, *а*.

Объектом описанных в статье исследований является АС УКВ-диапазона (100–400 МГц), размещенная на одном из ярусов АК (рис. 1, *б*) [1]. Она состоит из нескольких независимых компланарных пар рамочных антенн, установленных на круглом металлическом экране. Испытания показали, что кабели от АС верхних ярусов, проходящие между рамками антенной системы УКВ-диапазона, искажают ее пеленгационные характеристики: наблюдается увеличение уровня побочного излучения (УПИ) и ширины диаграммы направленности (ДН) антенны. УПИ в данном случае – максимальный уровень заднего и боковых лепестков ДН.

Предложенные решения и методика их исследования

В целях ослабления влияния кабелей было предложено воспользоваться ферритовыми материалами, подобранными по типу и размещенными на кабелях особым образом. Рассматривались два решения: покрывать кабель ферритовыми кольцами [2] и наматывать кабель на ферритовые стержни [3].

Для выяснения эффективности этих способов выполнялись специальные исследования. Направление поиска – выбор ферритовых материалов наиболее подходящего типа и размера. Ожидалось, что при правильном использовании они могут осуществлять подавление паразитных сигналов и экранирование [4].

Чтобы оценить эффективность предложенных решений, была выполнена серия экспериментальных исследований на конструкциях антенн двух типов:

вспомогательной модели, имитирующей кабель АС верхнего яруса, что позволило подобрать лучшие в заданном диапазоне частот ферритовые материалы и геометрию их размещения на кабеле;

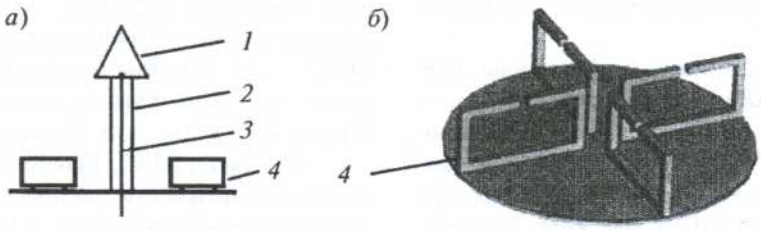


Рис. 1. Схема ярусной конструкции антенного комплекса (*а*) и ее фрагмент (*б*)

1 – антенна СВЧ; *2* – диэлектрическая стойка; *3* – сигнальный кабель; *4* – антенная система УКВ-диапазона на компланарных парах

компланарной паре рамочных антенн УКВ-диапазона в присутствии конструкций, имитирующих кабели разного вида (с ферритами и без них) между рамками, что позволило проверить эффективность предложенных решений.

Измерения производились на специализированном стенде антенных измерений по схеме, приведенной на рис. 2, *а*. Применялся метод вращающейся антенны, размещенной в дальней зоне (расстояние между излучателями 4 м, высота над землей 4 м). В качестве основного измерителя использовался анализатор спектра FSH6 фирмы ROHDE&SCHWARZ, работающий в режиме трекинг-генератора и позволяющий измерять коэффициент передачи гармонического сигнала (КП) с выхода на вход спектроанализатора в диапазоне 100–400 МГц с шагом 1 МГц.

Вспомогательная модель, имитирующая сигнальный кабель АС верхнего яруса, представляла собой несимметричный вибратор над экраном. Вибратор покрывался ферритовыми кольцами. Эффективность влияния ферритовых колец определенной марки характеризовалась величиной ослабления сигнала (в децибелах), принятого вибратором, вычисляемого по формуле

$$\text{Ослабление} = \text{КП}_{\text{отн}} = \text{КП} - \text{КП}_{\text{фер}},$$

где КП и $\text{КП}_{\text{фер}}$, дБ, – значения коэффициентов передачи в эксперименте с вибратором без ферритов и при добавлении ферритовых колец некоторой марки; $\text{КП}_{\text{отн}}$, дБ, – относительный коэффициент передачи.

В ходе исследований было сформулировано предположение: полученные для ферритовых колец выводы о влиянии параметров материала на излучение кабеля справедливы и для материалов ферритовых стержней. Это предположение впоследствии подтвердилось в экспериментах.

Частотные зависимости ослабления принимаемого вибратором сигнала для разных ферритовых материалов

Исследования производились для ряда отечественных марок ферритов (см. таблицу).

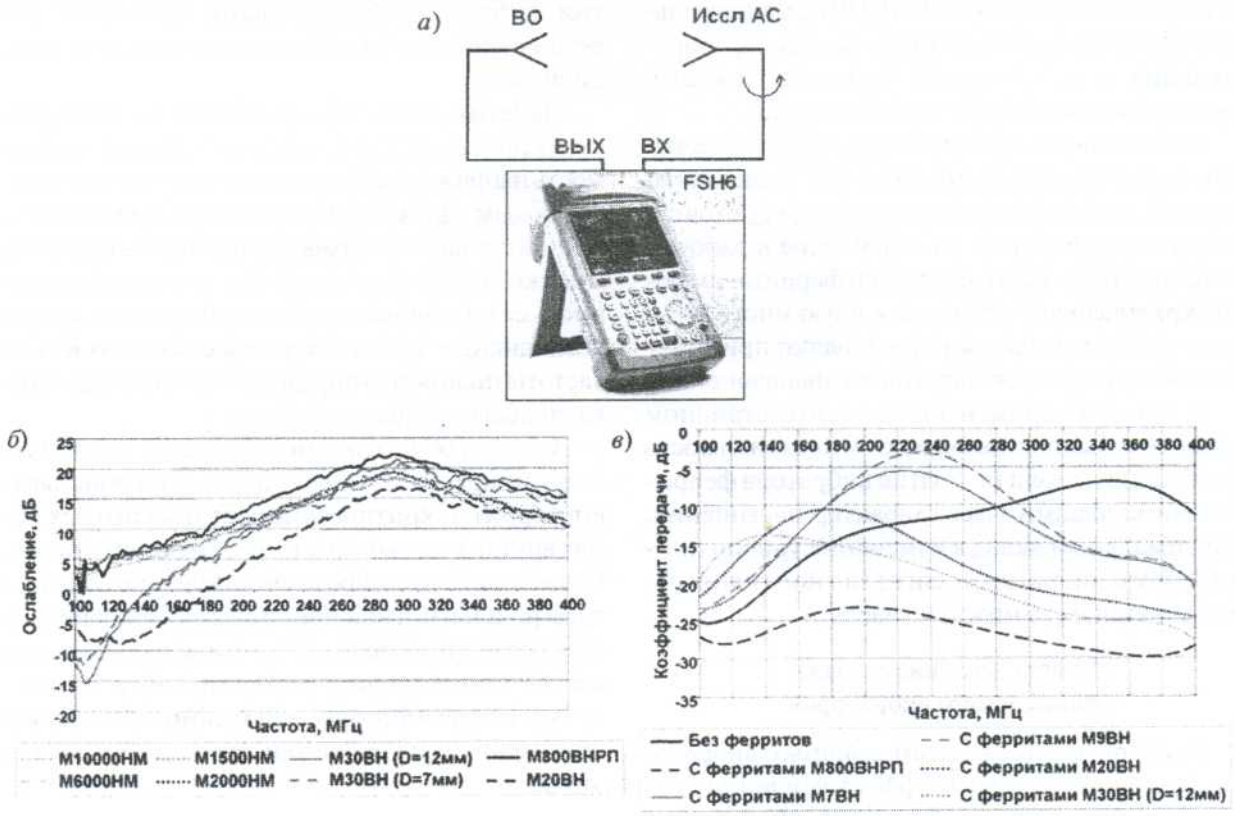


Рис. 2. Измерительная схема (*а*); результаты измерений частотных зависимостей ослабления сигнала (*б*) и коэффициента передачи измерительной цепи (*в*). Длина вибратора 240 мм (*б*) и 200 мм (*в*)

ВО – вспомогательный облучатель; *АС* – исследуемая антенная система; FSH6 – анализатор спектра; *ВХ, ВЫХ* – его вход и выход

Основные характеристики использованных ферритовых колец

Марка феррита	$\mu_н$	f_c, МГц	ρ, Ом·м	$D \times d \times h$, мм
М10000НМ	10000	0,05	0,1	7 × 4 × 2
М6000НМ	6000	0,1	1	
М2000НМ	2000	0,45	5	
М1500НМ	1500	1,7	0,5	
М800ВНРП	800	0,5	10^5	
М30ВН (D = 7 мм)	30	200	10^5	
М30ВН (D = 12 мм)	30	200	10^5	12 × 6 × 4,5
М20ВН	20	200	10^6	10 × 6 × 3
М9ВН	9	650	10^6	7 × 4 × 2
М7ВН	7	220	10^6	7 × 4 × 2

Принятые в таблице обозначения: $\mu_н$ – начальная магнитная проницаемость; f_c – критическая частота; ρ – удельное электрическое сопротивление; D, d, h – внешний, внутренний радиус и высота кольца соответственно.

На рис. 2, *б* приведены измеренные частотные зависимости ослабления сигнала, принимаемого вибратором длиной 240 мм, для восьми марок ферритов. Наибольшее ослабление сигнала дают кольца марки М800ВНРП с максимальным значением удельного электрического сопротивления и критической частотой, лежащей много ниже частот рабочего диапазона.

Увеличение толщины колец приводит к незначительному увеличению ослабления. Значение начальной магнитной проницаемости ферритов существенного влияния на ослабление в рабочем диапазоне не оказало. Все марки феррита, имеющие критическую частоту, лежащую много ниже частот рабочего диапазона, ослабляют принимаемый вибратором сигнал во всем диапазоне.

В ходе экспериментов на несимметричном вибраторе было установлено, что зависимость ослабления от доли покрытия вибратора ферритовыми кольцами имеет характер насыщения. Максимальный вклад в изменение уровня принимаемого вибратором сигнала вносят кольца, заполняющие его нижнюю часть.

Эффект усиления сигнала, принимаемого вибратором

Эксперименты с несимметричным вибратором позволили выявить эффект усиления сигнала, принимаемого вибратором. На рис. 2, *б* этому усилению соответствуют отрицательные значения ослабления в нижней части рабочего диапазона для ферритов М30ВН и М20ВН.

Приведем результаты экспериментальных исследований по полю, полученные для вибратора длиной 200 мм и позволяющие подробно рассмотреть эффект усиления сигнала, принимаемого вибратором с ферритами [5]. Зависимость абсолютного значения КП от частоты для вибратора без ферритов имеет резонансный характер с максимумом вблизи резонансной частоты вибратора 375 МГц (рис. 2, *в*). При покрытии вибратора ферритовыми кольцами его резонансная частота сдвигается в нижнюю часть диапазона.

Заметим, что в экспериментах по полю (см. схему рис. 2, *а*) не учитывалась частотная зависимость напряженности поля, излучаемого вспомогательным облучателем – логопериодической антенной с диапазоном рабочих частот 80–1000 МГц. Однако можно предполагать, что неравномерность ее излучения в рабочем диапазоне мала и резонансные зависимости абсолютного КП от частоты полностью определяются характеристиками исследуемого вибратора.

Следует обратить внимание на то, что усиление принимаемого вибратором сигнала вызывают ферриты, критические частоты которых лежат внутри или выше рабочего диапазона частот. Если исходить из определения критической частоты [6] и магнитной проницаемости в переменных магнитных полях [7], то участок усиления соответствует малому рассеянию энергии магнитного поля в феррите или малому ослаблению протекающих по вибратору токов, вызывающих это поле.

Исследование входного сопротивления

Для более глубокого исследования влияния ферритовых материалов на характеристики ан-

тенны регистрировались зависимости от частоты действительной и мнимой частей входного сопротивления несимметричного вибратора с ферритовыми кольцами. Эксперименты выполнялись по схеме, представленной на рис. 3, *а*, с использованием прибора Р4-37 (измеритель комплексных коэффициентов передачи).

Рис. 3, *б* и *в* иллюстрирует полученные результаты; вибратор был длиной 200 мм без ферритов и с ферритами М800ВНРП (критическая частота много ниже частот рабочего диапазона), М9ВН (критическая частота выше рабочего диапазона).

Из графиков видно, что покрытие вибратора ферритовыми кольцами М800ВНРП приводит к ухудшению согласования вибратора с подключенным к нему фидерным трактом (волновое сопротивление тракта 50 Ом) во всем рабочем диапазоне. Это важная причина ослабления принимаемого вибратором сигнала.

Покрытие вибратора ферритовыми кольцами М9ВН приводит к сдвигу зависимостей вещественной и мнимой частей входного сопротивления вибратора от частоты в нижнюю часть диапазона. При этом изменяется коэффициент согласования вибратора с подключенным к нему трактом таким образом, что лучшее согласование с выходной цепью происходит на более низких частотах. Этот эффект эквивалентен увеличению длины вибратора и может быть использован для миниатюризации антенных систем [5].

Таким образом, для ослабления излучения кабелей пригодны только ферритовые материалы со значениями критических частот, лежащими много ниже частот рабочего диапазона.

Влияние кабелей на диаграммы направленности компланарной пары

Исследовалось ослабление влияния кабеля на ДН компланарной пары рамочных антенн с помощью ферритов в диапазоне частот 100–400 МГц. Для этого требовалось измерить ДН компланарной пары по схеме рис. 2, *а* для пяти перечисленных ниже случаев устройства такой пары [3]:

две рамочные антенны над экраном ("чистая" компланарная пара);

пара дополнена металлическим штырем длиной 200 мм, имитирующим кабель, проходящий между рамками;

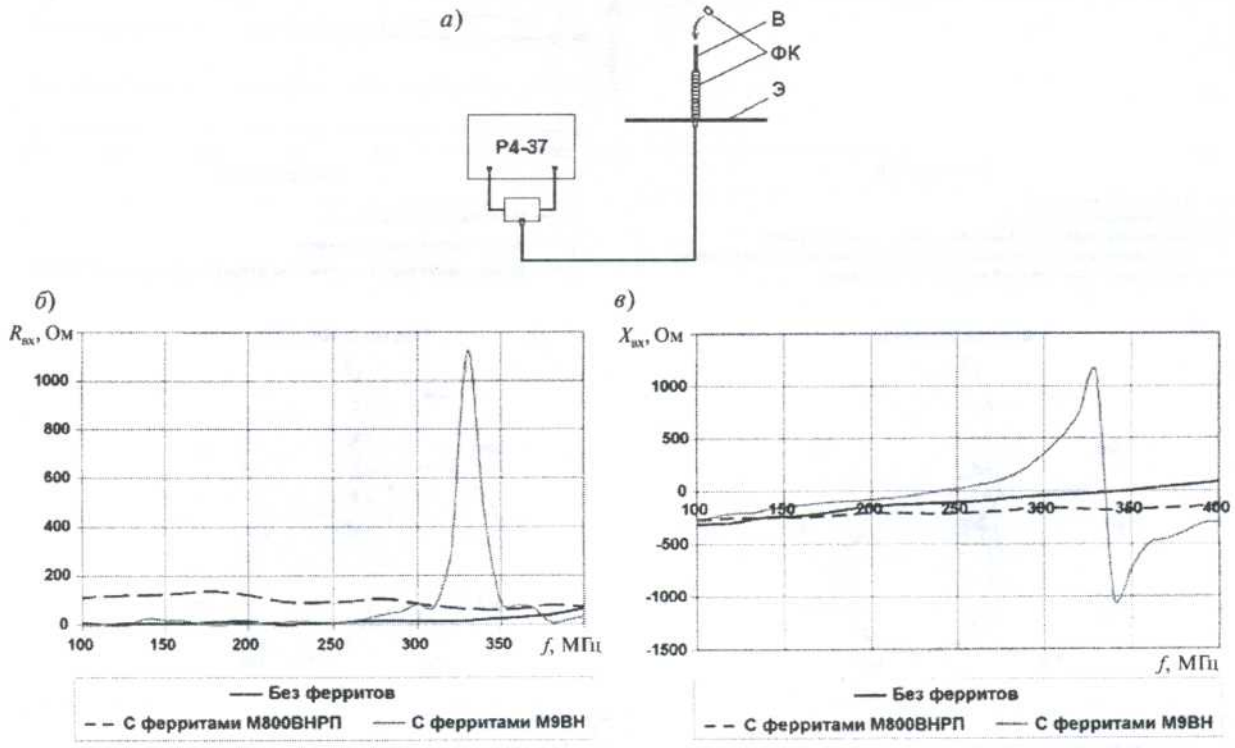


Рис. 3. Измерительная схема (*а*) и результаты измерений частотных зависимостей действительной (*б*) и мнимой (*в*) частей входного сопротивления вибратора

В – вибратор; *ФК* – ферритовые кольца; *Э* – экран; Р4-37 – измеритель комплексных коэффициентов передачи

между рамками помещен кабель, плотно намотанный на ферритовый стержень марки М150ВН (критическая частота 25 МГц). Длина помещаемой между рамками конструкции 200 мм;

между рамками помещен свитый спиралью кабель без ферритового сердечника. Длина конструкции 200 мм;

пара дополнена металлическим штырем длиной 200 мм, полностью покрытым ферритовыми кольцами М10000НМ.

Во всех случаях измерялись ДН компланарной пары. Штырь и оплетка кабеля, помещаемые между рамками, заземлялись (соединялись с экраном).

Эффективность того или иного способа ослабления влияния удобно анализировать по графикам зависимости УПИ и ширины ДН компланарной пары от частоты (приводимые ниже графики интерполированы полиномами).

Полученные результаты позволяют сделать вывод, что помещенный между рамками штырь (модель кабеля) нарушает работу компланарной пары (рис. 4, *а* и *б*): кривые УПИ и ширины ее ДН при наличии штыря лежат выше соответствующих кривых для чистой компланарной пары. Максимальное отклонение наблюдается в верхней части диапазона вблизи значения 375 МГц – резонансной частоты несимметричного вибратора длиной 200 мм. Помещая вместо штыря намотку кабеля на ферритовый стержень, получаем кривые, близкие к графикам для чистой компланарной пары. Это говорит об эффективном подавлении паразитного излучения кабеля, т. е. положительном влиянии на пеленгационные характеристики компланарной пары.

Измерение ДН для конструкции, включающей свитый спиралью кабель (без ферритового стержня), помещенный между рамками, выяви-

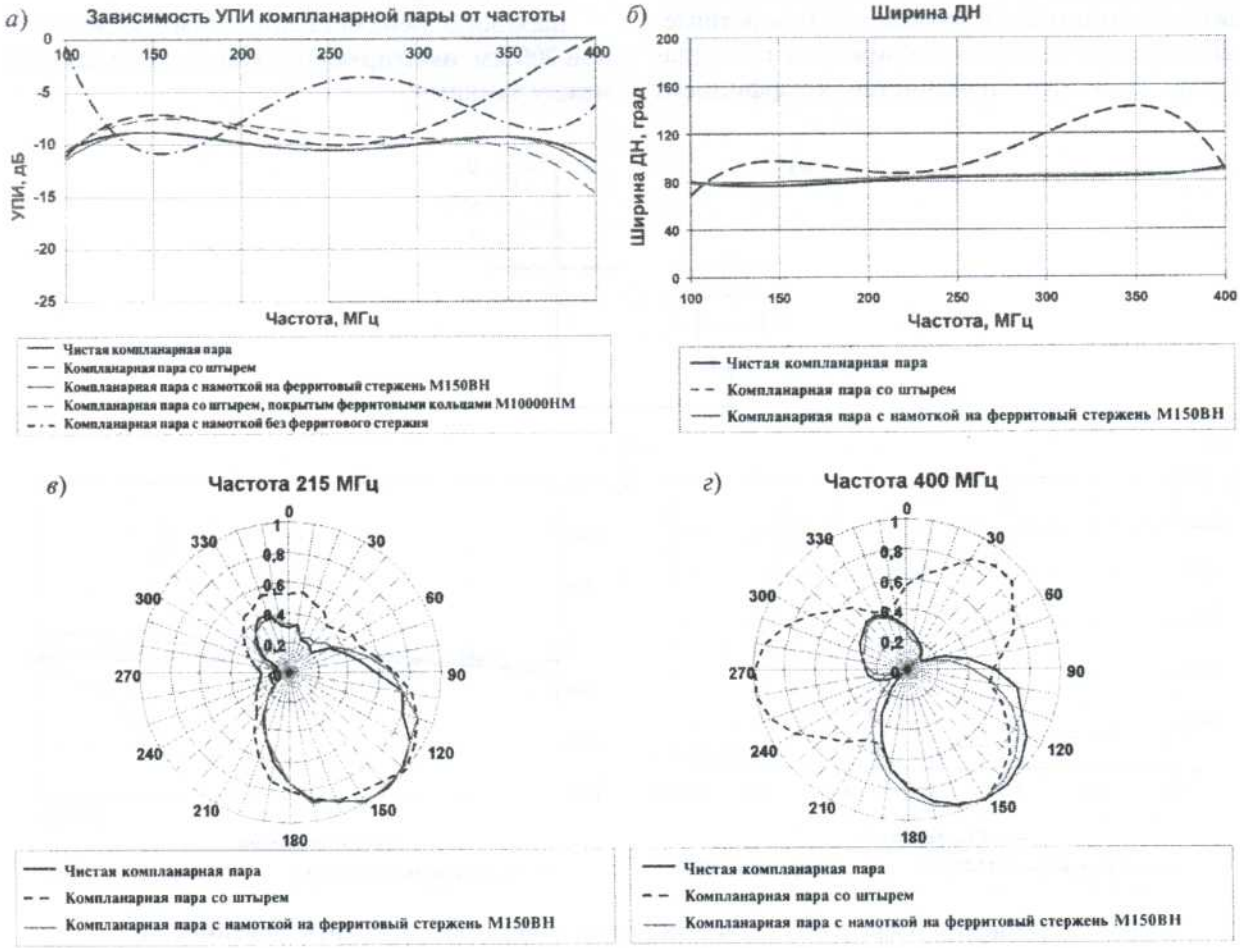


Рис. 4. Влияние различных конструкций кабелей на диаграммы направленности (ДН)

ло, что для эффективного ослабления паразитного излучения обязательно наличие ферритового сердечника в таком кабеле. Отклонение полученной частотной зависимости УПИ от соответствующей кривой для чистой компланарной пары иллюстрирует рис. 4, *а*.

Было проведено экспериментальное сравнение двух способов ослабления паразитного излучения: с металлическим штырем, покрытым ферритовыми кольцами, и с кабелем, намотанным на ферритовый стержень (рис. 4, *а*). Установлено, что оба предложенных способа эффективны, но ферритовый стержень М150ВН лучше восстанавливает пеленгационные характеристики КП во всем рабочем диапазоне частот, чем ферритовые кольца М10000НМ.

Эффективность предложенных конструктивных решений иллюстрируют ДН на двух частотах (рис. 4, *в*, *г*) для разных модификаций компланарной пары: "чистой", с металлическим штырем длиной 200 мм, и с намоткой кабеля на ферритовый стержень М150ВЧ (длина конструкции 200 мм).

Таким образом, применение ферритовых материалов разного вида (колец и стержней) со специально подобранными параметрами позволяет избежать искажения характеристик излучения антенных систем УКВ-диапазона многоярусных антенных комплексов.

СПИСОК ЛИТЕРАТУРЫ

1. **Горбунов В.П., Новиков Ю.Н.** Антенная система радиопеленгатора, работающего в сложной электромагнитной обстановке // XXXIII Неделя науки СПбГПУ. Ч. VI: Матер. Всерос. межвуз. науч.-техн. конф. студ. и асп. СПб.: Изд-во СПбГПУ, 2005. С. 10–12.

2. **Серединский Ю.В., Горбунов В.П., Новиков Ю.Н.** Применение ферритов для ослабления влияния кабелей в многоуровневых антенных комплексах // XXXVI Неделя науки СПбГПУ. Ч. VI: Матер. Всерос. межвуз. науч.-техн. конф. студ. и асп. СПб.: Изд-во СПбГПУ, 2008. С. 3–5.

3. **Серединский Ю.В., Горбунов В.П., Новиков Ю.Н.** Применение ферритовых стержней для ослабления влияния кабелей в многоуровневых антенных комплексах // XXXVII Неделя науки СПбГПУ. Ч. VI: Матер. Всерос. межвуз. науч.-техн. конф. студ. и асп. СПб.: Изд-во СПбГПУ, 2008. С. 3–5.

4. **Отт Г.** Методы подавления шумов и помех в электронных системах. М.: Мир, 1979. 318 с.

5. **Горбунов В.П., Серединский Ю.В., Новиков Ю.Н.** Увеличение эффективности несимметричного вибратора с помощью ферритовых колец // Труды СПбГПУ. 2008. № 507. С. 42–47.

6. Электронный каталог магнитомягких ферритов ([Электронный ресурс]: каталог содержит сведения о ферритовых материалах, производимых ОАО "НИИ ФЕРРИТ-ДОМЕН": http://www.ferrite-domen.com).

7. Физический энциклопедический словарь / Гл. ред. А.М. Прохоров. М.: Сов. энциклопедия, 1984. 944 с.

УДК 621.37

В.А. Сороцкий, М.А. Уткин

ВЛИЯНИЕ ЭЛЕКТРОМАГНИТНОЙ СВЯЗИ АНТЕНН РАДИОПЕРЕДАЮЩИХ УСТРОЙСТВ НА ПРОЦЕССЫ В КЛЮЧЕВЫХ ГЕНЕРАТОРАХ

Переход на современную элементную базу, а также использование более совершенных схемотехнических решений при реализации ключевых генераторов (КГ), используемых в качестве усилителей мощности радиопередающих устройств (РПдУ), открывают возможности для уменьшения массогабаритных характеристик оборудования существующих радиоцентров и создают предпосылки их модернизации, в том числе путем размещения в существующей инфраструктуре этих центров дополнительных радиопередатчиков. С учетом изложенного представляется целесообразным исследование особенностей совместного функционирования

двух или более РПдУ, обусловленных их взаимным влиянием друг на друга. К числу наиболее типичных ситуаций в этой связи следует отнести одновременную работу двух или более КГ на отличающихся частотах при наличии электромагнитной связи антенн их РПдУ. Для определенности в дальнейшем ограничимся рассмотрением только двух РПдУ, что нисколько не снижает практической ценности полученных результатов.

Исследование данной проблемы целесообразно ограничить рассмотрением квазистационарного режима работы КГ, поскольку в этом случае можно получить решение рассматриваемой задачи в общем виде. Что же касается анализа процессов в переходных режимах работы КГ, под которыми в данном случае подразумеваются процессы установления амплитуды и фазы колебаний в одном генераторе при манипуляции частотой и электронной перестройкой элементов согласующего устройства (если она используется) в другом генераторе, а также процессы пуска или выключения одного из генераторов при работающем втором генераторе, то решение подобной задачи с учетом высокого порядка системы дифференциальных уравнений, описывающих процессы в антенно-согласующих трактах (АСТ) обоих РПдУ (как правило, не ниже 10-го порядка), а также ненулевых начальных условий подразумевает применение численных методов и позволяет получить решение применительно к конкретной структуре и определенному набору параметров элементов АСТ. Данное обстоятельство переводит указанную задачу из теоретических в разряд задач практической направленности и заметно снижает актуальность ее рассмотрения в рамках настоящей статьи. Дополнительным аргументом в пользу такого решения может служить также и то обстоятельство, что широко известны приближенные методы оценки качества переходных процессов в линейных цепях, рассмотренные, например, в [1, 2], которые также могут быть использованы при анализе процессов установления амплитуды и фазы колебаний в КГ. Кроме того, в частном случае, когда переходные процессы в АСТ носят монотонный характер, в качестве оценки максимальных значений амплитуды и фазы колебаний, разумеется, могут быть использованы значения, полученные из решения квазистационарной задачи.

При анализе квазистационарных процессов в выходных цепях генераторов примем допущение о том, что напряжения и токи, формируемые каждым из генераторов, изменяются по гармоническому закону. Обоснованность такого предположения обусловлена тем, что коэффициент гармоник $K_Г$ выходного напряжения КГ может быть сравнительно невелик (например, при формировании на выходе КГ трехступенчатого напряжения $K_Г = 12{,}5\ \%$ [3]).

Расчет токов в выходных цепях генераторов можно осуществить на основании эквивалентной схемы, приведенной на рис. 1. На рисунке ключевые генераторы представлены в виде источников гармонического напряжения $e_1(t)$ и $e_2(t)$, формирующих колебания соответственно на частоте ω_1 и ω_2, а четырехполюсник T включает в себя все колебательные цепи рассматриваемой системы: эквивалентные схемы замещения антенн РПдУ1 и РПдУ2, согласующие и настроечные элементы каждой из этих антенн, а также эквивалентный четырехполюсник, учитывающий их взаимную электромагнитную связь.

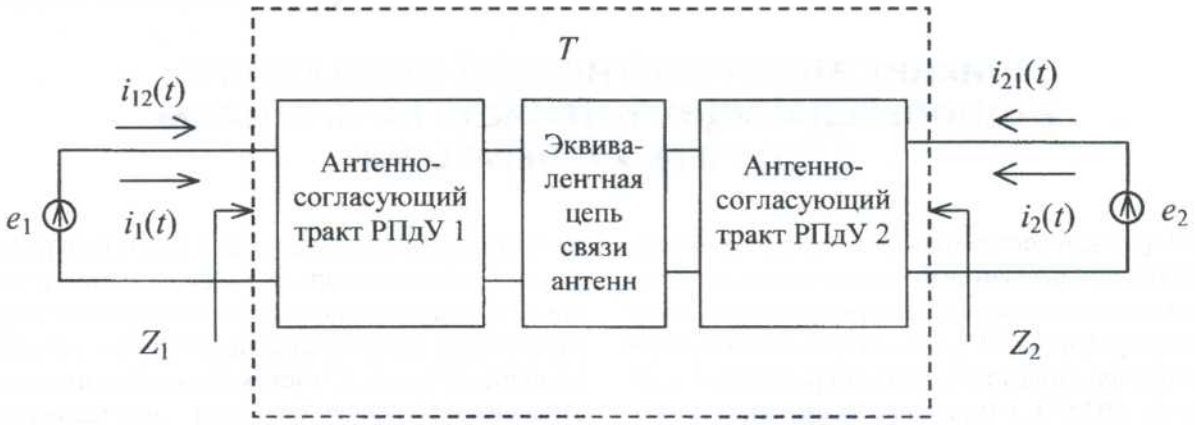


Рис. 1. Эквивалентная схема для расчета токов в выходных цепях генераторов

Токи $i_1(t)$ и $i_2(t)$ обозначают соответственно токи на выходе каждого из генераторов, протекающие в отсутствие электромагнитной связи между антеннами. Ток $i_{12}(t)$, протекающий в цепи первого КГ, обусловлен влиянием РПдУ2 на РПдУ1. По аналогии с этим ток $i_{21}(t)$ учитывает влияние РПдУ1 на РПдУ2.

Амплитуды первой гармоники напряжений, формируемых генераторами e_1 и e_2, можно записать в виде

$$U_1 = A_1 E_1 = \frac{4E_1}{\pi}\sum_{j=1}^{N_{S1}}\cos(\gamma_{1j}); \qquad (1)$$

$$U_2 = A_2 E_2 = \frac{4E_2}{\pi}\sum_{g=1}^{N_{S2}}\cos(\gamma_{2g}), \qquad (2)$$

где $A_{1,2}$ – первая гармоника коммутационной функции $a_{1,2}(t)$, определяющей форму выходного напряжения соответственно первого и второго ключевых генераторов; $E_{1,2}$ – напряжения питания генераторов; N_{S1}, N_{S2} и γ_{1j}, γ_{2g} – соответственно число ступеней в выходном напряжении и угловые координаты ступеней для каждого из генераторов.

В рассматриваемом приближении при совместной работе двух РПдУ на взаимосвязанные линейные цепи ток в выходной цепи каждого генератора содержит две гармонические составляющие с частотами $\omega_1 = 2\pi f_1$ и $\omega_2 = 2\pi f_2$, результатом суперпозиции которых будут биения этих колебаний:

$$\begin{aligned} i_{g1}(t) &= i_1(t) + i_{12}(t) = \\ &= I_1\sin(\omega_1 t + q_1) + I_{12}\sin(\omega_2 t + q_{12}); \\ i_{g2}(t) &= i_2(t) + i_{21}(t) = \\ &= I_2\sin(\omega_2 t + q_2) + I_{21}\sin(\omega_1 t + q_{21}). \end{aligned} \qquad (3)$$

Амплитуды и фазы токов, входящих в выражения (3), могут быть определены на основе принципа наложения и выражаются через параметры короткого замыкания четырехполюсника T (см. рис. 1) на соответствующих частотах:

$$\begin{aligned} I_1 &= U_1\Big/\left|\dot{Z}_{1KZ}(\omega_1)\right|,\ q_1 = -\operatorname{arctg}\frac{\operatorname{Im}\dot{Z}_{1KZ}(\omega_1)}{\operatorname{Re}\dot{Z}_{1KZ}(\omega_1)}; \\ I_2 &= U_2\Big/\left|\dot{Z}_{2KZ}(\omega_2)\right|,\ q_2 = -\operatorname{arctg}\frac{\operatorname{Im}\dot{Z}_{2KZ}(\omega_2)}{\operatorname{Re}\dot{Z}_{2KZ}(\omega_2)}; \\ I_{12} &= U_2\Big/\left|\dot{Z}_{TKZ}(\omega_2)\right|,\ q_{12} = -\operatorname{arctg}\frac{\operatorname{Im}\dot{Z}_{TKZ}(\omega_2)}{\operatorname{Re}\dot{Z}_{TKZ}(\omega_2)}; \\ I_{21} &= U_1\Big/\left|\dot{Z}_{TKZ}(\omega_1)\right|,\ q_{21} = -\operatorname{arctg}\frac{\operatorname{Im}\dot{Z}_{TKZ}(\omega_1)}{\operatorname{Re}\dot{Z}_{TKZ}(\omega_1)}, \end{aligned} \qquad (4)$$

где $\dot{Z}_{1KZ}(\omega_1)$ и $\dot{Z}_{2KZ}(\omega_2)$ – входные сопротивления четырехполюсника T со стороны каждого источника на его рабочей частоте при закороченном другом источнике; $\dot{Z}_{TKZ}(\omega_1)$ и $\dot{Z}_{TKZ}(\omega_2)$ – сопротивления передачи от одного источника на его рабочей частоте в цепь другого источника.

После нормировки суммарного тока каждого генератора к амплитуде составляющей тока собственной частоты выражения (3) можно переписать:

$$\begin{aligned} \tilde{i}_{g1}(t) &= i_{g1}(t)/I_1 = \sin(\omega_1 t + q_1) + K_1\sin(\omega_2 t + q_{12}); \\ \tilde{i}_{g2}(t) &= i_{g2}(t)/I_2 = \sin(\omega_2 t + q_2) + K_2\sin(\omega_1 t + q_{21}), \end{aligned} \qquad (5)$$

где

$$K_1 = \frac{U_2\left|\dot{Z}_{1KZ}(\omega_1)\right|}{U_1\left|\dot{Z}_{TKZ}(\omega_2)\right|},\quad K_2 = \frac{U_1\left|\dot{Z}_{2KZ}(\omega_2)\right|}{U_2\left|\dot{Z}_{TKZ}(\omega_1)\right|}.$$

Окончательные выражения для нормированных токов генераторов могут быть представлены в форме:

$$\begin{aligned} \tilde{i}_{g1}(t) &= B_1(t)\sin[\omega_1 t + q_1 + Q_1(t)]; \\ \tilde{i}_{g2}(t) &= B_2(t)\sin[\omega_2 t + q_2 + Q_2(t)], \end{aligned} \qquad (6)$$

где

$$B_1(t) = \left\{1 + K_1^2 + 2K_1\cos[(\omega_2 - \omega_1)t + q_{21} - q_1]\right\}^{\frac{1}{2}}; \qquad (7)$$

$$B_2(t) = \left\{1 + K_2^2 + 2K_2\cos[(\omega_1 - \omega_2)t + q_{12} - q_2]\right\}^{\frac{1}{2}}; \qquad (8)$$

$$Q_1(t) = \operatorname{arctg}\frac{K_1\sin[(\omega_2 - \omega_1)t + q_{21} - q_1]}{1 + K_1\cos[(\omega_2 - \omega_1)t + q_{21} - q_1]}; \qquad (9)$$

$$Q_2(t) = \operatorname{arctg}\frac{K_2\sin[(\omega_1 - \omega_2)t + q_{12} - q_2]}{1 + K_2\cos[(\omega_1 - \omega_2)t + q_{12} - q_2]}. \qquad (10)$$

Из этих выражений видно, что амплитуды и фазы токов генераторов периодически изменяются во времени вследствие биений колебаний. Дополнительные фазовые сдвиги $Q_1(t)$ и $Q_2(t)$, появляющиеся в выходных токах генераторов вследствие взаимного влияния рассматриваемых РПдУ, могут привести к серьезным негативным последствиям. Так, в случае использования генераторов, выполненных на основе неполностью управляемых ключевых приборов (тиристоров) по схеме резонансного инвертора с естественной коммутацией ключей, появление дополнительного фазового сдвига в выходном токе КГ может нарушить устойчивость его работы и привести к

опрокидыванию. С точки зрения устойчивости КГ указанного типа необходимо, чтобы максимальное значение дополнительного фазового сдвига, вызванного биениями, не превосходило заданной допустимой величины, которая зависит, главным образом, от времени выключения тиристорных ключей. В случае применения в составе РПдУ транзисторных КГ появление заметного по величине дополнительного фазового сдвига в выходном токе генератора также способно оказать негативное влияние на его характеристики, поскольку указанный фазовый сдвиг эквивалентен рассогласованию КГ с нагрузкой и может привести к возрастанию мощности, рассеиваемой в КГ.

Оценим максимальную величину отклонения фазы тока генераторов, возникающего в результате биений, в зависимости от разноса частот генераторов. Для этого вычислим производные $\partial Q_1/\partial\phi_1$ и $\partial Q_2/\partial\phi_2$, где $\phi_1 = (\omega_2 - \omega_1)t$, $\phi_2 = (\omega_1 - \omega_2)t$, и приравняем их нулю. Учитывая сходство выражений (9) и (10), ограничимся вычислением только одной из производных, поскольку для второй производной результаты будут очевидны. Получим:

$$\frac{\partial Q_1}{\partial\phi_1} = \frac{K_1\cos(\vartheta_1)[1 + K_1\cos(\vartheta_1)] + K_1^2\sin^2(\vartheta_1)}{[1 + K_1\cos(\vartheta_1)]^2 + [K_1\sin(\vartheta_1)]^2}, \quad (11)$$

где $\vartheta_1 = \phi_1 + q_{21} - q_1$.

Приравняв числитель в (11) нулю и проведя несложные вычисления, определим значение параметра ϕ_1, при котором $Q_1(t)$ принимает максимальное значение:

$$\phi_{1o} = \arccos(-K_1). \quad (12)$$

Параметр ϕ_2, при котором $Q_2(t)$ принимает максимальное значение, запишем по аналогии с (12):

$$\phi_{2o} = \arccos(-K_2). \quad (13)$$

Подстановка (12) и (13) в выражения (9) и (10) позволяет определить экстремумы $Q_1(t)$ и $Q_2(t)$, значения которых при $\omega_1 \neq \omega_2$ определяются выражениями

$$Q_{1,2\,\max} = \arcsin(K_{1,2}). \quad (14)$$

При этом экстремумы огибающих тока в выходной цепи генераторов равны:

$$B_{1,2} = 1 + K_{1,2}. \quad (15)$$

Таким образом, наличие электромагнитной связи между антеннами РПдУ и обусловленные ею биения колебаний в выходных цепях ключевых генераторов этих РПдУ вызывают как увеличение максимальных значений токов, так и появление дополнительных фазовых сдвигов этих токов. При этом максимальные значения обоих параметров определяются коэффициентами K_1 и K_2. Следует также обратить внимание на то, что полный размах пульсации фазы равен удвоенному значению $Q_{1,2\,\max}$.

Оценим важность учета влияния рассмотренного выше эффекта на примере двух КГ, работающих на индивидуальные антенны, аналогичные антенне, рассмотренной в [4]. Семейства зависимостей $Q_{1\max}$ (сплошные линии) и $Q_{2\max}$ (прерывистые линии), рассчитанные для различных значений нормированных частот генераторов $f_1 = f_1/F_{\text{норм}}$ и $f_2 = f_2/F_{\text{норм}}$, а также различных значений отношения $g_a = U_2/U_1$, приведены на рис. 2–4. В качестве частоты нормировки использована верхняя граница диапазона СНЧ.

Как видно из приведенных на рис. 2 зависимостей, дополнительный фазовый сдвиг тока в выходной цепи каждого из генераторов при равных амплитудах генерируемых ими колебаний и близких значениях рабочих частот может доходить до 8–10°. Для сравнения отметим, что в случае использования тиристорных ключевых генераторов, работающих в режиме с естественной коммутацией полупроводниковых вентилей и выполненных, например, на быстродействующих тиристорах типа ТБ 253-800 или ТБ 253-1000 с максимальным значением времени выключения 63 мкс [5], необходимый для их надежного выключения фазовый сдвиг тока при емкостной расстройке на верхней границе диапазона СНЧ составляет всего 6,8°. Для приборов указанного типа с меньшим временем выключения, например 32 или 25 мкс, указанный фазовый сдвиг тока соответственно равен 3,5 и 2,7°. Таким образом, если не учитывать дополнительный фазовый сдвиг тока в выходных цепях генераторов, обусловленный биениями колебаний, и не ввести необходимый запас по величине емкостной расстройки их нагрузки, электромагнитная связь антенн РПдУ, вероятнее всего, приведет к так называемому опрокидыванию КГ. Что же касается КГ, выполненных на основе транзисторов, то, хотя проблема опрокидывания для них и не актуальна, тем не менее тот факт, что полный размах пульсации фазы тока нагрузки при равных амплитудах колебаний генераторов

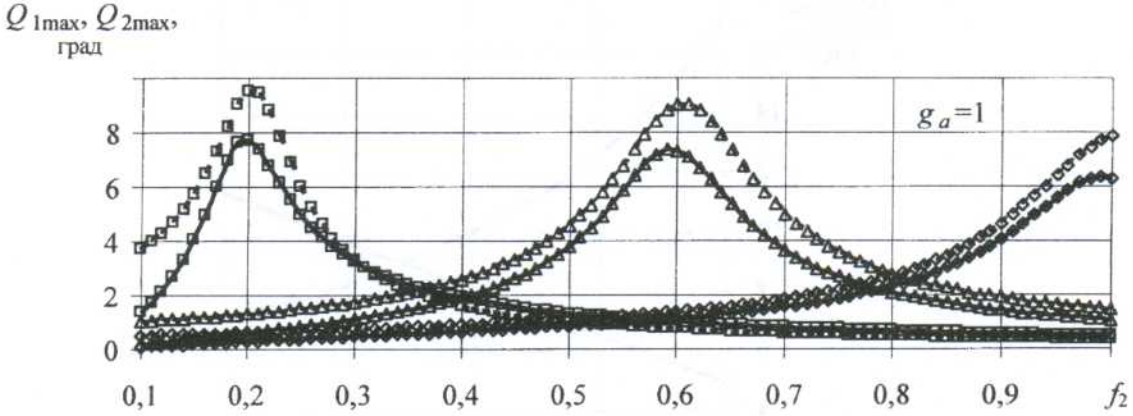


Рис. 2. Зависимости Q_{1max} и Q_{2max} от частоты при $g_a = 1$

f_1 = 0,2( ); 0,6 ( ); 1 ( ); 0,2 ( ); 0,6 ( ); 1 ( )

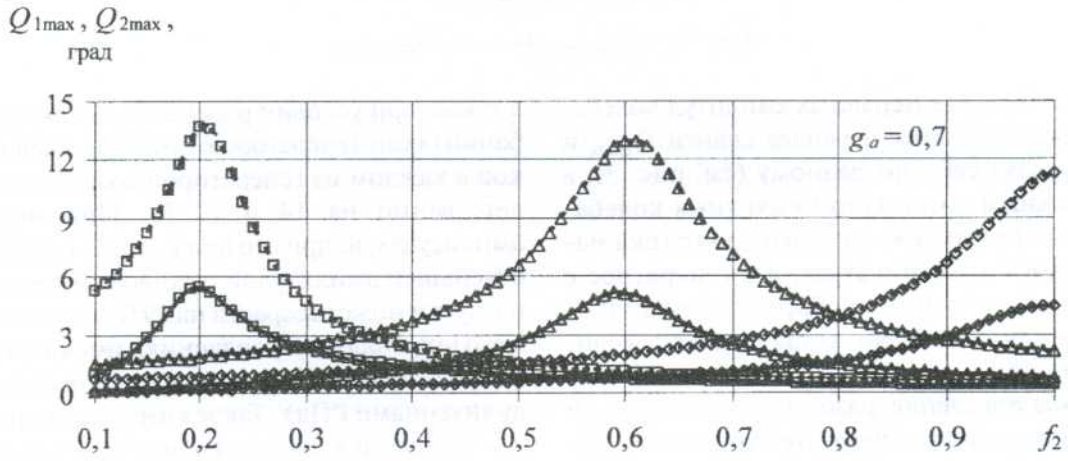


Рис. 3. Зависимости Q_{1max} и Q_{2max} от частоты при $g_a = 0,7$

f_1 = 0,2( ); 0,6 ( ); 1 ( ); 0,2 ( ); 0,6 ( ); 1 ( )

(см. рис. 2) может достигать $2Q_{1,\,2\,max} = 16–20°$, позволяет прогнозировать ухудшение энергетических характеристик генераторов. Это может быть обусловлено, в частности, попаданием рабочей точки КГ в режим с емкостной расстройкой, в котором очень высока вероятность протекания сквозных токов по цепи диод встречного тока – транзистор каждой из стоек КГ.

Как следует из зависимостей, приведенных на рис. 2, экстремумы Q_{1max} и Q_{2max} имеют место при равенстве частот обоих КГ. Это является следствием того, что на этих частотах сопротивления передачи от одного источника в цепь другого источника ($\dot{Z}_{T\,KZ}(\omega_1)$ и $\dot{Z}_{T\,KZ}(\omega_2)$ в (4)) принимают минимальные значения вследствие настройки АСТ обоих РПдУ на одну и ту же частоту. В то же время по мере увеличения разноса частот дополнительный фазовый сдвиг уменьшается до сравнительно небольших значений, составляющих величину порядка нескольких десятых долей градуса. Таким образом, используя эти зависимости, можно определить необходимый разнос частот РПдУ по заданному допустимому значению пульсации фазы токов КГ, вызванной биениями.

До сих пор амплитуды колебаний обоих генераторов предполагались одинаковыми. Однако на практике это условие выполняется далеко

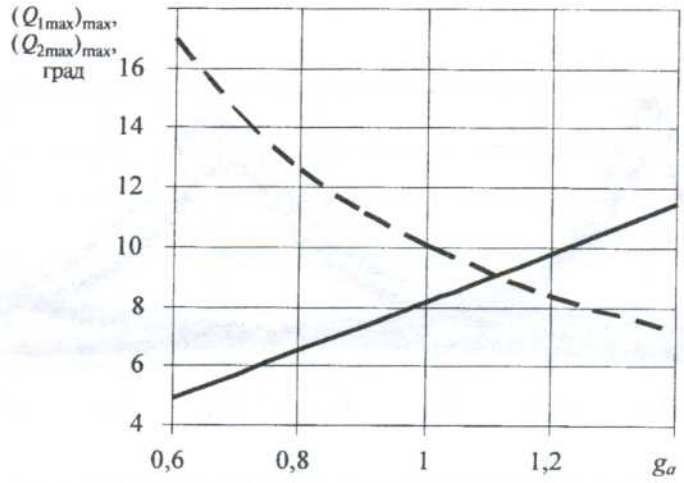


Рис. 4. Зависимости максимальных фазовых сдвигов $(Q_{1max})_{max}$ и $(Q_{2max})_{max}$ от g_a

(——) – $(Q_{1max})_{max}$, (— —) – $(Q_{2max})_{max}$

не всегда. В случае неравных амплитуд колебаний дополнительные фазовые сдвиги Q_{1max} и Q_{2max} поведут себя по-разному (см. рис. 3): в КГ с меньшей амплитудой выходных колебаний дополнительный фазовый сдвиг тока нагрузки будет увеличиваться, а в генераторе с большей амплитудой, наоборот, уменьшаться по сравнению со случаем, когда амплитуды одинаковы ($g_a = 1$).

Данное поведение фазовых сдвигов Q_{1max} и Q_{2max} наглядно иллюстрируется кривыми, приведенными на рис. 4, где представлены зависимости maximum maximorum $\{Q_{1,\,2\,max}\}$ – соответственно $(Q_{1max})_{max}$ и $(Q_{2max})_{max}$ как функции ga. Различие $(Q_{1max})_{max}$ и $(Q_{2max})_{max}$ на рис. 2 при равных амплитудах колебаний, соответствующих случаю $g_a = 1$, обусловлено неидентичностью параметров антенн диапазона СНЧ, характеристики которых были использованы в рассматриваемом примере.

Завершая анализ квазистационарных процессов в выходных цепях КГ, обратимся к выражению (15), которое позволяет оценить экстремумы огибающих тока в выходных цепях генераторов. Как и следовало ожидать, с учетом проведенного выше анализа поведения зависимостей Q_{1max} и Q_{2max} $B_{1,\,2}$ достигают максимума при совпадении частот КГ. В частности, для использованных выше параметров антенн, включая и параметры их электромагнитной связи, а также при условии равенства амплитуд колебаний обоих генераторов, пиковые значения токов в каждом из генераторов возрастают соответственно на 14 и 17 %. При неравных амплитудах, например при $g_a = 0{,}6$, в генераторе с меньшей амплитудой пиковое значение тока нагрузки может возрасти на 29 % по отношению к его номинальному значению, имеющему место при отсутствии электромагнитной связи между антеннами РПдУ. Такое возрастание тока может привести к выходу КГ из строя вследствие перегрузки ключевых приборов. Поэтому условие соблюдения необходимого разноса частот обоих РПдУ важно выполнить не только с целью минимизации дополнительных фазовых сдвигов токов в выходных цепях КГ, но и для уменьшения вероятности возникновения перегрузки полупроводниковых приборов по току.

Таким образом, в результате проведенного анализа процессов в КГ двух РПдУ, работающих на отличающихся частотах, показано, что электромагнитная связь антенн этих РПдУ может привести не только к ухудшению характеристик генераторов, но в ряде случаев и к возникновению аварийных режимов работы. Уменьшить указанные негативные эффекты, обусловленные взаимным влиянием РПдУ друг на друга, можно за счет рационального разноса их рабочих частот, опираясь на приведенные выше результаты расчета.

СПИСОК ЛИТЕРАТУРЫ

1. **Айзинов М.М.** Анализ и синтез линейных радиотехнических цепей в переходном режиме. М.; Л.: Энергия, 1964. 283 с.

2. Теория автоматического регулирования / Под ред. В.В. Солодовникова. Т. 1, 2. М.: Машиностроение, 1985.

3. **Сороцкий В.А., Урлапова О.М., Уткин М.А.** О выборе количества ступеней в выходном напряжении ключевых генераторов сверхмощных РпДУ СНЧ диапазона // Практическая силовая электроника: Науч.-техн. сб. 2004. Вып. 16. С. 36–38.

4. **Велихов Е.П., Жамалетдинов А.А, Собчаков Л.А.** и др. Опыт частотного электромагнитного зондирования земной коры с применением мощной антенны СНЧ-диапазона // Докл. АН. 1994. Т. 338, № 1. С. 106–109.

5. Мощные полупроводниковые приборы. Тиристоры: Справочник / Замятин В.Я., Кондратьев Б.В., Петухов В.М. М.: Радио и связь, 1988. 576 с.

УДК 621.37

В.А. Сороцкий, М.А. Уткин

ИССЛЕДОВАНИЕ ХАРАКТЕРИСТИК МНОГОЯЧЕЙКОВЫХ ТИРИСТОРНЫХ ГЕНЕРАТОРОВ НА ЦИФРОВОЙ МОДЕЛИ

Проведенный в известных работах [1–3 и др.] анализ многоячейковых тиристорных генераторов (ТГ), применяемых для генерирования колебаний повышенной частоты, в основном базируется на представлении тиристора идеальным ключом. Это обусловлено достаточно высокой сложностью аналитических методов расчета ТГ, вследствие чего, в частности, методы расчета, основанные только на допущении об идеальности тиристорных ключей, принято относить к числу точных [4]. Вместе с тем при работе ТГ на повышенных частотах его характеристики существенным образом зависят от проявления нелинейных и инерционных свойств тиристоров, которые не удается учесть при аналитическом исследовании генератора. Кроме того, для улучшения характеристик генератора в схему ТГ могут вводиться дополнительные нелинейные элементы: насыщающиеся дроссели – для ограничения скорости нарастания тока через тиристор и рекуперационные диоды [2] – для стабилизации режима работы генератора при изменении нагрузки и рабочей частоты. Учёт влияния этих элементов с помощью аналитических методов исследования вкупе с упоминавшимися выше их ограниченными возможностями также сопряжен с серьезными трудностями.

Указанные проблемы в значительной степени могут быть устранены при исследовании рассматриваемых генераторов на цифровых моделях. Подобные модели могут быть разработаны, в частности, на базе универсальных пакетов программ схемотехнического моделирования, таких, как Micro-Cap, DesignLab, MATLAB Simulink и др.

Цель настоящей статьи – рассмотреть особенности построения цифровых моделей многоячейковых ТГ при использовании универсальных пакетов программ, а также исследовать влияние нелинейных свойств тиристоров и перечисленных выше дополнительных нелинейных элементов на основные характеристики генераторов. Разработка моделей ТГ и расчеты их характеристик выполнены в среде Micro-Cap. Рассмотренные подходы к построению цифровой модели ТГ могут быть использованы и в других универсальных пакетах программ схемотехнического моделирования.

Модель тиристорного ключа

Модель однооперационного тиристора в пакете Micro-Cap представлена макромоделью со следующим списком параметров:

VALUE = SCR(IH, IGT, TON, VTMIN, VDRM, DVDT, TQ, K1, K2).

Здесь *VALUE* – имя модели; *IH* – ток удержания; *IGT* – ток включения; *TON* – время включения; *VTMIN* – минимальное напряжение во включённом состоянии; *VDRM* – максимальное повторяющееся импульсное напряжение в закрытом

состоянии; *DVDT* – максимально допустимая (критическая) скорость нарастания напряжения в закрытом состоянии; *TQ* – время выключения; *K1* и *K2* – поправочные коэффициенты соответственно для *DVDT* и *TQ*.

Данная макромодель вполне адекватно описывает поведение тиристора при частоте переключения $f \leq 10$ кГц. Однако на более высоких частотах ее использование может привести к заметной погрешности определения потерь в тиристоре и, как следствие, к погрешности вычисления энергетических характеристик ТГ в целом, поскольку она не учитывает ряд важных свойств реальных приборов. В частности, в приведенном выше описании макромодели не отражен в явном виде один из важнейших параметров, характеризующих работу тиристоров на повышенных частотах – время обратного восстановления t_{RR}. Более того, как показали расчеты с использованием библиотечной макромодели тиристора, значение t_{RR}, оцениваемое по току, протекающему через прибор, определяется параметром *TON*.

К числу недостатков библиотечной макромодели тиристора следует отнести и то, что напряжение на включенном тиристоре, задаваемое с помощью параметра *VTMIN*, полагается неизменным. Другими словами, библиотечная макромодель не позволяет учесть инерционность распространения включенного состояния по площади структуры тиристора [5], которая оказывает существенное влияние на величину остаточного напряжения на приборе. Например, для отечественного тиристора серии ТБЧ153-1000-14, как показывают экспериментальные измерения, процесс распространения включенного состояния по площади прибора занимает до 10 мкс, в течение которых напряжение на аноде включенного тиристора изменяется по закону, близкому к экспоненциальному, прежде чем достигнет значения *VTMIN*. Поэтому, как нетрудно определить, для тиристоров указанного типа при частоте генерируемых колебаний $f \geq 25$ кГц длительность интервала распространения включенного состояния, характеризуемого шести – десятикратным превышением эффективного падения напряжения на проводящем приборе по сравнению с *VTMIN*, составляет не менее половины интервала протекания тока через открытый тиристор, в связи с чем пренебрежение данным эффектом неизбежно приводит к заметной погрешности при расчете КПД генератора.

Таким образом, для того чтобы моделирование многоячейковых ТГ с помощью универсальных пакетов программ приводило к адекватным результатам, библиотечная макромодель тиристора нуждается в определенной корректировке, предполагающей, во-первых, установление связи между параметрами t_{RR} и *TON*, и, во-вторых, учёт непостоянства напряжения на включённом тиристоре вследствие эффекта распространения проводящего состояния прибора.

Связь параметра *TON* библиотечной макромодели с временем включения t_{gt} и временем обратного восстановления тиристора t_{RR} может быть установлена путем имитационного моделирования на компьютерной модели испытательного стенда. При этом определение параметров t_{gt} и t_{RR} следует производить в соответствии с ГОСТ 19138.5-85 и 24461-80 в режиме, заданном в технических условиях (ТУ) на прибор. В качестве примера в таблице представлены результаты вычисления времени включения t_{gt} и времени обратного восстановления t_{RR}, полученные при варьировании параметра *TON* в макромодели тиристора серии ТБЧ153-1000-14. Измерение времени обратного восстановления проводилось при прямоугольной форме тока анода с амплитудой $I_T = 500$ А и параметрах, заданных в ТУ на тиристоры указанного типа: максимальном обратном токе $I_{RRM} \approx 80$ А, обратном напряжении $U_{RRM} \approx -100$ В, длительности импульса управления $t_G = 5$ мкс, длительности импульса тока анода $t_i = 20$ мкс.

Результаты вычисления времени

TON, мкс	t_{gt}, мкс	t_{RR}, мкс
0,25	0,28	1,0
0,5	0,50	2,0
1,0	0,99	5,0
1,5	1,46	7,6
2,0	1,92	9,9

Из таблицы видно, что отличие по величине времени включения тиристора t_{gt} (быстрая фаза спада напряжения) и параметра *TON* составляет не более 10 %, а длительность интервала обратного восстановления связана с параметром *TON* соотношением $t_{RR} = (4–5) \cdot TON$. Соответственно, для того чтобы библиотечная макромодель тиристора обеспечивала время обратного восстановления $t_{RR} = 2$ мкс, приведенное в ТУ на тиристор серии ТБЧ153-1000-14, необходимо

задать значение параметра $TON = 0{,}5$ мкс. При этом если параметр TON уже использован, остается открытым вопрос, как задать в макромодели требуемое время включения тиристора, ответ на который будет дан ниже.

Учёт процесса распространения включенного состояния по площади прибора в рассмотренной макромодели тиристора можно осуществить включением последовательно с тиристором источника напряжения экспоненциальной формы с заданным значением постоянной времени [2]. Доработанная с учётом указанного эффекта макромодель тиристора представлена на рис. 1.

В этой макромодели медленный экспоненциальный спад напряжения на тиристоре отражается включением источника напряжения E_1, управляемого напряжением, которое формируется дифференцирующей цепью C_1R_3. Параметры экспоненты выбирают исходя из экспериментальных данных для конкретного типа тиристора. Приведенные на рис. 2 параметры источника дополнительного напряжения получены по осциллограмме процесса установления напряжения при включении тиристора серии ТБЧ153-1000-14.

Дальнейшее усовершенствование макромодели может быть связано с уточнением поведения тиристора при его включении. Необходимый закон изменения напряжения [7], регистрируемого прибором на этапе быстрого спада напряжения в процессе включения, и требуемое значение t_{gt} можно реализовать, например, путём введения второго источника экспоненциального напряжения с требуемой постоянной времени. При синусоидальной форме тока, протека-

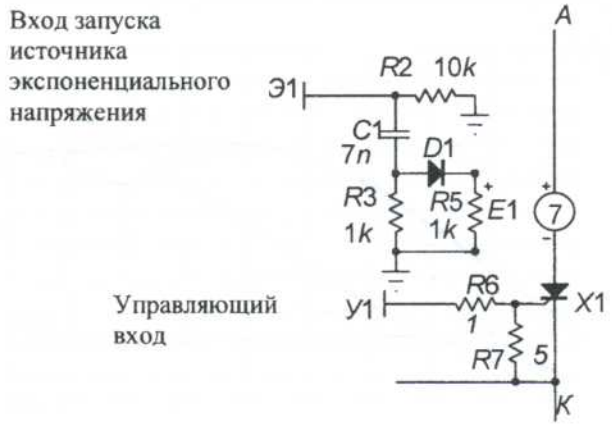


Рис. 1. Эквивалентная схема макромодели тиристорного ключа

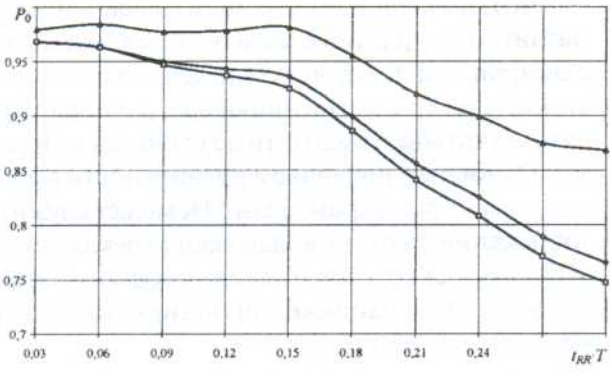


Рис. 2. Зависимости потребляемой мощности от t_{RR}/T

ющего через прибор, точность определения рассеиваемой в нем на этапе включения мощности будет определяться тем, насколько точно в модели отражен именно закон изменения напряжения. Поэтому предложенный подход следует считать вполне оправданным.

Модель насыщающегося дросселя

Насыщающийся дроссель, нередко называемый еще магнитным ключом (МК), применяют для ограничения скорости нарастания тока через тиристор из-за сравнительно медленного распространения включенного состояния по площади прибора [5]. На практике его выполняют в виде проводника, на котором надето требуемое для получения заданного сечения магнитного сердечника число ферритовых колец. Модель насыщающегося дросселя в пакете программ Micro-Cap состоит из двух частей: модели магнитного сердечника и модели катушки индуктивности, намотанной на этом сердечнике.

Минимальный набор параметров, который необходимо задать для описания катушки индуктивности, включает компоненты *PART* и *VALUE*, первая из которых обозначает имя элемента, а вторая, в отличие от случая моделирования линейных катушек, обозначает не величину индуктивности, а число витков, содержащихся в обмотке. Для указанной выше конструкции МК $VALUE = 1$.

В качестве модели магнитного сердечника в программе Micro-Cap используется одна из разновидностей модели Джилса–Атертона, в которой безгистерезисная кривая намагниченности ферромагнетика изменяется по закону гиперболического котангенса в зависимости от напряженности

магнитного поля H [6]. В общем случае в модели магнитного сердечника используется восемь параметров, три из которых определяют геометрические размеры магнитопровода, а оставшиеся пять – учитывают магнитные свойства материала. Основным признаком правильности выбора параметров макромодели МК может служить достижение требуемой задержки переключения МК, обмотка которого состоит из одного витка, при известном напряжении на тиристоре в момент его включения U_0:

$$t_d = \Delta B_M S_M / U_0, \qquad (1)$$

где ΔB_M – полное изменение индукции при перемагничивании магнитопровода; S_M – площадь его сечения.

Цифровая модель генератора

Рассмотренные выше макромодели тиристорных и магнитных ключей были использованы при разработке цифровой модели (ЦМ) трехъячейкового ТГ [1, 2]. Параметры разрядных цепей и сопротивления нагрузки, использованные при моделировании, рассчитаны по методике [1] исходя из условия получения мощности в нагрузке 100 кВт и реализации на рабочей частоте "гладкого" режима, определяемого из условия равенства нулю производной тока тиристора в момент окончания импульса прямого тока. Выбор "гладкого" режима обусловлен тем, что потери во включаемых параллельно тиристорам демпфирующих RC-цепочках достигают в этом режиме минимального значения.

Полученная ЦМ, по существу, универсальна, поскольку позволяет исследовать влияние на основные характеристики ТГ целого набора эффектов, которые, как уже отмечалось выше, не могут быть учтены при использовании аналитических методов расчета. При этом не требуется видоизменять схему ТГ, достаточно лишь скорректировать значение параметров соответствующих элементов схемы. Так, например, чтобы исключить влияние задержки, вносимой МК, с учетом (1) в его макромодели достаточно положить площадь поперечного сечения магнитопровода равной нулю. Аналогично, если требуется исключить учет инерционности распространения включенного состояния по площади прибора в тиристорах и перейти к использованию библиотечной макромодели тиристорного ключа, достаточно устремить к нулю постоянную времени управляемого источника напряжения E1 (см. рис. 1) или его коэффициент передачи.

С помощью ЦМ было проведено исследование влияния на основные характеристики ТГ инерционности тиристора, определяемой длительностью интервала времени обратного восстановления. В число исследуемых характеристик ТГ вошли: потребляемая мощность P_0, мощность в нагрузке $P_н$, электронный КПД генератора η (потери мощности в разрядных цепях не учитывались), мощности потерь соответственно в тиристорном ключе $P_Т$, в демпфирующей RC-цепи $P_{дц}$ и магнитном ключе $P_{мк}$, а также максимальный ток тиристора I_{max} и максимальное напряжение на нём U_{max}. Результаты расчетов, выполненных в диапазоне частот 15–90 кГц, представлены на рис. 2–4. Здесь использованы следующие обозначения: кривые, относящиеся к случаю использования библиотечной макромодели тиристорного ключа, обозначены ромбоидальным маркером; кривые, рассчитанные с помощью модифицированной макромодели тиристора, т. е. учитывающие в том числе и медленный экспоненциальный спад напряжения на включенном тиристоре, выделены с помощью маркеров квадратной формы. Наконец, кривые, соответствующие случаю использования модифицированной модели тиристора и магнитных ключей, выделены с помощью треугольного маркера. Кроме того, для удобства анализа все полученные результаты нормированы к соответствующим значениям, полученным при аналитическом расчете по методике [1].

Проанализируем полученные результаты. В области сравнительно низких частот ($t_{RR}/T \le 0{,}1$), где инерционные свойства тиристоров проявля-

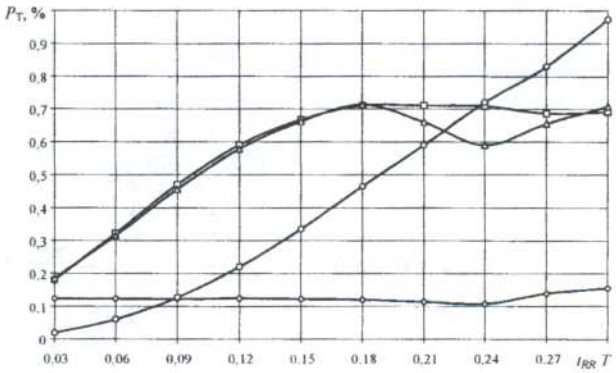


Рис. 3. Зависимости мощности потерь в тиристоре и магнитном ключе от t_{RR}/T

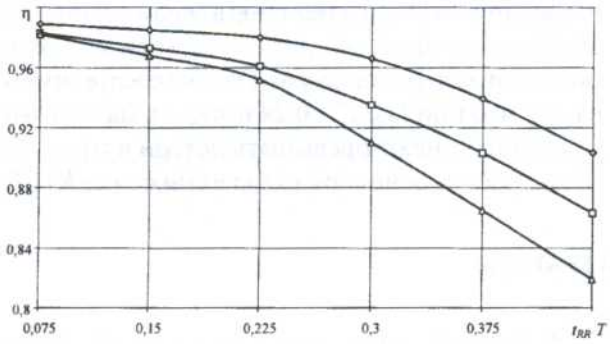


Рис. 4. Зависимости КПД генератора от t_{RR}/T

ются не столь заметно, расчет потребляемой мощности с помощью библиотечной и модифицированной макромоделей дает одинаковые результаты (см. рис. 2). На более высоких частотах ($0{,}1 < t_{RR}/T < 0{,}3$) увеличение остаточного напряжения на тиристорных ключах, учитываемое в модифицированной модели, приводит к небольшому, в пределах 3 %, ожидаемому снижению потребляемой мощности. Гораздо важнее здесь обратить внимание на заметный спад потребляемой мощности, наблюдаемый при $t_{RR}/T > 0{,}15$ и достигающий 17 %. Причина его появления связана с уменьшением по абсолютной величине начального напряжения на коммутирующей емкости в каждом цикле работы ячеек ТГ вследствие ее разряда током обратного восстановления тиристора [8]. Введение МК в схему ТГ не только увеличивает амплитуду тока тиристоров и, соответственно, потребляемую генератором мощность, но и замедляет ее спад по мере роста отношения t_{RR}/T (см. рис. 2). Данное обстоятельство также можно объяснить с точки зрения влияния токов обратного восстановления тиристора: МК существенно ограничивает их величину.

Приведенные на рис. 3 зависимости мощности потерь в тиристорном ключе, нормированной к номинальной расчетной мощности ТГ, подтверждают актуальность разработки модифицированной макромодели тиристора: пренебрежение эффектом распространения проводящего состояния по структуре тиристора может привести к существенной погрешности. В частности, в случае использования тиристоров серии ТБЧ153-1000-14 мощность потерь в приборе при $t_{RR}/T > 0{,}18$ возрастает приблизительно в пять раз по сравнению с тем случаем, когда этим эффектом пренебрегают. Заслуживает внимания и то, что добавление в схему ТГ насыщающихся дросселей не приводит к заметному снижению рассеиваемой в тиристорах мощности (см. рис. 3). Это можно объяснить тем, что МК ограничивает ток на начальном этапе включения тиристора на протяжении всего нескольких микросекунд (как правило, 1–4 мкс). Поэтому основной эффект, создаваемый МК, заключается в ограничении скорости нарастания тока и, как следствие этого, в ограничении мгновенной рассеиваемой мощности в тиристоре на указанном временном интервале. При усреднении рассеиваемой в тиристоре мощности за период вклад МК в уменьшение этого показателя оказывается не столь заметен. Интересно другое обстоятельство, связанное с МК. Как следует из рис. 3 (кривая с маркерами круглой формы), мощность, рассеиваемая в самом МК, включая и шунтирующее его сопротивление, растет достаточно быстро и при $t_{RR}/T > 0{,}24$ начинает превалировать над мощностью, рассеиваемой в самом тиристоре. Таким образом, оказывается, что повышение надежности работы тиристоров за счет ограничения скорости нарастания тока в них с помощью МК на повышенных частотах может сопровождаться ухудшением КПД генератора, что и подтверждается зависимостями, приведенными на рис. 4. С помощью этих зависимостей нетрудно оценить, что при повышении параметра t_{RR}/T в пределах с 0,2 до 0,3 снижение КПД, обусловленное ростом потерь в МК, увеличивается с 3 до 7 %.

Подводя итоги, можно формулировать следующие выводы. Макромодели тиристоров, входящие в состав встроенных библиотек универсальных пакетов программ схемотехнического моделирования, обеспечивают приемлемую точность только в диапазоне сравнительно низких частот, характеризуемых условием $t_{RR}/T \le 0{,}1$. Для того чтобы макромодели адекватно отражали свойства реальных приборов в области более высоких частот, их необходимо модифицировать в части более точного учета изменения остаточного напряжения на включенном тиристоре, а также тока обратного восстановления. Несоблюдение первого из условий может привести к заметной ошибке при определении мощности потерь в тиристорных ключах, обусловив заниженные значения. Так, для тиристоров серии ТБЧ153-1000-14 указанное различие при расчете мощности потерь, как показали расчеты, может достигать пятикратной величины.

Влияние времени обратного восстановления тиристоров на основные характеристики ТГ начинает заметно проявляться при $t_{RR}/T > 0{,}15$ и связано с ухудшением условий перезаряда емкости коммутирующего контура в каждой ячейке ТГ. Применение магнитных ключей с целью ограничения скорости нарастания тока через тиристор в целом достаточно слабо влияет на общий уровень средней за период мощности потерь в приборах. Однако следует иметь в виду, что при $t_{RR}/T > 0{,}24$ потери в магнитных ключах начинают превышать потери в тиристорах и вносить основной вклад в снижение КПД.

СПИСОК ЛИТЕРАТУРЫ

1. **Артым А.Д.** Ключевые генераторы гармонических колебаний. Л.: Энергия, 1972. 170 с.

2. **Бальян Р.Х., Сиверс М.А.** Тиристорные генераторы и инверторы. Л.: Энергоиздат, 1982. 223 с.

3. **Дмитриков В.Ф., Тонкаль В.Е., Островский М.Я.** Теория ключевых формирователей гармонических колебаний. Киев: Наукова думка, 1993. 312 с.

4. **Мерабишвили П.Ф., Ярошенко Е.М.** Нестационарные электромагнитные процессы в системах с вентилями. Кишинев: Штиница, 1980. 208 с.

5. **Герлах В.** Тиристоры: Пер. с нем. М.: Энергоатомиздат, 1985. 328 с.

6. **Разевиг В.Д.** Система сквозного проектирования электронных устройств DesignLab 8.0. М.: Солон-Р, 2000. 698 с.

7. **Дерменжи П.Г., Кузьмин В.А., Крюкова Н.Н.** и др. Расчет силовых полупроводниковых приборов / Под ред. В.А. Кузьмина. М.: Энергия, 1980. 184 с.

8. **Сороцкий В.А.** Влияние инерционности тиристорных ключей на энергетические характеристики многоячейкового генератора // Радиотехника. 2003. № 4. С. 68–70.

УДК 681.3.06

Е.А. Макарова

УПРАВЛЕНИЕ ПОВЕДЕНИЕМ МАКРОЭКОНОМИЧЕСКИХ АГЕНТОВ НА ОСНОВЕ МЕХАНИЗМА САМОВЫРАВНИВАНИЯ ФИНАНСОВЫХ ПОТОКОВ

Проблемы исследования динамики воспроизводственных процессов на макроэкономическом уровне приобретают особую остроту в условиях, когда системный кризис экономики России, сопровождающийся нарушениями основных воспроизводственных пропорций, был смягчен, но до конца не преодолен, а затем усугубился под влиянием мирового финансового кризиса. В рамках этого направления проводятся интенсивные исследования по созданию экономико-математических моделей и программных инструментариев для решения различных проблем макроэкономического анализа [1].

Для задач анализа пропорций воспроизводственного процесса и механизмов распределения финансовых потоков в динамике разрабатываются макроструктурные модели с использованием различных инструментариев: межотраслевого баланса; интегрированных социальных матриц – SAM-матриц (social account matrix); системной динамики, вычислимых моделей общего равновесия – CGE-моделей (Computable General Equilibrium), моделей межвременного равновесия, SFC-моделей (stock-flow consistent), макроэкономических и межотраслевых моделей, разрабатываемых международным сообществом INFORUM [2–9].

Несмотря на большое количество работ в исследуемом направлении, следует признать, что из-за сложности решаемых проблем некоторые особенности динамики воспроизводственных процессов не нашли полного отражения в разрабатываемых модельных комплексах. В частности, вопросы анализа динамики воспроизводственных процессов, которые рассматриваются с учетом замыкания макроэкономического кругооборота финансовых потоков и влияния запасов макроэкономических агентов, остаются недостаточно изученными.

Чтобы обеспечить поддержку процедур анализа и формирования сценариев управления макроэкономической системой (МЭС), разрабатывается информационно-аналитическая система имитационного моделирования (ИАСИМ) процессов функционирования МЭС на основе разработанного ранее комплекса динамических моделей макроэкономического кругооборота потоков, формирующих воспроизводственный процесс [10]. Отметим, что под макроэкономической системой понимается система общественного производства, осуществляемого в масштабах страны, выполняемого взаимодействующими секторами экономики (агентами) и регулируемого органами государственного управления.

Комплекс динамических моделей макроэкономического кругооборота включает четыре динамические модели поведения макроэкономических агентов: фирм, домохозяйств (населения), банков и государства [11]. Взаимосвязанное функционирование макроэкономических агентов охватывает все стадии воспроизводственного процесса и формирует макроэкономический кругооборот [1]. В данной статье решается задача разработки алгоритма управления поведением одного отдельно взятого макроэкономического агента на основе информации о запасах накопленных ранее ресурсов. Для соблюдения целостности представления материала предварительно необходимо кратко изложить системные принципы построения динамических моделей поведения макроэкономических агентов, среди которых выделены следующие.

Во-первых, макроэкономическим агентом (МЭА) называется действующее лицо в экономике, которое может принимать экономические решения, выступать в качестве производителя или покупателя экономических благ (товаров, услуг). Каждый макроэкономический агент: имеет запасы финансовых ресурсов; в соответствии с его ролью в воспроизводственном процессе выполняет один или несколько функциональных процессов, что сопровождается формированием расходов (оттоков) финансовых ресурсов; получает доходы (притоки) по одному или нескольким каналам от других МЭА; формирует сальдо темпов доходов и расходов; формирует запасы, которые в зависимости от знака сальдо либо увеличиваются, либо уменьшаются. Отметим, что согласно логике макроэкономического кругооборота построение замкнутого воспроизводственного цикла обеспечивается за счет соблюдения правила: расходы одного МЭА обязательно становятся доходами другого МЭА.

Во-вторых, ситуация, при которой сальдо потоков равно нулю и объем запасов не изменяется, соответствует *динамически равновесному режиму* функционирования МЭА. Формирование условий сбалансированности доходов и расходов для каждого МЭА позволяет вывести основные известные макроэкономические тождества [1]. Только при соблюдении баланса расходов и доходов по всем МЭА одновременно возможно функционирование всей МЭС в целом на динамически равновесных режимах [10].

В-третьих, *динамически неравновесные режимы* функционирования МЭС возникают при нарушении равенства расходов и доходов хотя бы по одному МЭА. Восстановление равновесия МЭС осуществляется с помощью механизма самовыравнивания, который реализован при построении модели управляемого поведения МЭА.

Модель управляемого поведения МЭА разработана в классе динамических нелинейных моделей с логическими элементами. Реализация модели выполнена с помощью библиотек приложения Simulink среды MATLAB, а также с использованием модулей, логика работы которых программируется на языке MATLAB [11].

Обобщенная схема функциональной модели управляемого поведения МЭА представлена в виде параллельных каналов формирования доходов и нескольких, тоже параллельных, каналов формирования расходов для реализации функциональных процессов. Цель управления каждым функциональным процессом формирования расходов состоит в обеспечении планового темпа расхода ресурсов. Формирование плана осуществляется в три этапа. Первый этап связан с учетом автономных, не зависящих от информации о поведении других секторов, планов по расходам (например, автономное потребление населением). На втором этапе учитывается эта информация, например, в виде зависимости потребления населения от располагаемых доходов, поступающих от фирм как оплата труда. Полученные на этом этапе планы должны быть сбалансированы по потокам для обеспечения равновесного режима. Третий этап – корректировка планов на основе информации о запасах с помощью алгоритма, основанного на механизме самовыравнивания финансовых потоков.

Рассмотрим особенности работы алгоритма на примере упрощенной модели МЭА, который выполняет один функциональный процесс (имеет один отток) и получает доход по одному каналу (имеет один приток). Это необходимо для обеспечения наглядности выполняемых далее преобразований функциональной схемы модели.

С позиций экономического содержания функциональная схема модели управляемого поведения i-го МЭА (рис. 1) должна быть представлена в виде параллельного соединения каналов формирования притока (дохода) с темпом $\dot{In}_i(t)$ и оттока (расхода) с темпом $\dot{Out}_i(t)$, а также последовательно соединенных сумматора для вычисления сальдо потоков $\dot{Fl}_i(t)$ и интегратора для расчета текущего объема $St_i(t)$ накопленных запасов финансовых ресурсов. Интегратор в этом способе представления схемы является завершающим элементом. Блок реализации функционального процесса представлен непрерывной динамической моделью, построенной в классе управляемых систем на основе принципа комбинированного управления.

Особенность предлагаемой модели заключается в построении канала расхода ресурсов, который содержит логический блок f_i, реализующий алгоритм корректировки планового темпа расхода $\dot{R}_i^0(t)$ на основе информации о текущем располагаемом объеме запасов $St'_i(t)$. Вычисление коэффициента k_{ci} выполняется по формуле

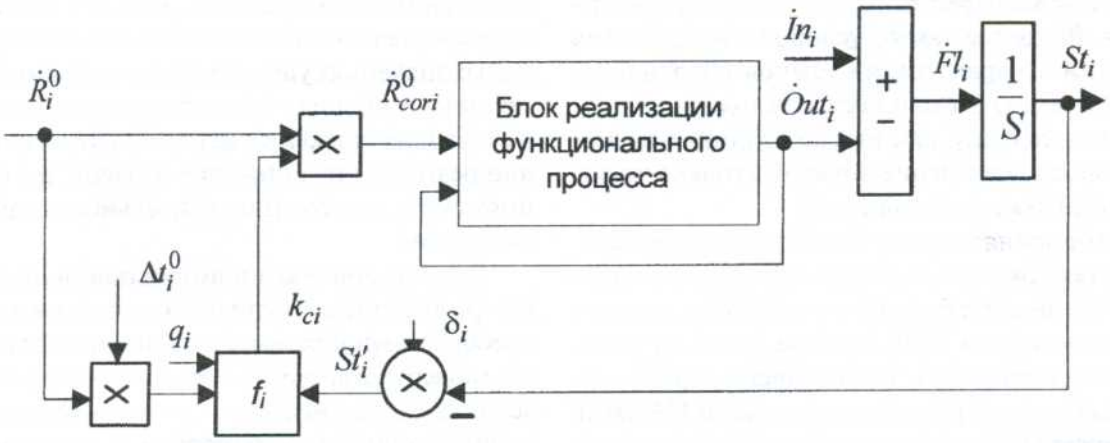


Рис. 1. Функциональная схема модели управляемого поведения i-го МЭА (с позиций экономической теории)

$$k_{ci} = \begin{cases} 1, \text{если } \dot{R}_i^0 \Delta t_i^0 \le St_i' \le \dot{R}_i^0 \Delta t_i^0 q_i \\ \dfrac{St_i'}{\dot{R}_i^0 \Delta t_i^0}, \text{если } St_i' < \dot{R}_i^0 \Delta t_i^0 \\ \dfrac{St_i'}{\dot{R}_i^0 \Delta t_i^0 q_i}, \text{если } St_i' > \dot{R}_i^0 \Delta t_i^0 q_i \end{cases}, \qquad (1)$$

где $St'_i(t) = Sti(t) - \delta_i(t)$, $\delta_i(t)$ – неприкосновенный запас; Δt_i^0 – интервал времени, в течение которого выполнение функционального процесса должно быть обеспечено запасами ресурсов; q_i – коэффициент, определяющий, во сколько раз текущий объем запасов должен превышать требуемый для того, чтобы принять решение об увеличении темпов расхода ресурсов (рис. 2).

Алгоритм вычисления коэффициента k_{ci} реализован программно на языке MATLAB.

Вычисление корректирующего коэффициента k_{ci} выполняется на основе анализа информации об объеме текущих запасов. При этом проводится анализ достаточности текущих запасов для расхода ресурсов с плановым темпом в течение времени Δt_i^0 ($\Delta t_i^0 \ge 1$). Причины изменений объема запасов заключаются либо в изменении темпов поступления доходов от других МЭА, либо в корректировке планов в связи с изменением поведения других агентов. Любое изменение объема запасов свидетельствует о наступлении динамически неравновесного режима функционирования МЭА.

В случае недостаточности запасов (ситуации для зоны *1*, $k_{ci} < 1$) выполняется снижение темпов с коэффициентом, умножение на который дает возможный темп, такой, что обеспечивается текущими запасами на планируемый интервал времени. Снижение темпов расхода ресурсов, связанное с уменьшением запасов, – вынужденное, отследить момент возникновения таких ситуаций для человека (системного аналитика) не представляется возможным, поэтому для алгоритма управления "на уменьшение" применяется автоматический режим. Контроль условия достаточности запаса ресурсов осуществляется на каждом шаге моделирования в течение всего периода моделирования.

Для благоприятных ситуаций роста запасов (ситуации для зоны *3*, $k_{ci} > 1$) при превышении запасами некоторой пороговой величины, которая в q ($q \ge 1$) раз больше требуемого объема,

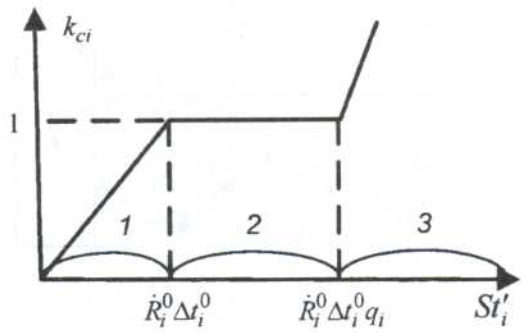


Рис. 2. График нелинейной зависимости $k_{ci} = f_i(St_i')$

выполняется пересчет планов в сторону увеличения. Введение такого условия для принятия решений о наращивании темпов объясняется тем, что темпы расхода ресурсов должны увеличиваться не сразу, как только возникло незначительное превышение запасов, а только по достижении некоторого предела.

Такое принятие решение об увеличении плановых темпов расхода ресурсов должно основываться на анализе большого количества данных о состоянии всей МЭС в целом, поэтому алгоритм корректировки "на увеличение" реализован в автоматизированном режиме. В ИАСИМ предложены два режима работы алгоритма управления: режим корректировки темпов расхода ресурсов как "на уменьшение", так и "на увеличение" (флаг $regU_i = 2$), и режим корректировки темпов расхода ресурсов только "на уменьшение" (флаг $regU_i = 1$).

Для ситуаций зоны *2* ($k_{ci} = 1$) МЭА поддерживает плановый темп расхода ресурсов $\dot{R}_i^0(t)$, несмотря на колебания объема запасов в допустимых пределах, определяемых формулой (1). Размер зоны нечувствительности (резистентности агента к внешним возмущениям) определяется как горизонтом планирования запасов Δt_i^0, так и параметром q. Поскольку от значений этих параметров зависит характер поведения МЭА на динамически неравновесных режимах, отнесем их к управляющим и назовем параметрами управления "демпфирующими" запасами. Значения этих параметров определяют не только объем запасов, при котором начинаются какие-либо изменения темпов расхода, но и, как следствие, время начала этих изменений. Например, при малых значениях управляющих параметров, т. е. при малых объемах "демпфирующих" запасов, в ситуациях снижения запасов МЭА очень поздно реагирует на снижение запасов, а в ситуациях роста запасов (положительного сальдо) – очень рано.

С целью анализа динамики поведения МЭА в перечисленных группах ситуаций выполним преобразования схемы. Рассмотренный вид схемы модели согласуется с ее экономическим содержанием, сформулированным ранее в соответствии с экономической теорией. Эксперименты, проведенные для различных форм изменения во времени темпа дохода $\dot{In}_i(t)$, показали, что предложенный алгоритм корректировки плановых темпов расхода ресурсов обеспечивает приближение темпа расхода ресурсов $\dot{Out}_i(t)$ к форме изменения темпа поступления дохода $\dot{In}_i(t)$. Функциональная схема модели, представленная с позиций теории управления (рис. 3), объясняет этот эффект.

Схема получена перемещением сумматора, вычисляющего сальдо $\dot{Fl}_i(t)$ (а теперь ошибку), и интегратора, вычисляющего объем запаса $St'_i(t)$, в начало схемы. Теперь очевидным становится наличие второго контура управления по обратной связи с задающим воздействием в виде темпа $\dot{In}_i(t)$ дохода (притока). Следовательно, модель управляемого поведения МЭА двухконтурная и предназначена для управления темпом расхода ресурсов $\dot{Out}_i(t)$. В ситуациях, когда $k_{ci} = 1$, цель управления запишется: $\dot{Out}_i(t) \to \dot{R}_i^0(t)$,

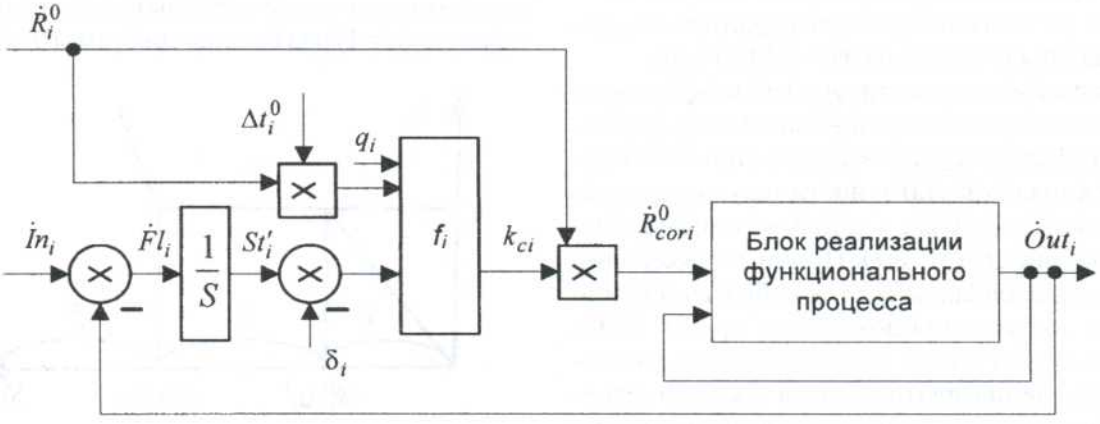


Рис. 3. Функциональная схема модели управляемого поведения *i*-го МЭА (с позиций теории управления)

в остальных ситуациях $\dot{O}ut_i(t) \to \dot{I}n_i(t)$. Происходит как бы переключение цели управления: для зоны нечувствительности это отработка первоначального плана; для наклонных участков нелинейной характеристики – приближение темпов расхода к текущим темпам получения дохода. Иными словами, на динамически неравновесных режимах обеспечивается самовыравнивание темпов потоков расходов и доходов макроэкономического агента.

Эффект самовыравнивания может быть объяснен и с третьей стороны – с позиций теории систем, если привлечь к анализу системный принцип неравновесного состояния [12]. Согласно последнему система находится в динамически неравновесном состоянии, если нарушено равенство между двумя противоположно направленными силами системы: $\frac{dX}{dt} = k\big(F(+) - F(-)\big)$. Для МЭА в качестве силы $F(+)$, способствующей движению системы, рассматриваются темпы формирования доходов. В общем случае при моделировании работы всей МЭС в целом график ее имеет произвольную форму, в простейшем случае – это прямая (рис. 4). В качестве противодействующей силы $F(-)$ рассматриваются темпы расходов. Вид поточно-запасной характеристики для темпа расхода ресурсов повторяет нелинейную зависимость (1). Регулирующим системным фактором X выступают запасы. Тогда уравнение движения системы примет вид:

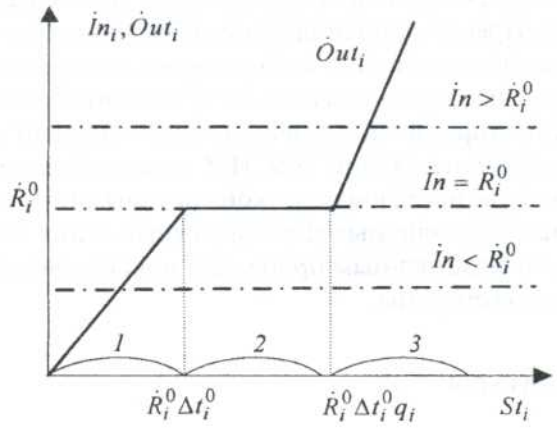


Рис. 4. Графический вид зависимостей силы, способствующей (темпа дохода $\dot{I}n_i(t)$) и противодействующей (темпа расхода $\dot{O}ut_i(t)$) движению МЭА (с позиций теории систем)

$$\frac{dSt_i}{dt} = \big(\dot{I}n_i(t) - \dot{O}ut_i(t)\big). \qquad (2)$$

Преобразования уравнения (2), выполненные при условии модели идеальной реализации функционального процесса с коэффициентом передачи, равным единице, а также при нулевом неприкосновенном запасе $\delta_i(t)$, позволяют получить дифференциальное уравнение инерционного звена, входом которого является темп дохода $\dot{I}n_i(t)$, а выходом – объем запасов $St_i(t)$. Коэффициент передачи k_p и постоянная времени τ определяются управляющими параметрами Δt_i^0 и q. Для ситуаций группы *1* $k_p = \Delta t_i^0$; $\tau = \Delta t_i^0$, а для группы *3* – $k_p = \Delta t_i^0 q_i$; $\tau = \Delta t_i^0 q$. Введем коэффициент $\rho = \frac{1}{k_p} = \frac{1}{\tau}$ (равный тангенсу угла наклона кривой расхода) и назовем его коэффициентом самовыравнивания, так как его увеличение способствует более быстрому достижению нового равновесного состояния, и наоборот. Поведение агента тем более реактивно, т. е. тем быстрее наступит равновесие темпов потоков и тем меньше новое равновесное состояние будет отличаться от прежнего, чем больше коэффициент самовыравнивания ρ. Однако при этом согласно логике планирования запасов объем оставшихся запасов будет меньше за счет меньшего коэффициента передачи, т. е. меньших значений управляющих параметров Δt_i^0 и q. Иными словами, более высокая реактивность поведения присуща агенту в случае, если "демпфирующих" запасов у него меньше, и наоборот. При этом большая реактивность МЭА позволяет ему реагировать изменением расходов позже – при снижении запасов и раньше, если запасы растут.

Таким образом, полученные три варианта интерпретации модели управляемого поведения МЭА позволяют сформировать триединый взгляд на предлагаемую модель, которая воплотила в себе закономерности и экономической теории, и теории управлении, а также теории систем. Полученная триада теорий, формирующих математическую основу описания поведения МЭА, представляет собой отображение триады подходов, составляющих основу концепции исследования и моделирования МЭС. Концепция основана на интеграции подходов: системного, динамического (воплощенного в теории управления), а также когнитивного и сценарного, нашедших отражение в экономической теории [10].

Проведены исследования модели поведения МЭА для ситуаций группы *1* с отрицательным сальдо потоков. Применяется режим работы алгоритма только "на уменьшение" $regU_i = 1$, плановый темп дохода $\dot{I}n^0 = \dot{R}_i^0 = 10$, объем запаса $St_i^0 = 59$, "страховочный" период планирования $\Delta t_i^0 = 3$. На рис. 5 представлены: базовый эксперимент 1 с возмущением, подаваемым в момент времени $t = 23$, в виде снижения темпа дохода $\Delta\dot{I}n = -5$; эксперимент 2, отличающийся начальным значением запаса $St_i^0 = 95$; эксперимент 3, в котором увеличен горизонт планирования запасов $\Delta t_i^0 = 5$. Первые два эксперимента не отличаются по динамическим характеристикам (равные постоянные времени и коэффициенты передачи, новые установившиеся значения запасов $St_i = 15$). Только момент начала снижения темпов для более высоких начальных запасов в эксперименте 2 наступает позже. В третьем эксперименте за счет большего резервирования запасов агент становится более инерционным, медленнее реагирует на снижение запасов, но раньше; поэтому и запасов сохраняет при восстановлении равновесного режима больше $St_i = 25$ (объясняется большим коэффициентом передачи). Равновесие восстановлено $\dot{I}n_i(t) = \dot{O}ut_i(t)$ во всех ситуациях.

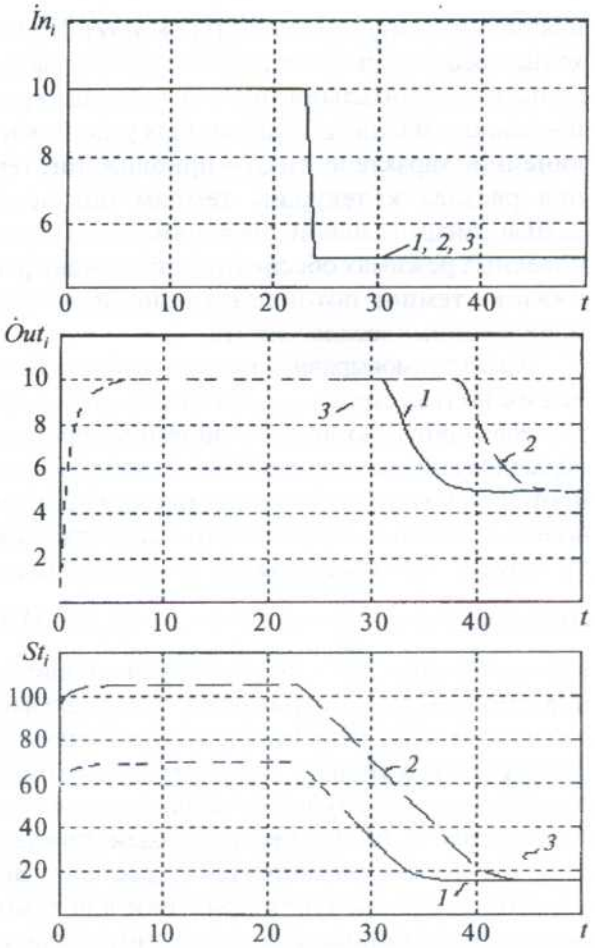


Рис. 5. Графики переходных процессов функционирования МЭА

Таким образом, предложенный алгоритм управления поведением макроэкономических агентов обеспечивает самовыравнивание темпов доходов и расходов на неравновесных режимах функционирования. Предложенные варианты интерпретаций функциональной схемы модели поведения МЭА позволили: объяснить экономическое содержание структуры модели, сформулировать цели управления в различных неравновесных ситуациях, а также описать динамические особенности процесса перехода МЭА из одного динамически равновесного состояния в другое. Исследование динамики взаимосвязанного функционирования МЭА в составе МЭС показало, что неравновесные режимы, возникшие в поведении одного агента, распространяются на других агентов. Предложенный алгоритм обеспечивает восстановление динамически равновесного режима функционирования всей МЭС в целом. Применение управляемых динамических моделей МЭА и разрабатываемой ИАСИМ целесообразно в качестве исследовательской системы для целей анализа различных сценариев управления воспроизводственным процессом макроэкономической системы.

СПИСОК ЛИТЕРАТУРЫ

1. **Тарасевич Л.С., Гребенников П.И., Леусский А.И.** Макроэкономика: Учебник. 7-е изд., испр. и доп. М.: Высшее образование; Юрайтиздат, 2009. 654 с.

2. **Макаров В.Л., Бахтизин А.Р., Сулакшин С.С.** Применение вычислимых моделей в государственном управлении. М.: Научный эксперт, 2007. 304 с.

3. **Попков Ю.С.** Макросистемные модели пространственной экономики. М.: Комкнига, 2008. 240 с.

4. **Абрамова Е.А., Белоусов Д.Р., Михайленко К.Е.** Экономические итоги развития российской экономики в 2006 г. и прогноз на 2008–2010 гг. // Проблемы прогнозирования. 2008. № 1. С. 55–72.

5. **Петров А.А., Поспелов И.Г., Шананин Х.А.** Опыт математического моделирования экономики. М.: Энергоатомиздат, 1996. 558 с.

6. **Узяков М.Н.** Отрасль в системе межотраслевых связей: возможности анализа и прогнозирования. М.: ТЕИС, 2002. 224 с.

7. **Клоцвог Ф.Н., Костин В.А.** Макроструктурные модели – инструмент народнохозяйственного прогнозирования // Проблемы прогнозирования. 2004. № 6. С. 17–27.

8. **Dos Santos C.H., Zezza G.** (2004) A Post-Keynesian Stock-flow Consistent Macroeconomic Growth Model Preliminary Results [WWW document] / The Levy Economics Institute <http://ideas.repec.org/p/lev/wrpap/402.html> (01.09.2009).

9. **Миронова Е.С.** Обзор итогов XVI международной конференции по межотраслевому моделированию INFORUM // Проблемы прогнозирования. 2009. № 2. С. 162–164.

10. **Ильясов Б.Г., Дегтярева И.В., Макарова Е.А., Габдуллина Э.Р.** Концепция системного моделирования процессов кругооборота денежных потоков // Проблемы управления и моделирования в сложных системах: Тр. X Междунар. конф. Самара: Изд-во Самарского НЦ РАН, 2008. С. 153–160.

11. **Ильясов Б.Г., Дегтярева И.В., Макарова Е.А., Габдуллина Э.Р.** Моделирование динамики кругооборота финансовых потоков с учетом накопления финансовых ресурсов // Вестник компьютерных и информационных технологий. 2009. № 1 (61). С. 28–38.

12. Системный анализ и принятие решений. Словарь-справочник: Учеб. пособие для вузов / Под ред. В.Н. Волковой, В.Н. Козлова. М.: Высш. шк., 2004. 616 с.

УДК 007:681.518.2

В.Л. Токарев, С.В. Орлов

ЗАДАЧА АВТОМАТИЗАЦИИ ПРОЦЕССА ФОРМИРОВАНИЯ ЛИНГВИСТИЧЕСКИХ МОДЕЛЕЙ ДЛЯ ПОДДЕРЖКИ ПРИНЯТИЯ РЕШЕНИЙ

Интегрированные системы поддержки принятия решений (ИСППР) обладают большими возможностями для компьютерной поддержки решений сложных, трудноформализуемых задач. В отличие от известных интеллектуальных систем (например, экспертных) ИСППР по имеющейся информации выстраивают приближенные модели системы, относительно которой принимается решение (СОПР), с точностью, достаточной для получения на ее основе рационального решения [1]. Получаемые модели используются для ряда этапов поддержки принятия решений. Их применение в качестве элементов знаний наряду с экспертными оценками дает возможность использовать математические процедуры поиска решений, делает систему гибкой и универсальной, обеспечивают повышение уровня данных, позволяет оперировать с качественными переменными и др. Эти свойства ИСППР обеспечивают им широкую область применения и позволяют решать такие трудноформализуемые задачи, как управление сложными СОПР; прогнозирование их состояния, недоступного для непосредственных измерений; диагностика в условиях неопределенности и др.

Большинство известных методов и подходов к решению задач приближенного моделирования, рассматриваемых, например, в работах [2, 3] и многих других, в общем случае предполагает проведение процедур кластеризации информативных признаков (состояний СОПР) с заранее задаваемыми параметрами кластеров. При этом инженер по знаниям оказывает сильное влияние на процесс формирования моделей, качество которых будет существенно зависеть от знания им предметной области и субъективных факторов.

Предлагаемый в данной статье метод построения приближенной (лингвистической) модели предметной области СОПР заключается в полностью автоматическом определении (итерационным приближением) наилучших параметров, при которых достигается экстремальный уровень

адекватности модели, обусловливаемый достижением глобального минимума значения предложенного тестового критерия. Он включает два этапа: начальный и настроечный.

Целью *начального* этапа является получение на основе имеющейся априорной информации I_a и некоторой обучающей выборки данных W_N^O лингвистической модели $L_N^{(1)}$ в первом приближении с помощью метода приближенной формализации, описанного в [1]. Суть метода заключается в следующем.

1. Из множества выходных переменных Y выбирается подмножество таких, значения которых существенно влияют на качество принимаемых решений. Для упрощения будем считать, что такое подмножество состоит из одной переменной y ($p = 1$).

Формируется обучающая выборка данных $W_N^O = [X^n \mid Y]$, представленная в виде матрицы размерности $N \times (n + 1)$. Здесь n – число информативных входных переменных, N – число наблюдений (строк матрицы). Каждая строка $[x_1, x_2, ..., x_n, y]$ матрицы W_N^O соответствует одному моменту наблюдения.

2. Все значения переменных $x_1, x_2, ..., x_n, y$ переводятся в шкалу термов соответствующих лингвистических переменных (ЛП):

$$L(x_i) = \langle A_i, \mu_{A_i}(x_i) \rangle, x_i \in X_i \to A_i, A_i = \{a_{i1}, a_{i2}, ..., a_{ij}, ..., a_{im_i}\},$$

$$L(y) = \langle B, \mu_B(y) \rangle, y \in Y \to B, B = \{b_1, b_2, ..., b_j, ..., b_{m^o}\}.$$

Изначально формы функций принадлежностей $\mu_{A_i}(x_i)$ и $\mu_B(y)$ принимаются треугольного вида с пересечением на уровне β (см. рисунок). Значение β определяет степень размытости ЛП. Оно рассчитывается для каждой переменной исходя из типа шкалы, в которой могут быть измерены ее значения, или субъективно, исходя из уверенности ЛПР в оценках значений.

3. С учетом представления значений переменных x_i и y в виде термов ЛП $L(x_i)$ и $L(y)$ матрица W_N^O преобразуется в матрицу $L_N^{(0)} = [L(x) \mid L(y)]$, в которой каждая строка $l_j(x, y) \in L(x, y), j = 1, ..., N$ представляет собой импликацию

$$l_j(x, y) = \left(\bigwedge_{i=1}^{n} a_{ik} \right) \to b_q, \qquad (1)$$

$$a_{ik} \in A_i, \quad b_q \in B, \quad j = 1, ..., N,$$

где $A_i = \Im(x_i)\{a_{i1}, a_{i2}, ..., a_{ik}, ..., a_{im_i}\}$; $B = \Im(y)\{b_1, b_2, ..., b_q, ..., b_{m^o}\}$; $\Im(x_i)$, $\Im(y)$ – нечеткие преобразования значений x_i, y в значения ЛП, соответственно в A_i и B.

Импликации группируются в подмножества G_s, $s = 1, ..., p$, $p \le m^o$, каждое из которых представляет собой набор правил определения s-го терма выходной ИП.

4. Для каждой строки $l_j(x, y) \in L(x, y)$, ($j = 1, ..., N$) определяется информационная мера

$$d(c_{js}) = q_{js} \cdot (1 - \overline{q}_{js}), \qquad (2)$$

которая оценивает принадлежность конъюнкции $c_{js} = \bigwedge_{i=1}^{n} a_{iv}$, $v \in 1, ..., m_i$, $c_{js} \in l_j(x, y)$ лингвистической модели. Здесь q_{js} и $\overline{q}_{js}$ – вероятности различения и неразличения конъюнкцией c_{js} выходного терма b_s. В результате образуется матрица $L_N^{(1)} = [L(x) \mid L(y) \mid d(c_{js})]$ – первое приближение лингвистической модели.

5. По контрольной выборке данных $W_M^K = [X^n \mid Y]$ оценивается степень соответствия модели $L_{N'}^{(1)}$ реальной системе путем вычисления значения тестового критерия

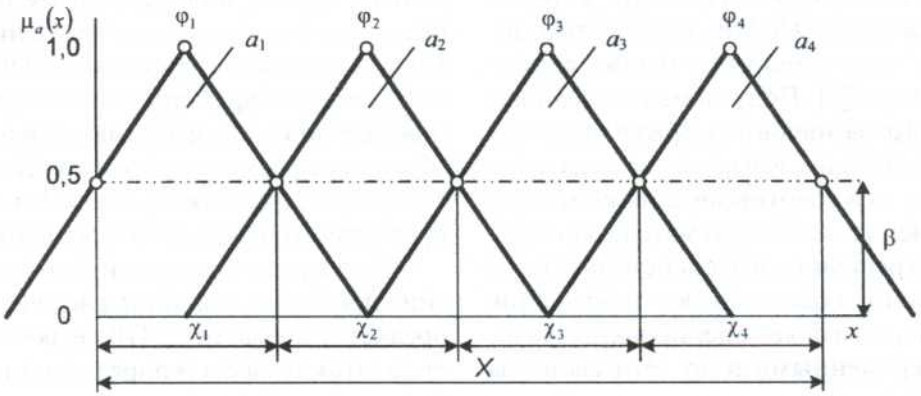


График функций принадлежности для четырех интервалов

$$J(e) = 1/N\sum_{i=1}^{N}|e_i|, \quad e_i = q_i - \overline{q}_i \quad q \in 1, ..., m^o, \quad (3)$$

который должен быть равен нулю при полном соответствии модели моделируемой системе. Здесь q_i и $\overline{q}_i$ – значения индексов термов b_s выходной ЛП, полученные соответственно из матриц W_M^K и W_N^O с помощью модели $L_{N'}^{(1)}$.

На *настроечном этапе* выполняется коррекция функций принадлежности термов входных количественных переменных с целью достижения глобального минимума значения критерия (3).

Задача определения наилучших параметров модели на этом этапе ставится следующим образом. Имеется n' входных количественных переменных модели с заранее заданными степенями размытости $\beta_i \in B$ их лингвистических переменных $L(x_i)$, каждая из которых имеет m_i базовых термов, $i = 1...n'$. Термы имеют функции принадлежности $f_{ij} \to [0; 1]$, определяющие степень принадлежности $\mu_{ij}(x)$ каждого значения i-й переменной j-му терму соответствующей ЛП. Графически все функции принадлежности имеют треугольную форму и образуются двумя прямыми линиями, пересекающимися в точке $(\varphi_j; 1)$ и расположенными друг от друга на расстоянии χ_j на уровне β_i по оси абсцисс (см. рисунок), $\varphi_j \in \Phi_i$, $\chi_j \in X_i$, $j = 1...m_i$, $i = 1...n'$. Требуется найти такие φ_j и χ_j, чтобы достичь глобального минимума значения критерия $J(e)$.

Задача определения наилучших параметров функции $J(\Phi, X) \to \min$ разбивается на две последовательные задачи многомерной оптимизации:

$$J(\Phi) \to \min; \quad (4)$$

$$J(X) \to \min \text{ при } \chi_j > 0. \quad (5)$$

Их решение может быть найдено при помощи градиентных методов в сочетании с методом случайного поиска параметров. Далее показано использование модифицированного градиентного метода наискорейшего спуска для решения задачи безусловной многомерной глобальной минимизации (4).

В данном методе целевая функция $J(\Phi)$ минимизируется по направлению, в котором она быстрее всего убывает, т. е. в направлении, обратном градиенту $\text{grad}(J(\Phi))$. Вектор градиента, компонентами которого являются частные производные, с достаточной точностью может быть рассчитан при помощи четырехточечной формулы конечно-разностного отношения

$$\frac{\partial J(\Phi)}{\partial \varphi_i} \approx \frac{1}{12h} \times$$
$$\times \left(J(\varphi_i - 2h) - 8 \cdot J(\varphi_i - h) + 8 \cdot J(\varphi_i + h) - J(\varphi_i + 2h)\right),$$

где h – отклонение от точки, в которой производится расчет значения частной производной.

Вдоль направления антиградиента функция $J(\Phi)$, не являющаяся унимодальной, зависит от одной параметрической переменной t, для нахождения минимума которой можно воспользоваться методом полного перебора параметра в заданном интервале (глубине) поиска. Спуск из точки начального приближения $\Phi_0 = (\varphi_1, ..., \varphi_z)$ против градиента до минимума t определяет новую точку $\Phi_1 = (\varphi_1, ..., \varphi_z)$, в которой вновь определяется градиент и делается следующий спуск, $z = \sum_{i=1}^{n'} m_i$. Окончание процесса сопровождается выполнением одного из условий (6):

$$\max|\Phi_{k+1} - \Phi_k| < \varepsilon; J(\Phi) = 0. \quad (6)$$

Итеративная процедура многомерной оптимизации на основе градиентного метода будет выглядеть следующим образом:

$$\Phi_{k+1} = \Phi_k - t \cdot \text{grad}(J(\Phi)).$$

Задачу (5) также можно решить методами безусловной многомерной оптимизации, если максимизировать значение критерия (3) при выходе $\chi_i \in X$ из области ограничений.

Очередное измерение нового значения критерия $J(e)$ в заключение каждого шага процедуры коррекции термов входных переменных $G(\mu_{a_i}(x_i), i = 1, ..., n')$ происходит по модели $L_{N'}^{(t^*)}$, образованной от текущей $L_{N'}^{(t)}$ за счет удаления статистически маловероятных конъюнкций согласно следующему алгоритму исключения противоречий (АИП):

1) если одна конъюнкция c_j определяет все термы b_s, $s = 1, ..., p$, то она исключается из модели;

2) одинаковые конъюнкции c_j, определяющие один и тот же терм b_s, заменяются одной;

3) конъюнкции группируются в подмножества G_s, $s = 1, ..., p$, $p \le m^o$, где упорядочиваются по убыванию значений информационных мер, вычисляемых по выражению (2);

4) конъюнкции с неудовлетворительными значениями информационных мер, для которых выполняется условие $d(c_{js}) < \widetilde{c}_s$, исключаются из модели. Здесь $\widetilde{c}_s$ – граница значимости подмножества G_s, вычисляемая как значение $d(c_{ks})$

первой по порядку конъюнкции в подмножестве G_s, для которого выполняется условие $d(c_{ks}) >= \sum_{i=k+1}^{n_s} d(c_{is})$, $k = 1, ..., n_s - 1$.

Исходя из этого настроечный этап представляет собой итерационную процедуру изменения параметров ФП термов

$$L_{N'}^{(t+1)} = G\left(L_{N'}^{(t)}, W_N^K(t)\right), \quad t = s = 2, ..., \qquad (7)$$

продолжающуюся до тех пор, пока не будет получена модель с минимальным значением критерия $J(e)$.

На каждой итерации процедуры (7) решаются задачи многомерной оптимизации (4), (5). Количество итераций метода случайного поиска определяет количество итераций t процедуры $G(...)$.

Сходимость процедуры (7) к $L_{N'}^{(t)} \mid \min J(e)$ обеспечивается: 1) ограниченностью шума $\sigma_v^2 < \infty$; 2) подстройкой функций принадлежности процедурой $G(...)$, направленной на безусловную минимизацию функции $J(e)$; 3) выбором оптимального значения s, а также параметров градиентного метода t и ε, обеспечивающих его сходимость.

В результате формируется лингвистическая модель, которая в максимальной степени адекватна моделируемой системе при требуемом уровне селективности.

Рассмотренный метод автоматизации процесса приближенного моделирования в отличие от известных методов обладает достаточной гибкостью и устойчивостью к субъективным факторам за счет автоматического определения наилучших параметров модели в соответствии с начальными условиями.

Лингвистические модели достаточно хорошо себя зарекомендовали [4, 5] и нашли широкое применение во многих сферах человеческой деятельности.

СПИСОК ЛИТЕРАТУРЫ

1. **Токарев В.Л.** Основы теории обеспечения рациональности решений: Монография. Тула: Изд-во ТулГУ, 2000. 120 с.

2. **Загоруйко Н.Г., Ёлкина В.Н., Лбов Г.С.** Алгоритмы обнаружения эмпирических закономерностей. Новосибирск: Наука, 1985.

3. **Ивахненко А.Г., Юрачковский Ю.П.** Моделирование сложных систем по экспериментальным данным. М.: Радио и связь, 1987. 120 с.

4. **Токарев В.Л.** Логико-лингвистические модели в задачах управления сложными объектами // Автоматизация и современные технологии. 1999. № 3. С. 35–39.

5. **Орлов С.В.** Об одном механизме постановки диагноза в экспертной системе медицинской диагностики // Нечеткие системы и мягкие вычисления. 2007. Т. 2, № 2. С. 41–48.

УДК 519.866

П.Н. Победаш

АНАЛИЗ ДВУХКРИТЕРИАЛЬНОЙ МОДЕЛИ ОПТИМАЛЬНОГО УПРАВЛЕНИЯ РЕАЛЬНЫМИ ИНВЕСТИЦИЯМИ С НЕОПРЕДЕЛЕННЫМ СПРОСОМ

Анализируя инвестиционные проекты (ИП), лица, принимающие решение (ЛПР), всегда осуществляет два этапа их оценки – предварительной и детальной проработки проектов. При этом часто важнее оценить и научно обосновать решение, принятое на предварительном этапе, так как именно на этой стадии ошибка в оценке проекта может привести к выбору неэффективного (отказу от эффективного) ИП, что влечет за собой неоправданно высокие затраты. На фоне многочисленных работ по инвестиционному анализу с учетом целей нескольких экономических агентов особенно заметно, что мало публикаций, в которых принимают во внимание неопре-

деленность денежных потоков, инициируемых рассматриваемым проектом, и спроса на производимую продукцию, а также максимальную фондоотдачу используемых основных производственных фондов (ОПФ) и т. п. В предлагаемой статье для управления процессом реального инвестирования (т. е. инвестирования в ОПФ – оборудование, станки и т. п.) на этапе предварительного анализа ИП используется многокритериальный оптимизационный подход, позволяющий оценить наибольшую привлекательность проекта с учетом интересов каждого его участника.

Рассмотрим задачу, обобщающую постановку, принятую в [1], на случай двух критериев, когда учитываются интересы двух лиц – предприятия и государственного органа (налогового центра). Сформулируем ее следующим образом. Предприятие располагает собственным начальным капиталом и предполагает производить продукцию нескольких видов, спрос на которую неизвестен в силу инновационности проекта, отсутствия соответствующей статистики и других причин. При этом известны технико-экономические характеристики ОПФ, участвующих в производстве: стоимость, срок службы, а также производительность единицы ОПФ и стоимость единицы производимой продукции каждого вида. Требуется определить суммы инвестиций, выделяемые инвестором (налоговым центром) и предприятием на реализацию ИП в целом и по каждому виду ОПФ в отдельности, при которых их суммарные дисконтированные денежные потоки, порождаемые данным ИП за определенный период, максимальны. Будем понимать оптимальность инвестирования в смысле Парето.

Предполагаем, что выполнены следующие предпосылки: 1) учитываются налоги, составляющие большую часть затрат предприятия, – на добавленную стоимость (НДС), на прибыль (НП), на имущество (НИ), единый социальный (ЕСН) и отчисления в фонд оплаты труда (ФОТ); 2) на предприятии имеются достаточные запасы сырья; 3) срок T действия ИП меньше сроков T_k службы единицы ОПФ каждого типа: $T < T_k$ $(k = 1, ..., n)$; 4) на ОПФ каждого вида производится лишь один тип продукции. С учетом приведенных предпосылок сформулированная задача имеет вид двухкритериальной многошаговой задачи линейного программирования (МЗЛП), которую согласно [2] назовем моделью B1:

$$x_k(t+1) = x_k(t) + u_k(t)\ (k = 1, ..., n;\ t = 0, ..., T-1),$$

$$x_{n+1}(t+1) = -\sum_{k=1}^{n} x_k(t)/T_k + x_{n+1}(t) + \sum_{k=1}^{n} u_k(t)$$
$$(t = 0, ..., T-1),$$

$$\begin{aligned} x_{n+2}(t+1) &= -\alpha_2 x_{n+1}(t) + x_{n+2}(t) - \\ &- \sum_{k=1}^{n} u_k(t) + u_{2n+1}(t) + u_{2n+2}(t)\ (t = 0), \end{aligned} \qquad (1)$$

$$x_{n+2}(t+1) = \alpha_3 \sum_{k=1}^{n} \frac{x_k(t)}{T_k} - \theta x_{n+1}(t) + x_{n+2}(t) -$$

$$- \sum_{k=1}^{n} u_k(t) + \gamma \sum_{k=1}^{n} u_{n+k}(t)\ \ (t = 1, ..., T-1);$$

$$x_k(0) = 0\ (k = 1, ..., n+2);$$

$$x_{n+2}(t) \ge 0\ (t = 1, ..., T);$$

$$-\sum_{k=1}^{n} \frac{x_k(t)}{T_k} - \alpha_2 x_{n+1}(t) + (1-\beta)\sum_{k=1}^{n} u_{n+k}(t) \ge 0$$
$$(t = 1, ..., T-1);$$

$$u_{n+k}(t) \le \delta_k x_k(t)\ (k = 1, ..., n;\ t = 1, ..., T-1);$$

$$u_{2n+1}(0) \le I_0,\ u_{2n+2}(0) \le K_0,$$

$$u_k(t) \ge 0\ (k = 1, ..., n;\ t = 0, ..., T-1);$$

$$u_{n+k}(t) \ge 0\ (k = 1, ..., n;\ t = 1, ..., T-1);$$

$$u_{2n+1}(0) \ge 0,\ u_{2n+2}(0) \ge 0;$$

$$J = \{J_1, J_2\} \to \max, \qquad (2)$$

где

$$J_1 = -u_{2n+1}(0) - u_{2n+2}(0) +$$

$$+ \sum_{t=1}^{T-1} \frac{\left[\alpha_3 \sum_{k=1}^{n} \frac{x_k(t)}{T_k} - \theta x_{n+1}(t) + \gamma \sum_{k=1}^{n} u_{n+k}(t)\right]}{(1+r)^t} + \frac{\delta x_{n+1}}{(1+r}$$

$$J_2 = \sum_{t=T^2}^{T-1} \frac{\left[-\alpha_3 \sum_{k=1}^{n} \frac{x_k(t)}{T_k} + \theta x_{n+1}(t) + \rho \sum_{k=1}^{n} u_{n+k}\right.}{(1+r)^t}$$

– соответственно дисконтированные суммы собственных средств предприятия и налогового центра; $u_k(t)$ $(t = 0, ..., T-1)$, $u_{n+k}(t)$ $(k = 1, ..., n;\ t = 1, ..., T-1)$, $u_{2n+1}(0)$ и $u_{2n+2}(0)$ – стоимость приобретаемых ОПФ, выручка от реализации продукции k-го типа, внешние и внутренние инвестиции соответственно; $x_k(t)$ $(k = 1, ..., n)$, $x_{n+1}(t)$, $x_{n+2}(t)$ $(t = 0, ..., T)$ – накопленная стоимость

всех ОПФ k-го типа, остаточная стоимость всех ОПФ, текущие денежные средства предприятия в момент t; V_k, T_k, c_k и P_k – производительность, срок службы, стоимость единицы ОПФ и стоимость единицы продукции k-го типа; I_0, K_0 – суммы внешних и внутренних инвестиций, выделяемых на весь срок действия ИП; α_1, α_2, α_3, α_4 – ставки НДС, НИ, НП и ЕСН (НДС включается в цену продукции, поэтому можно считать, что $\alpha_1 = 0$); β – доля выручки от реализации, выделяемая на ФОТ; $\theta = (1-\alpha_3)\alpha_2$, $\delta_k = P_k V_k / c_k$ $(k = 1, ..., n)$, $\gamma = (1-\alpha_3)(1-\beta)$, $\rho = (1-\beta)\alpha_3 + \alpha_4\beta$; r – ставка доходности ИП; δ $(0 \le \delta \le 1)$ – доля остаточной стоимости всех ОПФ на момент $t = T$ от ее балансовой стоимости, определяемая в общем случае экспертно.

Согласно [3] многокритериальная МЗЛП (ММЗЛП) (1), (2) равносильна однокритериальной задаче с условиями (1) и максимизацией свертки критериев $J(\mu) = \mu_1 J_1 + \mu_2 J_2$, где $\mu \in M = \{(\mu_1; \mu_2) \in E^2 \mid \mu_i > 0\ (i = 1, 2);\ \mu_1 + \mu_2 = 1\}$ – вектор параметров, E^2 – двумерное евклидово пространство. Поскольку $\mu_2 = 1 - \mu_1$, то, обозначая $\mu = \mu_1$, перейдем от ММЗЛП (1), (2) к эквивалентной ей однокритериальной задаче (1) при условии:

$$J(\mu) = \mu J_1 + (1-\mu)J_2 \to \max\ (\mu \in (0; 1)). \quad (3)$$

Для задачи (1), (3) имеет место лемма [2].

Лемма. Для оптимальных значений переменных $u^*_{n+k}(t)$ и $x^*_k(t)$ $(k = 1, ..., n;\ t = 1, ..., T-1)$ модели (1), (3) справедливы равенства:

$$u^*_{n+k}(t) = \delta_k x^*_k(t)\ (k = 1, ..., n;\ t = 1, ..., T-1). \quad (4)$$

Здесь и далее звездочкой обозначены оптимальные значения переменных и критериев. Условие (4) означает: в оптимуме выручка от реализации равна стоимости произведенной по каждому виду продукции.

Подставляя (4) в условия задачи (1), (3) и перенумеровывая $u_{2n+1}(0)$ и $u_{2n+2}(0)$ как $u_{n+1}(0)$ и $u_{n+2}(0)$, получаем МЗЛП, эквивалентную указанной:

$$x_k(t+1) = x_k(t) + u_k(t)\ (k = 1, ..., n;\ t = 0, ..., T-1),$$

$$x_{n+1}(t+1) = -\sum_{k=1}^{n} x_k(t)/T_k + x_{n+1}(t) + \sum_{k=1}^{n} u_k(t)$$
$$(t = 0, ..., T-1),$$

$$x_{n+2}(t+1) = x_{n+2}(t) - \sum_{k=1}^{n} u_k(t) + u_{n+1}(t) + u_{n+2}(t)$$
$$(t = 0),$$

$$x_{n+2}(t+1) = \sum_{k=1}^{n} \gamma_k x_k(t) - \theta x_{n+1}(t) + x_{n+2}(t) - \sum_{k=1}^{n} u_k(t)$$
$$(t = 1, ..., T-1);$$

$$x_k(0) = 0\ (k = 1, ..., n+2);$$

$$x_{n+2}(t) \ge 0\ (t = 1, ..., T);$$

$$\sum_{k=1}^{n} \sigma_k x_k(t) - \alpha_2 x_{n+1}(t) \ge 0\ (t = 1, ..., T-1), \quad (5)$$

$$u_{n+1}(t) \le I_0;\ u_{n+2}(t) \le K_0\ (t = 0);$$

$$u_k(t) \ge 0\ (k = 1, ..., n;\ t = 0, ..., T-1),$$

$$u_k(t) \ge 0\ (k = n+1, n+2;\ t = 0);$$

$$\overline{J}(\mu) = -\mu\left[u_{n+1}(0) + u_{n+2}(0)\right] +$$
$$+ \sum_{t=1}^{T-1} \frac{\left[\sum_{k=1}^{n} \rho_k x_k(t) - (2\mu - 1)\theta x_{n+1}(t)\right]}{(1+r)^t} +$$

$$+ \frac{\mu \delta x_{n+1}(T)}{(1+r)^{T-1}} \to \max\ (\mu \in (0;1)),$$

где $\gamma_k = \alpha_3/T_k + \gamma\delta_k$; $\omega_k = -\alpha_3/T_k + \rho\delta_k$; $\sigma_k = (1-\beta)\delta_k - 1/T_k$; $\rho_k = \mu\gamma_k + (1-\mu)\omega_k$ $(k = 1, ..., n)$.

Модель, полученную из задачи (5) исключением второго неравенства – условия неотрицательности прибыли, назовем моделью C. Обозначим для большей содержательной наглядности $NPV^*(\mu) = \mu NPV_1^* + (1-\mu)\mathrm{NPV}_2^*$, $\overline{NPV}(\mu) = \mu\overline{NPV}_1^* + (1-\mu)\overline{NPV}_2^*$ $(\mu \in [0; 1])$ – оптимальные значения сверток целевых критериев в ММЗЛП $B1$ и C, где NPV_i^*, $\overline{NPV}_i^*$ $(i = 1, 2)$ – соответствующие Парето-оптимуму значения целевых функций в указанных задачах. При этом обозначение $NPV^*(\mu)$ (или $\overline{NPV}^*(\mu)$) вполне обосновано с экономической точки зрения, если полагать, что случайная величина NPV (соответственно $\overline{NPV}$) принимает значение NPV_1^* ($\overline{NPV}_1^*$) с вероятностью $\mu \in [0; 1]$, а также значение NPV_2^* ($\overline{NPV}_2^*$) с вероятностью $1 - \mu$. Тогда $NPV^*(\mu)$ (или $\overline{NPV}^*(\mu)$) можно трактовать формально как математическое ожидание, или среднее значение, чистой приведенной стоимости средств всех ЛПР, полученной в результате реализации ИП, формируемого по модели $B1$ (C). Имеет место теорема 1, представленная ниже.

Теорема 1. В задачах $B1$ и C существует оптимальное управление, причем $NPV^*(\mu) \le \overline{NPV}^*(\mu)$ $(\forall \mu \in [0; 1])$.

Отметим, что обозначение $NPV^*(\mu)$ (или $\overline{NPV}^*(\mu)$) вполне обоснованно, ибо содержательно эту величину можно трактовать как среднее значение чистой приведенной стоимости средств всех ЛПР, полученной в результате реализации ИП в соответствии с моделью $B1$ (C).

Поскольку по теореме 1 решение МЗЛП C существует, то, применяя дискретный принцип максимума [4] к указанной задаче, запишем двойственные статические задачи линейного программирования (ЗЛП) в следующем виде:

$$\min_{\lambda_1(t)} H_D(t); \lambda_1(t) \ge p_k^*(t+1) + p_{n+1}^*(t+1) - p_{n+2}^*(t+1)$$
$$(k = 1, ..., n),\ \lambda_1(t) \ge 0\ (t = T-1, ..., 1). \quad (6)$$

Здесь $H_D(t) = \nu(t)\lambda_1(t)$,

$$\nu(t) = \sum_{k=1}^{n} \gamma_k x_k^*(t) - \theta x_{n+1}^*(t) + x_{n+2}^*(t)\ (t = T-1, \quad ;$$

$$\min_{\lambda(t)} H_D(t); \lambda_1(t) \ge p_k^*(t+1) + p_{n+1}^*(t+1) -$$
$$- p^*_{n+2}(t+1)\ (k = 1, ..., n); \quad (7)$$

$$-\lambda_1(t) + \lambda_2(t) \ge p_{n+2}^*(t+1) - 1;\ -\lambda_1(t) +$$
$$+ \lambda_3(t) \ge p_{n+2}^*(t+1) - 1; \lambda_j(t) \ge 0\ (j = 1, ..., 3; t = 0),$$

где $H_D(t) = I_0\lambda_2(t) + K_0\lambda_3(t)$.

При этом двойственные уравнения движения имеют вид:

$$p^*_k(t) = p^*_k(t+1) - p^*_{n+1}(t+1)/T_k + \gamma_k[p^*_{n+2}(t+1) +$$
$$+ \lambda_1^*(t)] + \rho_k(1+r)^{-t}]\ (k = 1, ..., n);$$

$$p^*_{n+1}(t) = p^*_{n+1}(t+1) - \theta[p^*_{n+2}(t+1) + \lambda_1^*(t) +$$
$$+ (2\mu - 1)(1+r)^{-t}];$$

$$p^*_{n+2}(t) = p^*_{n+2}(t+1) + \lambda_1^*(t)\ (t = T-1, ..., 1), \quad (8)$$

где согласно [4] имеет место равенство

$$p_k^*(t) = 0\ (k = 1, ..., n+2, k \ne n+1); p_{n+1}^*(t) =$$
$$= \mu\delta(1+r)^{1-T}\ (t = T). \quad (9)$$

Заметим, что в силу теоремы 1 решение ЗЛП (6) и (7) существует. Найдем значение $\overline{NPV}^*$. Очевидно, в любом случае решением задачи (6) является значение, определяемое по формуле

$$\overline{\lambda}(t) = \max[0; \max_{k=1,...,n}\{p_k^*(t+1) + p_{n+1}^*(t+1) -$$
$$- p_{n+2}^*(t+1)\}]\ (t \in \{T-1, ..., 1\}). \quad (10)$$

Найдем условия, при которых выполняется равенство

$$\lambda_1^*(t) = 0\ (t = T-1, ..., 1), \quad (11)$$

т. е. в силу (10)

$$p_k^*(t+1) + p_{n+1}^*(t+1) - p_{n+2}^*(t+1) \le 0$$
$$(k = 1, ..., n; t = T-1, ..., 1). \quad (12)$$

Полагая в последнем неравенстве $t = T-1$ и учитывая (9), получаем: $\mu\delta(1+r)^{1-T} \le 0$, откуда следует, что $\delta = 0$, так как $\mu > 0$. Тогда условия (9) примут вид:

$$p_k^*(t) = 0\ (k = 1, ..., n+2)\ (t = T). \quad (13)$$

Поскольку общее решение следующего линейного разностного уравнения 1-го порядка $X_t = X_{t+1} + g(t)$ $(t = T-1, ..., 0)$ можно записать в виде формулы

$$X_t = X_T + \sum_{j=t}^{T-1} g(j)\ (t = T, ..., 0),$$

где $X_T = \text{const}$ – заданное конечное значение, то, применяя ее к уравнениям (8) и учитывая (11) и (13), получаем:

$$p_k^*(t) = \frac{[\rho_k + \theta(2\mu-1)/(rT_k)]}{r}\Big[(1+r)^{1-t} - (1+r)^1$$
$$- \frac{\theta(2\mu-1)(T-t)}{rT_k(1+r)^{T-1}}\ (k = 1, ..., n),$$

$$p_{n+1}^*(t) = \frac{\theta(2\mu-1)}{r}\Big[(1+r)^{1-T} - (1+r)^{1-t}\Big]; p_{n+2}^*(t) = 0;$$
$$p_k^*(T) = 0\ (k = 1, ..., n+2),$$
$$\lambda_1^*(t) = 0\ (t = T-1, ..., 1). \quad (14)$$

Подставляя (14) в (12) и учитывая, что (12) при $t = T-1$ истинно в силу (13), найдем:

$$-\frac{[rT_k\rho_k + \theta(2\mu-1)(1+r[T-1-t]-rT_k)]}{r^2T_k(1+r)^{T-1}} +$$

$$+\frac{[rT_k\rho_k + \theta(2\mu-1)(1-rT_k)]}{r^2T_k(1+r)^t} \le 0$$

$$(k = 1, ..., n; t = T-2, ..., 1).$$

Полагая $f_k(t) = \Lambda_k[(1+r)^{-t} - (1+r)^{1-T}] - \theta r(2\mu-1)(T-1-t)(1+r)^{1-T}\ (t = T-2, ..., 1)$, где $\Lambda_k = rT_k(\rho_k - \theta(2\mu-1)) + \theta(2\mu-1)\ (k = 1, ..., n)$, перепишем последнее неравенство в виде

$$f_k(t) \le 0\ (k = 1, ..., n; t = T-2, ..., 1). \quad (15)$$

Очевидно, что условие (15) осмысленно лишь для $T \ge 3$.

Анализируя случаи $\Lambda_k < 0$, $\Lambda_k = 0$ и $\Lambda_k > 0$ ($k \in \{1, ..., n\}$) и исследуя функции $f_k(t)$ на монотонность и экстремум при различных сочетаниях значений параметров $r > 0$; $\theta > 0$; $\mu \in (0; 1)$, получаем теорему.

Теорема 2. Если $r > 0$; $\delta = 0$ и имеют место условия

$$\begin{cases} \Lambda_k \le 0 \ (\mu \ge 1/2); \\ 0 < \Lambda_k \le \dfrac{\theta r(2\mu - 1)(T-2)}{(1+r)^{T-2} - 1} \ (\mu > 1/2) \end{cases}^{(T \ge 3; k = 1, ..., n)} \quad (16)$$

то оптимальные значения управляющих и фазовых переменных МЗЛП, двойственной к задаче C, определяются формулами (14).

Заметим, что при $T = 2$ для выполнения формул (14) достаточно, чтобы $r > 0$; $\delta = 0$. Можно обобщить теорему 2, отказавшись от условия $\delta = 0$ и потребовав выполнения неравенства (12), например, для $t = T - 2, ..., 1$. При этом, полагая в (6) $t = T - 1$ и учитывая (9), получаем ЗЛП вида $\min\limits_{\lambda_1(t)} \nu(t)\lambda_1(t), \lambda_1(t) \ge \mu\delta(1+r)^{1-T}; \lambda_1(t) \ge 0 \ (t = T-1)$, где

$$\nu(t) = \sum_{k=1}^{n} \gamma_k x_k^*(t) - \theta x_{n+1}^*(t) + x_{n+2}^*(t).$$

Несложно показать, что $\lambda_1^*(T-1) = \mu\delta(1+r)^{1-T}$ является решением этой задачи.

Пусть выполняются условия теоремы 2. Подставляя исходные данные задачи C и (14) в выражение целевого критерия $\overline{J}_D$ двойственной к ней МЗЛП и учитывая, что $\overline{J}_D^* = \overline{J}_P = \overline{NPV}^*(\mu)$, получаем оптимальное значение целевого критерия в задаче C: $\overline{NPV}^*(\mu) = I_0\lambda_2^*(0) + K_0\lambda_3^*(0)$. Принимая во внимание (14) при $t = 1$, последнее выражение и решая ЗЛП (7), имеем:

$$\overline{NPV}^*(\mu) = (I_0 + K_0)\eta, \quad (17)$$

где

$$\eta = \max[0; \bar{\chi} - 1], \ \bar{\chi} = \max[0; \max_{k=1, ..., n} \chi_k],$$

$$\chi_k = \frac{1}{r^2 T_k}\left(\Lambda_k - \frac{[\Lambda_k + \theta r(2\mu - 1)(T-1)]}{(1+r)^{T-1}}\right) (k = 1, ..., n). \quad (18)$$

Поскольку для оптимальных значений $NPV^*(\mu)$ и $\overline{NPV}^*(\mu)$ целевых критериев в моделях $B1$ и C по теореме 1 имеет место неравенство $NPV^*(\mu) \le \overline{NPV}^*(\mu)$, то в силу (17) и теоремы 2 получим теорему.

Теорема 3. Если $r > 0$; $\delta = 0$ и справедливы неравенства (16), то $NPV^*(\mu)$ удовлетворяет условию

$$NPV^*(\mu) \le (I_0 + K_0)\eta, \quad (19)$$

где величина η определяется формулами (18).

Заметим, что если 1-е неравенство (16) справедливо для всех $k = 1, ..., n$, то с учетом (18) и соотношений

$$\chi_k = \frac{1}{r^2 T_k}\left(\Lambda_k\left[1 - \frac{1}{(1+r)^{T-1}}\right] - \frac{\theta r(2\mu - 1)(T-}{(1+r)^{T-1}}\right.$$

$$\le 0 \ (k = 1, ..., n)$$

имеем $\bar{\chi} \le 0$, откуда следует, что $\eta = 0$. Поэтому по условию (19) получим: $NPV^*(\mu) \le 0$. С другой стороны, поскольку нулевые управления в модели $B1$ являются допустимыми и им соответствует $NPV(\mu) = 0$, то $NPV^*(\mu) \ge 0$, откуда найдем: $NPV^*(\mu) = 0$. Таким образом, получено следствие 1.

Следствие 1. Если $r > 0$; $\delta = 0$ и справедливо лишь 1-е неравенство (16), то оптимальное значение свертки критериев $J^*(\mu) = NPV^*(\mu)$ ИП, описываемого моделью $B1$, равно нулю: $NPV^*(\mu) = 0$.

Отметим, что из доказательства следствия 1 следует

$$\bar{\chi} = \max[0; \max_{k:\Lambda_k > 0} \chi_k]. \quad (20)$$

Рассматривая частный случай, когда $\eta = 0$, т. е. согласно (18) выполнено условие $\bar{\chi} \le 1$, и учитывая (20), из теоремы 3 получаем следствие 2.

Следствие 2. Если $r > 0$; $\delta = 0$, выполняется 2-е условие (16) и

$$\max_{k:\Lambda_k > 0}\left\{\frac{1}{T_k}\left(\Lambda_k\left[1 - \frac{1}{(1+r)^{T-1}}\right] - \frac{\theta r(2\mu - 1)(T-1)}{(1+r)^{T-1}}\right)\right.,$$

то $NPV^*(\mu) = 0$.

В свою очередь из следствия 2 получим еще одно следствие.

Следствие 3. Если $r > 0$; $\delta = 0$, выполнены 2-е неравенство (16) и

$$\max_{k:\Lambda_k > 0}\left(\frac{\Lambda_k}{T_k}\right) \le \frac{r\left[r\bar{T}(1+r)^{T-1} + \theta(2\mu - 1)(T-1)\right.}{\bar{T}\left[(1+r)^{T-1} - 1\right]},$$

где $\bar{T} = \max\limits_{k:\Lambda_k > 0} T_k$, то $\mathrm{NPV}^*(\mu) = 0$.

Таким образом, следствия 1–3 задают условия, когда ИП, описываемый МЗЛП (1), (2), невыгоден для инвесторов даже при его оптимальной реализации, а условие (19) – верхнюю

границу для $NPV^*(\mu)$, и могут использоваться для предварительного анализа проекта. Из теоремы 3 следует, что при $r > 0$; $\delta = 0$ и условиях (16) имеет место неравенство $0 \le NPV^*(\mu) \le \Gamma(r)$, где $\Gamma(r)$ – оценка на $NPV^*(\mu)$ сверху в правой части (19). Заметим, что, поскольку $\lim_{r\to+\infty} \Gamma(r) = +0$, то из последнего неравенства имеем: $\lim_{r\to+\infty} NPV^*(\mu) = +0$, т. е. оценка эффективности ИП становится все более определенной с увеличением ставки дисконтирования. Это вполне согласуется со следующим подходом к анализу инвестиционной привлекательности в условиях неопределенности и риска. В соответствии с работами [5–7 и др.] эффективность ИП в условиях риска (инфляции) оценивается при повышении ставки его доходности, т. е. в соответствии с формулой $r_i \approx r + i$, где r_i, r, i – коэффициенты дисконтирования соответственно с учетом инфляции и без нее и индекс инфляции i. Тогда при неограниченном росте инфляции эффективность проекта даже при наилучшей его реализации в рамках модели BI будет сколь угодно близка к нулю, т. е. проект находится на границе окупаемости.

Заметим, что полученные результаты подтверждены численно [8] и обобщают результаты, изложенные в работе [1], на случай, когда количество экономических агентов (и соответственно целевых критериев) равно двум.

Работа выполнена при финансовой поддержке АВЦП "Развитие научного потенциала высшей школы" (НИР 2.1.1/2710) и ФЦП "Научные и научно-педагогические кадры инновационной России» (НИР НК-136П/3).

СПИСОК ЛИТЕРАТУРЫ

1. **Медведев А.В., Победаш П.Н.** Параметрический анализ модели реальных инвестиций без ограничений на спрос с помощью дискретного принципа максимума // Вестник университетского комплекса. Вып. 4 (18) / НИИ СУВПТ. Красноярск, 2005. С. 186–196.

2. **Медведев А.В.** Применение z-преобразования к исследованию многокритериальных линейных моделей регионального экономического развития: Монография / Сиб. гос. аэрокосмич. ун-т. Красноярск, 2008. 228 с.

3. **Подиновский В.В., Ногин В.Д.** Парето-оптимальные решения многокритериальных задач. М.: Наука, 1982. 256 с.

4. **Победаш П.Н.** Решение оптимизационных задач реального инвестирования с переменным чис"ом управлений // Реформирование системы управления на современном предприятии: Сб. матер. V Междунар. науч.-практ. конф. / РИО ПГСХА. Пенза, 2005. С. 190–192.

5. **Куракина Ю.Г.** Оценка фактора риска в инвестиционных расчетах // Бухгалтерский учет. 1995. № 6. С. 22–27.

6. **Лукасевич И.Я.** Методы анализа рисков инвестиционных проектов // Финансы. 1998. № 9. С. 59–62.

7. **Глазунов В.Н.** Финансы фирмы. М.: Экономика, 2000. 246 с.

8. Конструктор и решатель дискретных задач оптимального управления (КАРМА): Программа для ЭВМ. Свид. о регистр. в Роспатенте 2008614387 от 11.09.2008 / Медведев А.В., Победаш П.Н., Смольянинов А.В., Горбунов М.А.

УДК 519.86; 598.2; 591.47; 681.142; 636.01

Н.А. Попцов, Д.Б. Бекурин, М.А. Марценюк

НЕЙРОСЕТЕВОЕ УПРАВЛЕНИЕ ЛОКОМОЦИЯМИ САМООБУЧАЮЩЕЙСЯ МОДЕЛИ БИООБЪЕКТА

В статье рассматривается модель амебы, состоящая из двух эллипсоидов, связанных между собой тонким каналом. Объект может в определенных пределах изменять длину канала, а также объем и форму эллипсоидов. Даже столь упрощенная модель амёбы имеет семь степеней свободы (шесть вращательных и поступательную) и 13 управляемых параметров. Целью является создание компьютерной модели рассмотренного выше биообъекта в приближении Стокса и

поиск методик самообучения плаванию. В качестве системы, управляющей локомоциями, использована нейросеть. Задача решается в следующей постановке: объект должен достигнуть цель – точку, имеющую заданные координаты, управляя локомоторными степенями свободы и сравнивая свое текущее положение с целевым.

Существуют разные методы достижения объектом цели. Первый метод подобен описанному в книге [3] эвристическому поиску. Этот метод (далее называемый стохастическим) мы реализовали с помощью нейросети с обратной связью. Суть его заключается в переборе доступных движений и анализе расстояния до цели на каждом шаге. На вход нейросети подаются текущие параметры объекта: три радиуса первого эллипсоида, три радиуса второго эллипсоида, три угла Эйлера связи относительно первого узла, три угла Эйлера второго узла относительно связи, расстояние между узлами. На выходе нейросети появляются новые параметры объекта. Если текущее движение было неудачным, то производится возврат в исходное состояние.

Объясним причину возврата после неудачного движения. В n-мерном пространстве степеней свободы существует n-мерный конус, внутри которого движения приближают объект к цели. Так как конус в большинстве случаев остроугольный, то чем больше степеней свободы, тем меньше вероятность того, что следующее случайно выбранное движение приблизит объект к цели. Следовательно, вероятность того, что объект в итоге приблизится к цели, станет ещё меньше, поэтому приходится возвращаться на шаг назад после неудачного движения. После обратного движения объект не возвращается в исходную точку из-за погрешности вычислений, которой можно было бы избежать. Мы не исправляем погрешность в связи с возможностью попадания объекта в вырожденное состояние, при котором любое движение приводит лишь к увеличению расстояния до цели.

Во время движения объекта его суммарный объем поддерживается постоянным, как будто вещество из одного узла перетекает в другой. Программа, в которой реализована эта методика, написана на языке программирования Java, главное окно программы изображено на рис. 1. В верхней части окна задаются координаты цели, отображаются координаты амебы и расстояние до цели, производится запуск и остановка

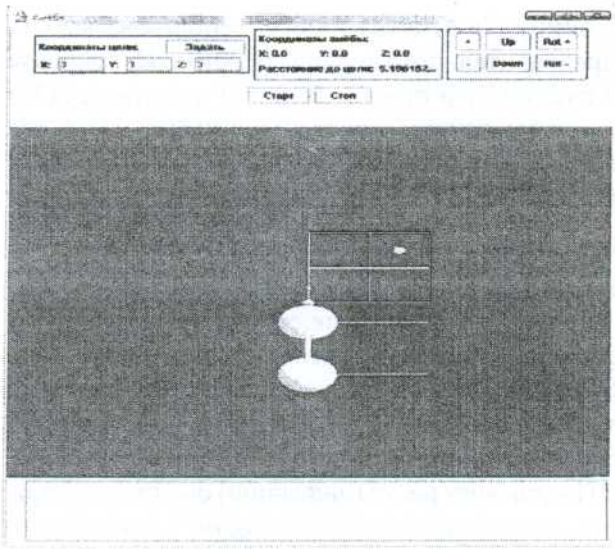


Рис. 1. Главное окно программы

движения амёбы, меняются параметры графического поля (приближение/удаление, вращение вокруг осей). В средней части изображено графическое поле, где с использованием OpenJL прорисовывается все происходящее. В нижнюю часть окна выводятся текущие параметры (номер шага, координаты амёбы).

Рассмотренный стохастический метод не всегда эффективен. Поэтому был также использован описанный в статье [2] метод выявления законов локомоторных движений. Он заключается в построении математической модели существующего локомоторного цикла. В этой статье рассмотрены кинематические и координационные синергии шагательных движений животных и человека. Мы же во втором методе будем искать локомоции для нашего объекта. На рис. 2 показано, как с помощью циклического изменения внутренних параметров можно продвинуться вдоль линии связи. Статья [1] иллюстрирует, как при совершении замкнутого локомоторного цикла наш объект способен также совершать поворот.

Во втором методе применяется четыре нейронных сети для управления амёбой. Каждую нейросеть мы используем для самообучения локомоторным циклам: первую – для самообучения циклическому изменению внутренних параметров, в процессе которого объект продвигается поступательно вперед или назад; вторую, третью и четвертую нейросети – для обучения повороту вокруг координатных осей.

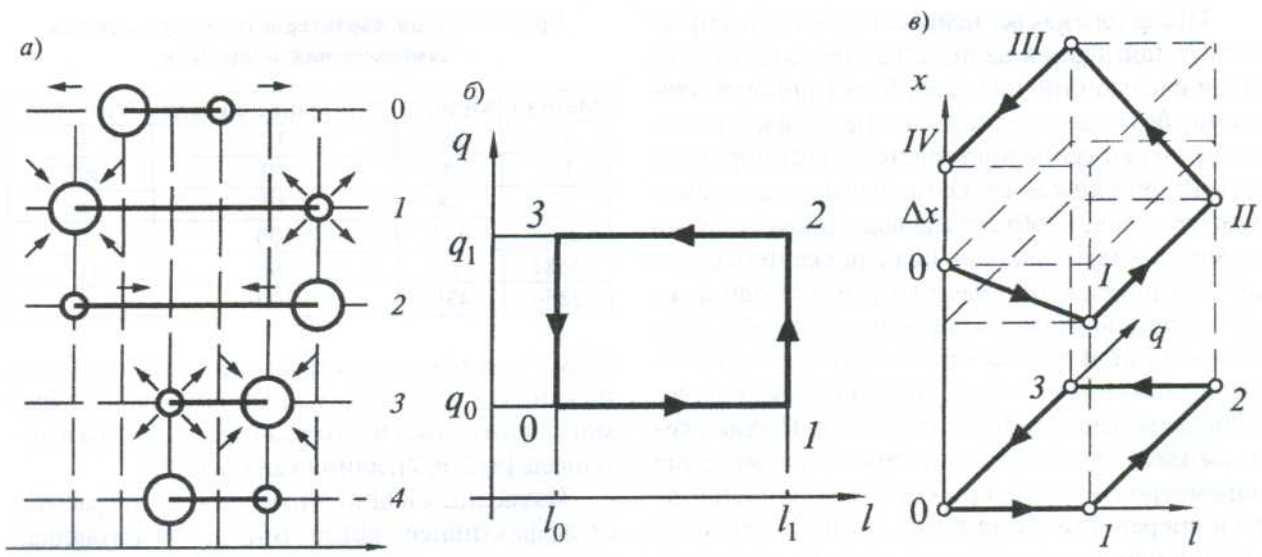


Рис. 2. Плавательное движение модели амебы: *а* – форма и положение плывущего тела, соответствующее точкам *1–4*; *б* – в пространстве форм; *в* – конфигурационное пространство и его проекция на пространство форм

Окно программы, в котором запускается самообучение нейросетей и отображаются результаты, изображено на рис. 3. В средней части окна помещена таблица, из которой можно узнать, какие внутренние параметры и в какой последовательности нужно менять, чтобы продвинуться в нужном направлении или повернуться вокруг заданной оси. В первой колонке ("Тип") изображен тип движения – движение вдоль канала или поворот вдоль осей, во второй колонке стоит номер шага, причем за один шаг можно одновременно менять радиусы обоих узлов или вращать канал относительно первого узла и вращать второй узел относительно канала. В третьей колонке изображено действие над узлом, имя которого находится в колонке "Узел". В колонках "Q1", "Q2", "Q3" помещены величины изменения внутренних параметров (если изменяется длина канала, то величина изменения находится в колонке "Q1"; при изменении радиуса первый радиус изменяется на величину "Q1", второй – на "Q2", третий – на "Q3"; изменения углов "Q1", "Q2", "Q3" соответствуют соответствующим изменениям углов Эйлера). Число шагов в последовательности равно четырем, причем на одном шаге можно менять только или длину связи, или радиусы, или углы.

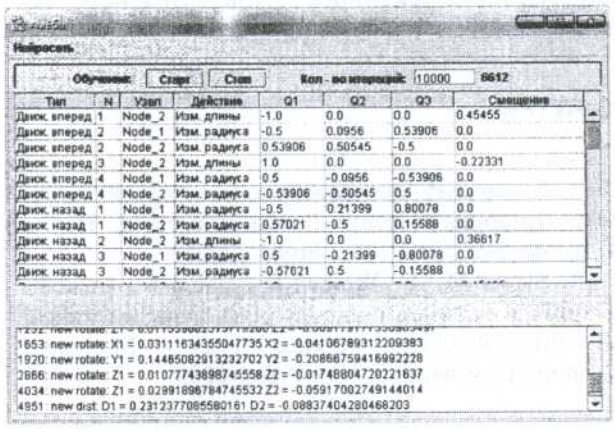


Рис. 3. Окно с результатами обучения

Рассмотрим процесс самообучения поступательному движению вдоль связи вперед или назад. Нейронная сеть представляет собой многослойный персептрон. Число нейронов во входном слое равно 14, в скрытом слое – 100, на выходе – 7. На вход нейронной сети подаются текущие параметры объекта, направление движения (число "+1", если вперёд; "–1" – если назад), номер шага. На выходе нейросети снимаются "новые" параметры объекта. Затем производится изменение параметров, и если результат удачный, то нейросеть запоминает текущие параметры, иначе – нейронная сеть возвращается к предыдущему состоянию. Затем снова происходит малое изменение параметров нейросети случайным образом и проверяется результат. И так далее, пока не достигнем заданного числа итераций. В среднем достаточно 2500 итераций для обучения нейронной сети.

После того как все нейросети обучены, объект может двигаться к цели. Алгоритм достижения цели следующий: сначала объект пробует движение без поворота; если оно приближает его к цели, то движение повторяется до тех пор, пока объект приближается. Если предыдущее движение неудачно, то объект сначала пробует первый поворот, затем движется. Если движение с поворотом приближает объект к цели, то подбирает угол путем повторения поворота до тех пор, пока объект приближается к цели. Затем, если поворот с движением не приближает объект к цели, выбираем следующий поворот и т. д. Под движением здесь понимается циклическое изменение параметров объекта, при котором он продвигается вперед или назад вдоль связи, а его внутренние параметры в начале и конце цикла совпадают. То же самое относится и к повороту, причем движение производится вперёд и назад, а поворот – по или против часовой стрелки вокруг оси.

Таким образом, объект не может измерить угол наклона канала между узлами и цели, а "знает" только о расстоянии до цели. Подобный подход был применен в [4]. Определить расстояние до цели можно, например, по концентрации вещества, из которого состоит цель (так поступает обычная амёба). Форма управления объектом совпадает с формой, описанной в первом методе. Координаты цели можно менять во время движения объекта, таким образом объект будет изменять направление своего движения.

Результат сравнения эффективности двух методов показан в таблице. В первой и второй колонках указано количество движений, необходимое для достижения цели, в последней – угол между каналом объекта и линией, связывающей цель с объектом. Начальные параметры объекта:

Сравнительная характеристика двух методов самообучения нейросети

Метод 1	Метод 2	Расстояние до цели	Угол, рад
334	104	10	0
175	289	10	π/2
323	38	10	π
2663	441	100	0
4687	4185	100	π/2
2856	430	100	π

узлы в виде сфер диаметром в одну единицу, длина канала – 4 ед., радиусы эллипсоидов (узлов) могут изменяться в интервале [0,5; 2], углы в интервале [π/2; π/2], длина канала [3,5].

Из таблицы видно, что второй метод работает эффективнее, когда линия, соединяющая объект и цель, параллельна каналу. Первый же метод эффективен только тогда, когда расстояние до цели небольшое и линия, соединяющая объект и цель, перпендикулярна каналу.

В результате данного исследования можно сделать следующие выводы:

стохастический метод эффективен только при малых расстояниях до цели и когда линия, соединяющая объект и цель, перпендикулярна каналу;

локомоторный метод зависит от результата обучения вспомогательных нейронных сетей и наиболее эффективен, когда линия, соединяющая объект и цель, параллельна каналу.

Если во втором методе изменить алгоритм обучения нейросетей (увеличить шаг продвижения вперёд и назад) и разработать нейросетевой алгоритм поворота объекта к цели (чтобы линия, соединяющая объект и цель, была параллельна каналу), то эффективность локомоторного метода повысится в несколько раз.

СПИСОК ЛИТЕРАТУРЫ

1. **Кислухин Н.М., Марценюк М.А., Бекурин Д.Б.** Компьютерная модель плавания самодеформирующегося тела в вязкой жидкости. Пермь, 2004.

2. **Смолянинов В.В.** Пространственно-временные задачи локомоторного управления // Успехи физических наук. 2000. Т. 170, № 10. С. 1063–1127.

3. **Люгер Д.Р.** Искусственный интеллект: стратегии и методы решения проблем. 4-е изд.: Пер. с англ. М.: Изд. дом "Вильямс", 2003.

4. **Svinin M.M., Yamada K., Ushio S., Ueda K.** A minimal simulation model of evolving periodical gaits for multi-legged walking robots. // Intelligent autonomous systems 6. IOS Press, 2000.

Вычислительные машины и программное обеспечение

УДК 004.451.2

М.Ю. Кудрин, Е.В. Соколов, А.Г. Тормасов

ВЫЯВЛЕНИЕ СОСТОЯНИЙ ГОНКИ С ПОМОЩЬЮ ГРАФА СОВМЕСТНОГО ИСПОЛНЕНИЯ ПОТОКОВ

Состояние гонки (race condition) – это ситуация, возникающая, когда несколько потоков одновременно обращаются к одному и тому же ресурсу, причем хотя бы один из потоков выполняет операцию записи, и порядок этих обращений точно не определен. Как правило, ошибки, вызванные состояниями гонки, трудно обнаружить и исправить.

Обычный способ борьбы с такими ситуациями – поиск разработчиком участков кода в программном продукте, различный порядок исполнения которых может привести к возникновению состояний гонки. Другим способом является внедрение в программный продукт дополнительной функциональности, которая на этапе тестирования позволяет определить, возникло ли состояние гонки в конкретном варианте исполнения программы [1]. К сожалению, оба способа не дают права утверждать, что возникновение состояний гонки в программе невозможно при любом варианте ее исполнения.

Существуют строгие математические доказательства отсутствия состояния гонки для определенных алгоритмов. Правда, чаще всего рассуждения не носят универсального характера и могут быть применимы только к рассматриваемым алгоритмам (см., например, доказательства для различных алгоритмов в [2]). При анализе нового алгоритма доказательство приходится строить с нуля.

Есть еще одна проблема – помимо состояний гонки, которые нарушают правильную работу программы, существуют те, которые не влияют на корректное решение поставленной задачи (будем в дальнейшем называть их разрешенными).

Требуется некоторый подход, который позволил бы точно сказать, существуют ли в программе состояния гонки, приводящие к её некорректной работе (в дальнейшем будем называть их неразрешенными состояниями гонки). Кроме того, удобно, если бы этот подход давал возможность быстро определить, какие именно варианты исполнения программы ошибочны. В данной статье такой подход предложен для программ, удовлетворяющих ряду условий.

Постановка задачи

Как было указано выше, критерием неразрешенности состояния гонки является его влияние на корректную работу программы. Значит, нам необходим некоторый способ определения корректности исполнения программы. При этом очевидно, что не существует абсолютного понятия корректности работы произвольной программы. Мы можем только проверить, что конкретная программа правильно решает определенную задачу. Иногда существует способ сделать это, анализируя текущее состояние общих ресурсов для потоков исполнения и, возможно, ранее считанные состояния, которые хранят потоки. Будем рассматривать именно этот случай в следующей формулировке:

существует два потока, каждый из которых исполняет заданный набор команд;

одна или несколько общих ячеек памяти;

некоторые из операций, выполняемых потоками, затрагивают общие ячейки памяти.

Требуется определить, возникает ли неразрешенное состояние гонки исходя из значений ячеек памяти после завершения работы потоков и, возможно, ранее считанных значений, которые хранят потоки.

Очевидно, что задачу можно решить, перебрав все возможные варианты совместного

исполнения двух потоков. Однако на практике этот метод неприменим из-за сложности перебора. Нетрудно показать, что если каждый из потоков совершает n операций, то нужно выполнить $O(C_{2n}^{n}) = O\left(\frac{2^{2n}}{\sqrt{n}}\right)$ действий.

Мы рассмотрим подход, который позволяет существенно уменьшить число необходимых действий, гарантируя при этом точное определение наличия либо отсутствия состояний гонки в программе.

Представление работы потоков

Каждый из потоков выполняет некоторую последовательность операций. Исходя из определения состояния гонки логично предположить, что нас будут интересовать операции, так или иначе связанные с разделяемыми ресурсами – в нашем случае общими ячейками памяти. В зависимости от отношения к общим ячейкам разобьем операции на четыре группы:

- Операции чтения. Поток считывает значение одной из общих ячеек памяти. Обозначение одной операции – R.
- Операции записи. Поток записывает значение в одну из общих ячеек памяти. Обозначение одной операции – W.
- Операции хранения и модификации ранее считанного значения. Поток хранит либо выполняет некоторые действия (например, арифметические) над ранее считанным значением одной из общих ячеек памяти. Обозначение одной такой операции – V.
- Другие операции. Операции, не входящие ни в один из вышеописанных классов. Обозначение одной операции – X.

В дальнейшем при обозначении операции будем указывать два индекса: верхний соответствует номеру общей ячейки памяти, над которой производится операция, нижний – номеру потока, выполняющего операцию. Например, R_1^3 означает, что первый поток считывает значение третьей ячейки. Для операций, которые мы обозначили через X, верхний индекс не имеет смысла, поэтому его можно опускать.

Введем понятие состояния исполнения двух потоков относительно ячейки памяти. Так будем называть вектор, содержащий три компоненты (a, b, c), где a – значение ячейки памяти в данный момент времени; b – считанное значение ячейки, которым оперирует первый поток; c – считанное значение ячейки, которым оперирует второй поток. Компоненты этого вектора изменяются при чтении и записи, а также когда один из потоков модифицирует считанное значение. Будем считать, что при операции записи в ячейку памяти переносится именно это модифицированное значение. До запуска потоков данный вектор может быть записан в виде $(m, -, -)$, где m – начальное значение ячейки памяти, а прочерки означают, что ни один из потоков еще не считывал значение этой ячейки.

Запишем правила, по которым изменяется состояние исполнения потоков относительно ячейки памяти с номером i при выполнении операций:

$R_1^j(a, b, c) \to (a, a, c)$, если $i = j$;
$R_2^j(a, b, c) \to (a, b, a)$, если $i = j$;
$W_1^j(a, b, c) \to (b, b, c)$, если $i = j$;
$W_2^j(a, b, c) \to (c, b, c)$, если $i = j$;
$V_1^j(f)(a, b, c) \to (a, f(b), c)$, если $i = j$;
$V_2^j(f)(a, b, c) \to (a, b, f(c))$, если $i = j$;
$X_k^j(a, b, c) \to (a, b, c)$, если $i = j$, $k = 0,1$;
$O_k^j(a, b, c) \to (a, b, c)$, если $i \neq j$, $k = 0,1$, O – любая из операций R, W, V, X.

Предположим, что первый поток совершает некоторую операцию A_1^j, а второй поток – операцию B_2^k. Возможны два варианта: операция A_1^j будет выполнена раньше операции B_2^k и операция A_1^j будет выполнена позже операции B_2^k. Возникает вопрос, в каких случаях состояние исполнения потоков после выполнения операций может быть различным.

Лемма 1. Состояние исполнения потоков относительно ячейки памяти с номером i после выполнения двух операций может зависеть от их порядка только том в случае, если $j = k = i$ и выполнено одно из следующих условий:

A = R, B = W (первый поток считывает значение ячейки памяти, а второй пишет туда же);

A = W, B = R (второй поток считывает значение ячейки памяти, первый поток записывает в эту же ячейку);

A = W, B = W (оба потока выполняют операцию записи в одну ячейку).

Доказательство. Если $j \neq i$, то $A_1^j(a, b, c) \to$ $\to (a, b, c)$, а значит, состояние исполнения потоков после выполнения A_1^j и B_2^k будет одинаковым вне зависимости от порядка операций. Аналогично рассматривается случай для $k \neq i$. Итак, $j = k = i$.

Рассмотрев все варианты операций A и B, нетрудно убедиться, что только в трех случаях итоговое состояние исполнения может быть различным. Запишем его в неопределенных коэффициентах (α, β и γ), принимающих значения 0 и 1. Подстановка нуля должна давать итоговое состояние при одном порядке операций, подстановка единицы – при другом.

A = R, B = W:

$R_1^i(a, b, c) \to (a, a, c)$, $W_2^i(a, a, c) \to (c, a, c)$;

$W_2^i(a, b, c) \to (c, b, c)$, $R_1^i(c, b, c) \to (c, c, c)$.

Итоговое состояние исполнения – $(c, \alpha a + \overline{\alpha} c, c)$.

A = W, B = R:

$W_1^i(a, b, c) \to (b, b, c)$, $R_2^i(b, b, c) \to (b, b, c)$;

$R_2^i(a, b, c) \to (a, b, a)$, $W_1^i(a, b, a) \to (b, b, a)$.

Итоговое состояние исполнения – $(b, b, \beta a + \overline{\beta} b)$.

A = W, B = W:

$W_1^i(a, b, c) \to (b, b, c)$, $W_2^i(b, b, c) \to (c, b, c)$;

$W_2^i(a, b, c) \to (c, b, c)$, $W_1^i(c, b, c) \to (b, b, c)$.

Итоговое состояние исполнения – $(\gamma b + \overline{\gamma} c, b, c)$.

Определение 1. Если существует хотя бы одна общая ячейка памяти, такая, что состояние исполнения потоков относительно нее после выполнения двух операций может зависеть от их порядка, то будем называть эти операции некоммутирующими, в противном случае скажем, что они коммутируют между собой.

Построение графа совместного исполнения потоков

Пусть первый поток совершает k операций, а второй – n операций. Под графом совместного исполнения потоков будем понимать ориентированный граф

$$G := (V, A),\ V = \bigcup_{\substack{i=1,k+1 \\ j=1,n+1}} v_j^i,$$

$$A = \bigcup_{\substack{i=1,k \\ j=1,n+1}} (v_j^i, v_j^{i+1}) \cup \bigcup_{\substack{i=1,k+1 \\ j=1,n}} (v_j^i, v_{j+1}^i).$$

Пример подобного графа представлен на рис. 1.

В таком графе существует только одна вершина v_1^1, в которую не входит ни одной дуги. Будем называть её начальной вершиной. Вершину v_{n+1}^{k+1}, единственную, из которой не исходит ни одной дуги, назовём конечной. Для произвольной вершины v_j^i назовем ее уровнем число, равное $i + j$. Начальная вершина имеет минимальный уровень в графе, конечная вершина – максимальный. Кроме того, нетрудно заметить, что для любых двух смежных вершин в данном графе

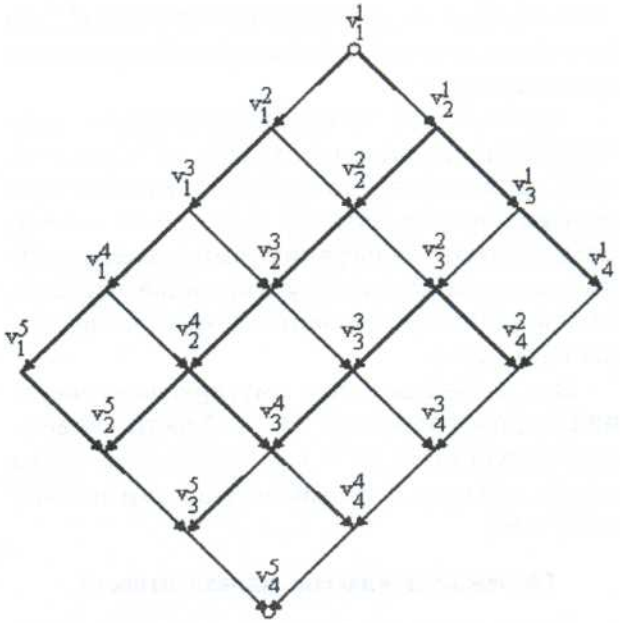


Рис. 1. Пример графа совместного исполнения потоков при $k = 4$, $n = 3$

уровни различаются на единицу, причем меньше уровень той вершины, откуда исходит дуга. Отсюда, в частности, следует, что в таком графе нет циклов. В дальнейшем, рассматривая графы совместного исполнения потоков, будем опускать изображения стрелок на дугах. Интерес к подобному графу обусловлен следующим утверждением.

Лемма 2. Назовем ориентированный путь из начальной вершины в конечную полным путем. Существует взаимно-однозначное соответствие между вариантами совместного исполнения потоков и полными путями в соответствующем графе совместного исполнения потоков.

Доказательство. Рассмотрим некоторый вариант совместного исполнения потоков. Будем строить соответствующий полный путь из вершины v_1^1 следующим образом: на m-м шаге для вершины v_j^i в путь включается дуга (v_j^i, v_j^{i+1}), если m-я операция в варианте исполнения принадлежит первому потоку, и дуга (v_j^i, v_{j+1}^i), если m-я операция в варианте исполнения принадлежит второму потоку. Так как первый поток всего выполняет k операций, а второй – n, то последней вершиной будет v_{n+1}^{k+1}. Аналогично, если у нас есть некоторый полный путь, то в соответствующем варианте совместного исполнения m-я операция будет принадлежать первому потоку, если m-я дуга имеет вид (v_j^i, v_j^{i+1}), и второму, если m-я дуга

вида (v_j^i, v_{j+1}^i). А раз последняя вершина v_{n+1}^{k+1}, то всего будет рассмотрено k операций первого потока и n второго.

Определение 2. Исходя из данного построения будем говорить, что дуги (v_j^i, v_j^{i+1}), где $j = 1, n+1$, представляют i-ю операцию, выполняемую первым потоком. Дуги (v_j^i, v_{j+1}^i), где $i = 1, k+1$, представляют j-ю операцию, выполняемую вторым потоком. Полный путь представляет соответствующий ему вариант совместного исполнения потоков.

Для любых двух некоммутирующих операций i и j поставим знак "+" в области, ограниченной дугами (v_j^i, v_j^{i+1}), $(v_j^{i+1}, v_{j+1}^{i+1})$, $(v_{j+1}^i, v_{j+1}^{i+1})$ и (v_j^i, v_{j+1}^i). На рис. 2 такими операциями являются R_1^1 и W_2^1.

Определение классов эквивалентности

Рассмотрим два различных пути на графе совместного исполнения потоков из вершины v_y^x в конечную вершину v_{n+1}^{k+1}. Пусть (a, b, c) – состояние исполнения потоков относительно некоторой ячейки памяти. Проделаем следующее: будем двигаться по дугам графа вдоль каждого из путей и преобразовывать состояние исполнения потоков согласно операциям, которые эти дуги представляют. Если, дойдя до вершины v_{n+1}^{k+1}, мы получим одинаковые состояния, и это верно при рассмотрении любой из общих ячеек памяти, то будем считать эти пути эквивалентными. Таким образом, все пути из v_y^x в v_{n+1}^{k+1} разбиваются на классы эквивалентности. Очевидно следующее утверждение.

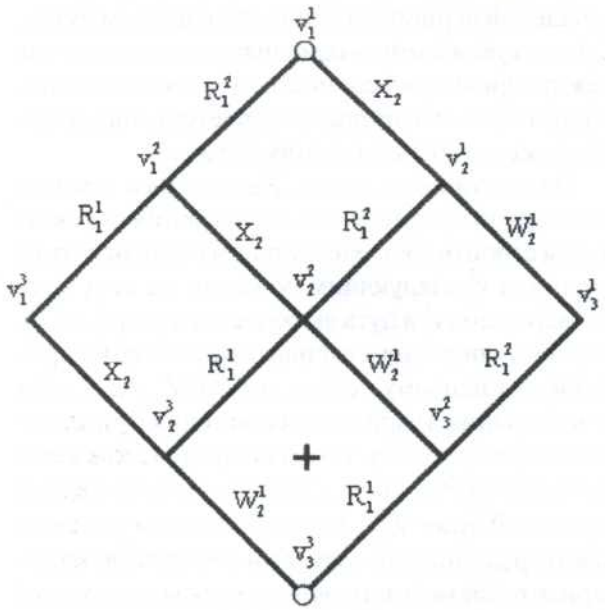


Рис. 2. Пример указания некоммутирующих операций на графе

***Утверждение* 1.** Два пути из вершины v_y^x в вершину v_{n+1}^{k+1}, которые отличаются друг от друга только парой дуг – (v_j^i, v_j^{i+1}), $(v_j^{i+1}, v_{j+1}^{i+1})$ в одном пути и (v_j^i, v_{j+1}^i), $(v_{j+1}^i, v_{j+1}^{i+1})$ в другом, принадлежат одному классу эквивалентности, если i-я операция первого потока и j-я операция второго коммутируют между собой (в области, ограниченной этими дугами (v_j^i, v_j^{i+1}), $(v_j^{i+1}, v_{j+1}^{i+1})$, $(v_{j+1}^i, v_{j+1}^{i+1})$, (v_j^i, v_{j+1}^i), на графе совместного исполнения потоков не указан знак "+").

Нас прежде всего интересуют классы эквивалентности полных путей.

Построение представителей классов эквивалентности

Пусть задан некоторый граф совместного исполнения потоков. Определим для него функцию $f(i,j)$, $1 \le i \le k+1, 1 \le j \le n+1$. Здесь k – число операций, выполняемых первым потоком, а n – число операций, выполняемых вторым потоком.

1) $f(i,j) = 1$ при $i = k+1, j = 1, ..., n+1$;

2) $f(i,j) = 1$ при $j = n+1, i = 1, ..., k$;

3) $f(i,j) = f(i+1,j) + f(i,j+1)$, если i-я операция первого потока и j-я операция второго – некоммутирующие, в противном случае $f(i,j) = f(i+1,j) + f(i,j+1) - f(i+1,j+1)$.

Теорема 1. Число классов эквивалентности полных путей равно $f(1,1)$ и может быть вычислено за $O(k \cdot n)$ операций.

Доказательство. Покажем, что функция $f(i,j)$ равна числу классов эквивалентности путей из вершины v_j^i в вершину v_{n+1}^{k+1}. Утверждение очевидно для п. 1 и 2, так как из таких вершин существует только один путь в вершину v_{n+1}^{k+1}.

Рассмотрим п. 3 для коммутирующих операций. Пусть $f(i+1,j)$, $f(i,j+1)$, $f(i+1,j+1)$ известны и равны числу классов эквивалентности для вершин v_j^{i+1}, v_{j+1}^i, v_{j+1}^{i+1}. Будем пользоваться следующим очевидным утверждением: если два пути из вершины v_{j+1}^{i+1} в вершину v_{n+1}^{k+1} эквивалентны, то и пути, полученные присоединением дуги $(v_j^{i+1}, v_{j+1}^{i+1})$ к данным путям, также будут эквивалентными. Аналогично для дуги $(v_{j+1}^i, v_{j+1}^{i+1})$. И наоборот, присоединение таких дуг к неэквивалентным путям порождает неэквивалентные

пути. Это может быть неверно в некоторых частных случаях, но результатом будет лишь увеличение числа классов эквивалентности для анализа.

Представим число классов эквивалентности путей из вершины v_j^{i+1} в виде $f(i+1,j) = a_1 + b_1$ (рис. 3). Здесь a_1 – число классов эквивалентности, в которых все пути проходят через вершину v_j^{i+2} (равно нулю, если такой вершины не существует или таких классов нет); b_1 – число классов эквивалентности, в которых хотя бы один путь проходит через вершину v_{j+1}^{i+1}. Но тогда b_1 должно равно $f(i+1, j+1)$ согласно утверждению о присоединении дуги, а значит $a_1 = f(i+1, j) - f(i+1, j+1)$. Аналогично для вершины v_{j+1}^{i} число классов эквивалентности, в которых все пути проходят через вершину v_{j+2}^{i}, будет равно $f(i, j+1) - f(i+1, j+1)$. Осталось заметить, что число классов эквивалентности для вершины v_j^i, в которых хотя бы один путь проходит через v_{j+1}^{i+1}, будет равно $f(i+1, j+1)$, так как i-я операция первого потока и j-я операция второго коммутируют между собой.

А значит, $f(i,j) = f(i+1,j) - f(i+1,j+1) + f(i,j+1) - f(i+1,j+1) + f(i+1,j+1) = f(i+1,j) + f(i,j+1) - f(i+1,j+1)$.

Случай некоммутирующих операций рассматривается аналогично за исключением того, что число классов эквивалентности для вершины v_j^i, в которых хотя бы один путь проходит через v_{j+1}^{i+1}, будет равно $2f(i+1, j+1)$, так как при-

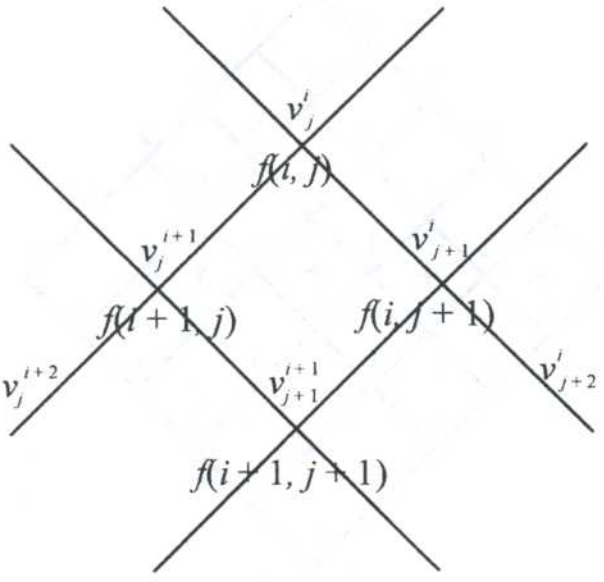


Рис. 3. Коммутирующие операции

соединение ребер (v_j^i, v_j^{i+1}), $(v_j^{i+1}, v_{j+1}^{i+1})$ и (v_j^i, v_{j+1}^i), $(v_{j+1}^i, v_{j+1}^{i+1})$ теперь порождает неэквивалентные пути. Значит, в этом случае $f(i,j) = f(i+1,j) - f(i+1, j+1) + f(i, j+1) - f(i+1, j+1) + 2f(i+1, j+1) = f(i+1, j) + f(i, j+1)$.

Опишем теперь алгоритм нахождения $f(1,1)$. Нетрудно заметить, что для нахождения $f(i,j)$, где $i+j=c$, нам достаточно знания $f(l, m)$ при всевозможных $l + m = c + 1$ и $l + m = c + 2$. Кроме того, из пунктов 1 и 2 нам известны $f(k+1, n)$, $f(k, n+1)$ и $f(k+1, n+1)$. Значит, мы можем последовательно вычислить $f(i, j)$, рассмотрев сначала случай $i+j=k+n$, затем $i+j=k+n-1$ и т. д. В итоге мы получим $f(1,1)$, при этом в силу такого построения число действий будет $O(kn)$. Теорема доказана.

Предположим, что классы эквивалентности занумерованы числами от единицы до $f(1,1)$, и опишем алгоритм построения пути, принадлежащего s-му классу эквивалентности. После выполнения алгоритма множество V_s содержит вершины, принадлежащие некоторому пути из s-го класса эквивалентности.

Алгоритм:

```
1. V_s = ∅, i = 1, j = 1, class = s
2. while (j ≤ n + 1) {
3.       while (1) {
4.             V_s = V_s ∪ v_j^i;
5.             if (i == k + 1 || f(i + 1, j) < class) break;
6.             i = i + 1;
7.       }
8.       if (j < n + 1)
9.             class = class – (f(i, j) – f(i, j + 1));
10.      j = j + 1;
11.      }
```

Нетрудно заметить, что построение пути требует $O(n+k)$ действий.

Оценка числа классов эквивалентности

Представим число операций, выполняемых первым потоком, в следующем виде: $k = \sum_{t=1}^{M} r_t^1 + w_t^1 + x_t^1 + v_t^1$. Здесь M – число общих ячеек памяти, для t-й ячейки r_t^1, w_t^1, x_t^1, v_t^1 – количество операций чтения, записи, хранения значений и других соответственно. Для второго потока $n = \sum_{t=1}^{M} r_t^2 + w_t^2 + x_t^2 + v_t^2$. Используя лемму 2, можно найти число некоммутирующих операций (число плюсов на графе): $\sum_{t=1}^{M} r_t^1 w_t^2 + r_t^2 w_t^1 + w_t^1 w_t^2$.

Теорема 2. Для числа классов эквивалентности полных путей C справедливо

$$1+\sum_{i=1}^{M} r_i^1 w_i^2 + r_i^2 w_i^1 + w_i^1 w_i^2 \le C \le$$
$$\le 2 \wedge \sum_{i=1}^{M} r_i^1 w_i^2 + r_i^2 w_i^1 + w_i^1 w_i^2.$$

Доказательство. Покажем, что превращение некоторой пары коммутирующих операций в некоммутирующие увеличивает число классов эквивалентности не меньше чем на единицу и не более чем в два раза. Действительно, рассмотрим пути, принадлежащие некоторому классу эквивалентности. Если хотя бы в двух из них порядок дуг, представляющих указанные операции, различен, то класс разбивается на два новых. Покажем теперь, что хотя бы один класс будет разбит. Рассмотрим два пути из утверждения 1 при $x=1$, $y=1$ и i, j, совпадающих с номерами операций. До превращения операций эти пути принадлежали одному классу эквивалентности, после – разным. Осталось заметить, что при отсутствии пар некоммутирующих операций класс эквивалентности будет единственным. Теорема доказана.

Рассмотрим пример. Пусть $n=k=5$, число некоммутирующих операций $\sum_{i=1}^{M} r_i^1 w_i^2 + r_i^2 w_i^1 + w_i^1 w_i^2 = 4$. Максимальное число классов эквивалентности полных путей 16, минимальное – 5 (рис. 4). Общее же число вариантов полных путей $C_{10}^5 = \frac{10!}{5!5!} = 252$.

Построение редуцированного графа и анализ результатов

В силу определения классов эквивалентности нам достаточно вычислить результирующие состояния исполнения только для одного представителя каждого из классов. Удалим из графа дуги, которые не входят в полные пути, соответствующие выбранным представителям классов. Мы получим редуцированный граф совместного исполнения потоков. При этом результирующее состояние исполнения должно остаться тем же самым. Вычислим его в неопределенных коэффициентах для всех ячеек памяти, двигаясь от начальной вершины графа к конечной. Для того чтобы определить, возможно ли возникновение неразрешенного состояния гонки, нам остается только проанализировать полученное состояние.

Часто признаком правильной работы программы является независимость некоторых компонент результирующего состояния от выбора варианта исполнения потоков (например, в одной из ячеек памяти после работы потоков все-

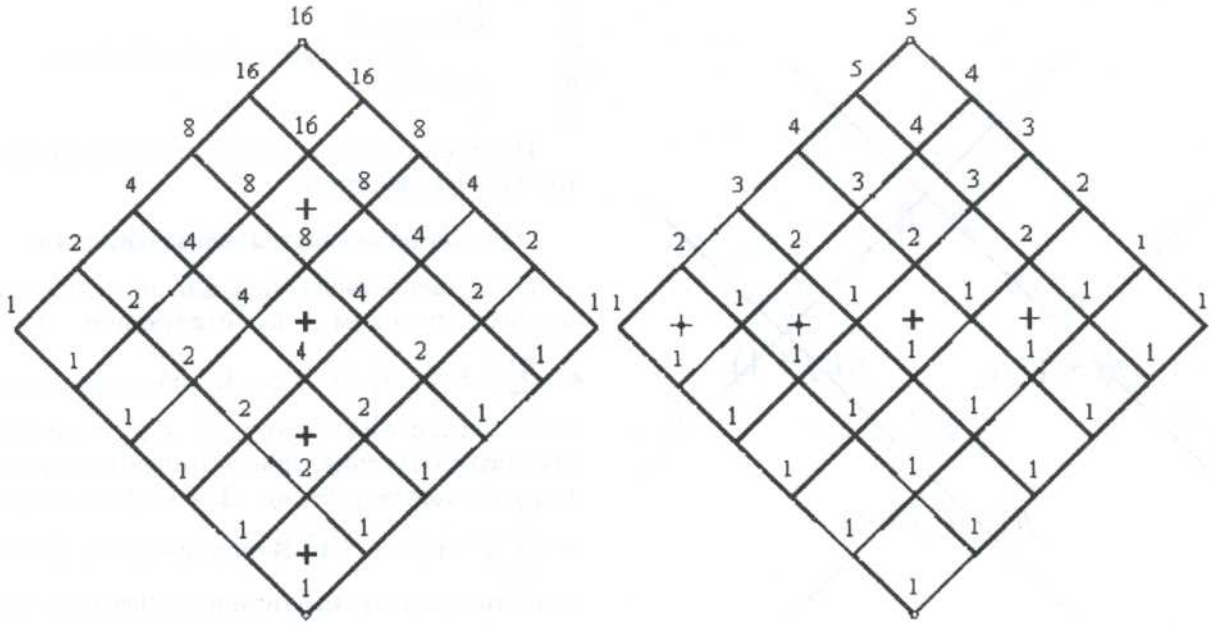


Рис. 4. Максимальное и минимальное число классов эквивалентности

гда должна быть записана некоторая константа *c*). Тогда, если существует два набора неопределенных коэффициентов, таких, что их подстановка дает различный результат, то можно утверждать, что возможно возникновение неразрешенного состояния гонки.

Были исследованы две простые задачи. В случае задачи об изменении значения ячейки в двух потоках число классов эквивалентности оказалось равным четырем при общем числе вариантов исполнения, равном 20. В случае задачи о транзакции число классов эквивалентности также равно четырем при общем числе вариантов исполнения 28.

В работе был описан подход, позволяющий гарантированно обнаружить все состояния гонки в программе, приводящие к ее неправильной работе. При этом необходимые расчеты минимизированы за счет объединения вариантов исполнения потоков программы в классы и проведения вычислений только для одного представителя каждого из классов.

СПИСОК ЛИТЕРАТУРЫ

1. **Serebryany K.** Data race test. URL: http://code.google.com/p/data-race-test.

2. **Herlihy M., Shavit N.** The Art of Multiprocessor Programming. Elsevier, 2008.

3. **Бурков В.Н., Заложнев А.Ю., Новиков Д.А.** Теория графов в управлении организационными системами. М.: Синтег, 2001.

4. **Rahul V. Patil, George B.** Concurrency: Tools And Techniques to Identify Concurrency Issues // MSDN Magazine, June 2008.

5. **Philip N. Klein, Hsueh Lu, Netze R.** Detecting Race Conditions in Parallel Programs that Use Semaphores. Springer Berlin/Heidelberg, 2008.

УДК 621.37

Д.А. Крупенко

АДАПТАЦИЯ СИСТЕМЫ ПРОЕКТИРОВАНИЯ AUTOCAD ДЛЯ РАЗРАБОТКИ КОНСТРУКЦИЙ МИКРОСБОРОК СВЧ

Современная микроэлектроника достигла такого уровня сложности, что изделия стали узкоспециализированными и системно-ориентированными, в связи с этим постоянно растут потребности в разработке новых устройств, позволяющих решать частные задачи. Так, в области проектирования СВЧ-устройств огромное значение приобретает разработка гибридных интегральных схем, в частности их разновидности, так называемых микросборок.

Микросборка, или электронный модуль СВЧ, – это изделие электронной техники для диапазона частот 3–30 ГГц, имеющее законченное конструктивное исполнение и состоящее из одного или нескольких функциональных узлов СВЧ, взаимозаменяемое и неремонтопригодное в условиях эксплуатации. Пример модуля СВЧ изображен на рис. 1.

Особенность микросборок состоит в том, что они представляют собой изделия частного применения, т. е. изготовляются для конкретного типа аппаратуры.

Создание модуля предполагает разработку его конструкции и технологии изготовления. К современным модулям СВЧ предъявляется большое число сложных, часто взаимоисключающих требований, среди которых:

высокий уровень электрических параметров с учетом конструктивно-технологических запасов;

прочность и устойчивость к внешним воздействующим факторам – механическим, климатическим и специальным;

надежность и длительная сохраняемость;

минимальные габариты, установочные и присоединительные размеры и масса;

приемлемые способы охлаждения и крепления в аппаратуре;

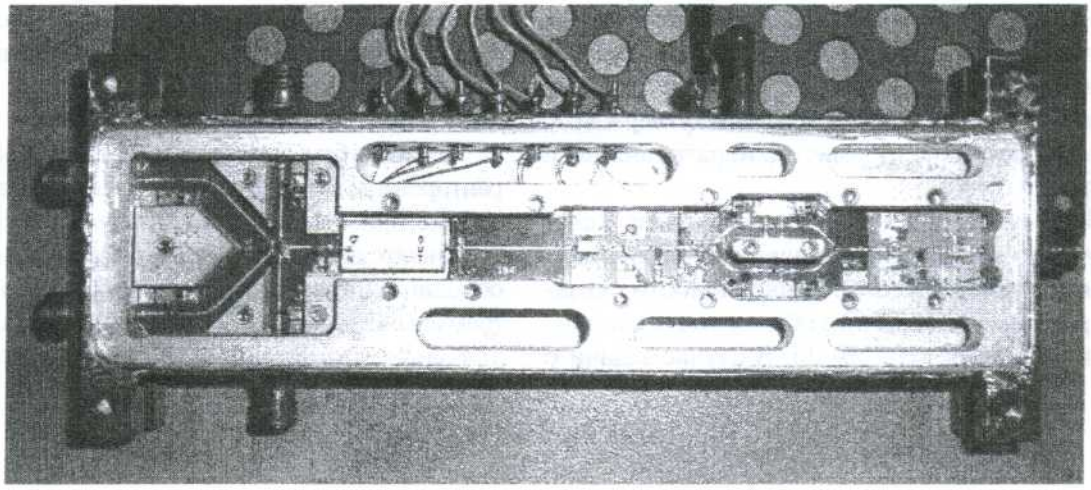


Рис. 1. Внешний вид модуля СВЧ

определенные типы вводов и выводов энергии, напряжения питания и сигналов управления;

производственная технологичность, требования стандартизации и унификации, высокий процент выхода годных изделий, низкая стоимость и др. [1].

Для того чтобы выполнить должным образом указанные выше требования, необходим тщательный подход к разработке конструкции узлов с самого начала. И это во многом зависит от системы автоматизированного проектирования, которая используется в качестве инструмента для проектирования микросборки. Стоит отметить, что на рынке до сих пор не существует специализированной системы автоматизированного проектирования, предназначенной для разработки конструкции электронных модулей СВЧ. Разработка узлов, как правило, производится в универсальных системах проектирования, которые имеются на предприятии, например AutoCAD, Компас, Solid Works и др. В связи с этим оптимальным и наименее затратным кажется адаптация имеющейся системы для работы по проектированию микросборок.

Наиболее популярная и распространенная система автоматизированного проектирования на сегодняшней день – AutoCAD. Его возможности можно расширить и с помощью встроенного языка программирования Visual LISP, и с использованием динамических блоков. Рассмотрим подробней возможность адаптации этой системы проектирования для разработки конструкции микросборок.

Процесс разработки СВЧ-узла после изучения технического задания начинается с этапа компоновки, а именно с прокладки микрополосковой линии. Конструктор в соответствии с электрической принципиальной схемой должен построить высокочастотный тракт. Элементная база, используемая при проектировании микросборок, довольно разнообразна, но основная масса элементов, таких, как конденсаторы, катушки индуктивности, резисторы, аттенюаторы, микросхемы, как правило, выполнена в одинаковых корпусах, хотя может иметь различные номиналы и выполнять различные функции. При разработке микрополосковой линии конструктору, работающему в системе проектирования AutoCAD, приходится неоднократно отрисовывать контуры элементов или копировать их из других чертежей, что неизбежно приводит к ошибкам. Эту работу можно значительно упростить, создав параметрические элементы, содержащие описание элементной базы.

Элемент может быть выполнен в виде динамического блока, характеризующегося определенными параметрами. Гибкость динамического блока обеспечивается сочетанием его параметров и действий. Выбор необходимых элементов из графической базы данных можно осуществить с помощью приложения, созданного на встроенном в AutoCAD языке программирования Visual LISP. Для этого необходимо использовать функции ActiveX, которые позволяют работать с файлами *.dwg, содержащими графическую базу данных, т. е. динамические блоки с описанием элементов конструкции СВЧ-узла. Вставка

описания элемента осуществляется командой vla-insertblock и указанием рабочего пространства, куда производится вставка элемента: vla-get-modelspace. Для связи LISP-приложения с рабочим пространством AutoCAD необходимо создать диалоговое окно, используя встроенный язык DCL, в котором пользователь сможет выбирать необходимый ему элемент конструкции узла (рис. 2, *а*). Интерпретатор Visual LISP вычисляет значение каждого введенного выражения и возвращает его в среду AutoCAD. Возвращаемое значение в свою очередь может быть либо передано для дальнейших вычислений в выражение более высокого уровня, либо сохранено в переменной с помощью функции setq. Используя эту функцию, можно определить значения переменных, которые работают с кнопками диалоговой формы, т. е. связать кнопки диалогового окна с LISP-приложением [2].

Если накоплено большое количество динамических блоков, соответствующих чертежным фрагментам, деталям, узлам, то их можно считать графической базой данных. Программы на Visual LISP в сочетании с пользовательскими меню могут организовывать просмотр, поиск, подключение к объектам их частей и т. п. [3]. Тогда работа конструктора в системе AutoCAD сведется к поиску нужных объектов (сборочных единиц, деталей) или частей чертежа, обращению к соответствующим LISP-программам и ответам на вопросы этих программ о параметрах того или иного узла или стандартного обозначения. Каждый такой блок может содержать описание конкретного типа элементов, например конденсатора. Конденсаторы, как и все остальные элементы, могут отличаться типом корпуса, или (в случае использования бескорпусных элементов) типоразмером. Для этого в описании блока необходимо установить параметр видимости и позволить пользователю выбирать нужный ему тип корпуса из выпадающего списка, вызываемого нажатием на специальную ручку рядом с изображением элемента (рис. 2, *б*). Также для упрощения установки элемента в нужное место в описание блока необходимо добавить параметр поворота на заданный угол (рис. 2, *в*). Таким образом может быть создано множество элементов, необходимое для проектирования микроплат.

При проектировании микросборок, как и практически в любой области проектирования, конструктору не обойтись без использования крепежных элементов (винтов, болтов, гаек, шайб). Автоматическую отрисовку элементов крепления в соответствии с ГОСТами можно осуществить, также используя параметрические элементы. Для вставки в рабочее пространство чертежа винта следует создать динамический блок, содержащий описание винта, с возможностью выбора диаметра резьбы (рис. 3, *б*) аналогично выбору типа корпуса радиоэлемента, описанного выше. Диаметр резьбы указывается в соответствии с ГОСТом для каждого конкретного типа винта. Длина выбирается перемещением специальной ручки вдоль оси винта, и она должна соответствовать стандартным длинам для каждого диаметра. Конструктор, не совершая дополнительных действий, получает возможность сразу определить необходимый тип крепежа, так

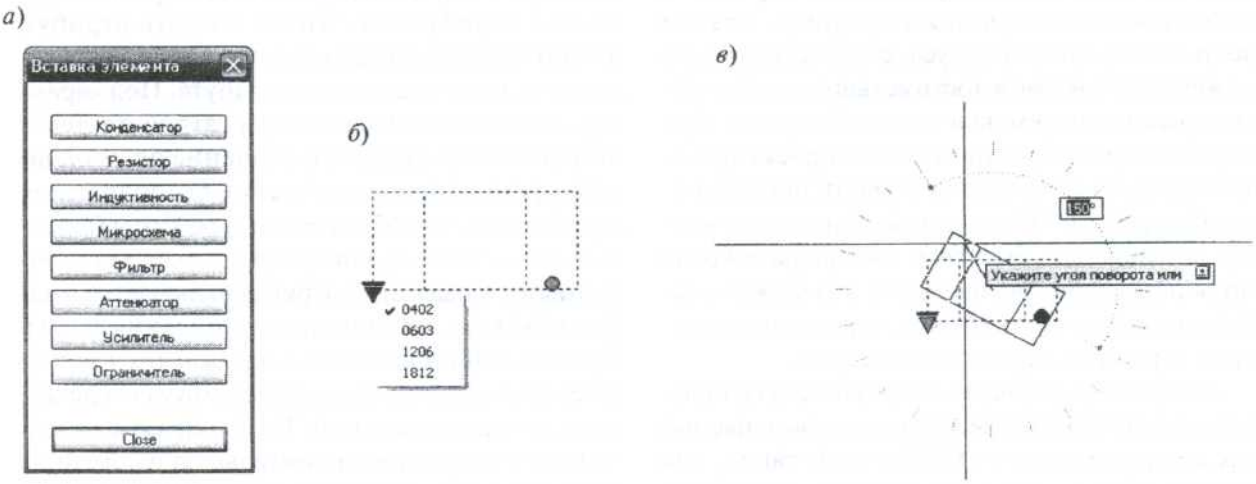


Рис. 2. LISP-приложение, выполняющее вставку радиоэлемента в рабочее пространство чертежа

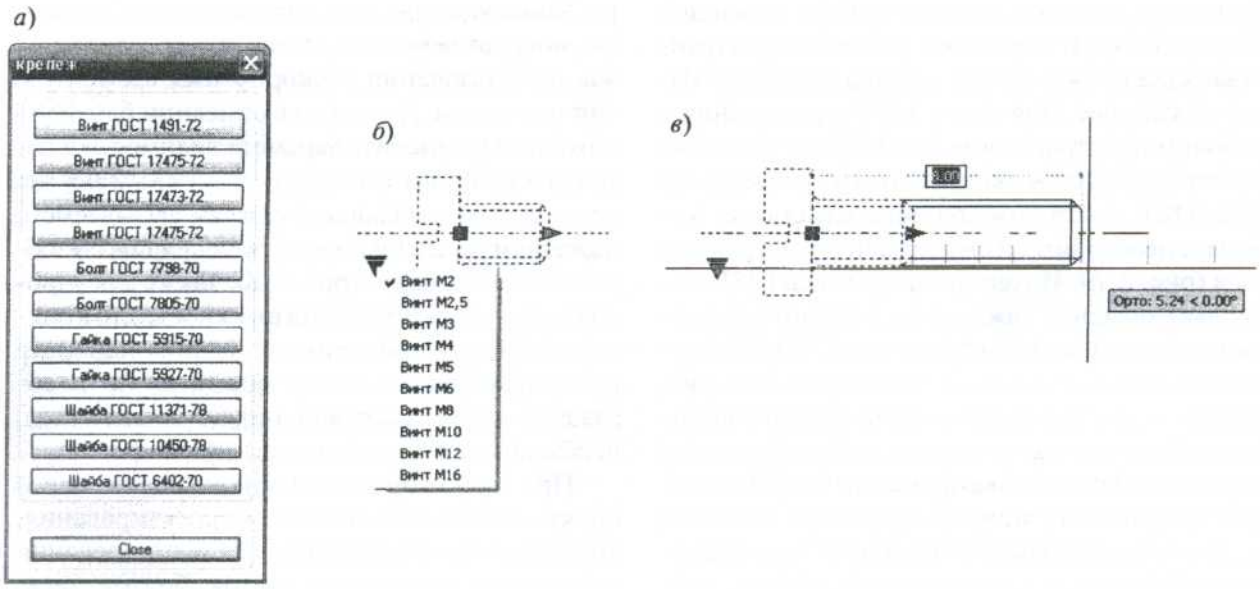


Рис. 3. LISP-приложение, выполняющее вставку элемента крепежа в рабочее пространство чертежа

как на экране изменение происходит в реальном времени (рис. 3, *в*). Например, при установке в корпус микросборки микроплаты с металлическим основанием разработчик может точно определить длину винта так, чтобы он не выступал за габарит корпуса и не был слишком коротким, обеспечивая надежную фиксацию платы к корпусу.

В качестве вводов и выводов энергии в модулях СВЧ наиболее широко применяются коаксиально-микрополосковые переходы – как отечественные (например, СРГ-50-751ФВ), так и импортные (переходы типа SMA различных типов) [4]. При проектировании микросборок конструктору необходимо выбрать тип перехода в соответствии с техническим заданием, а также учесть особенности корпуса, следовательно, он должен иметь возможность вставить графическое описание перехода в текущий чертеж. При разработке расширения системы проектирования AutoCAD для работы в области проектирования устройств СВЧ целесообразно предусмотреть возможность быстрого выбора различных типов переходов. Осуществить это можно, используя динамические блоки, содержащие описания переходов определенной серии.

Аналогичным способом можно создать различные конструктивные элементы, применяемые при конструировании СВЧ-узлов, таких, как простые корпуса, крышки, экраны, основания, откачные трубки, земляные штыри и пр.

Конечный итог работы конструктора, проектирующего микросборки, как и в других областях конструирования, – это комплект конструкторской документации, оформленный в соответствии со стандартами и ГОСТами. Так как система проектирования AutoCAD – продукт зарубежный, она не адаптирована для работы по правилам Единой системы конструкторской документации. Использование связки Visual LISP и динамических блоков, а также блоков с атрибутами может решить эту задачу. Для создания и удобного заполнения основной и дополнительных надписей целесообразно использовать блоки с редактируемыми атрибутами. Чтобы создать атрибут, нужно вначале создать его описание, куда заносятся характеристики атрибута. Под характеристиками понимаются имя атрибута, текст выдаваемой подсказки, значение по умолчанию, форматирование текста, расположение атрибута и его необязательные режимы (скрытый, постоянный, контролируемый, установленный). Созданный атрибут включается в набор объектов при создании описания блока. При очередной вставке блока выдаются запросы с заданными в описаниях атрибутов подсказками для ввода значений. Таким образом, заполнить основную и дополнительную надписи на чертеже можно, просто отвечая на вопросы программы.

Упрощение работы конструктора по созданию чертежей деталей и сборочных единиц требует организации процедуры автоматического подбора формата листа, в который вписывается чертеж. Используя внутренние переменные системы проектирования AutoCAD extmin и extmax, можно легко определить габариты графического изображения и в соответствии с полученными результатами выбрать из графической базы данных соответствующий формат листа, включающий в себя рамку формата, а также блоки с редактируемыми атрибутами основной и дополнительной надписи.

Основываясь на вышеизложенном, можно сделать выводы о том, что система автоматизированного проектирования AutoCAD может быть адаптирована для использования в качестве инструмента разработки конструкции СВЧ-узлов. Встроенный в систему язык программирования VisualLISP позволяет организовать графическую базу данных, содержащую описания используемых при разработке микросборок стандартизованных деталей и узлов, а также деталей общей применяемости. Язык диалоговых окон DCL позволяет организовать удобный доступ к этой графической базе данных. Таким образом, можно создать приложение на базе одной из самых популярных систем автоматизированного проектирования AutoCAD, которое будет адаптировано для конструирования СВЧ-узлов, а также оформления комплекта конструкторской документации в соответствии с правилами Единой системы конструкторской документации и которое сохранит при этом все возможности самой системы AutoCAD.

СПИСОК ЛИТЕРАТУРЫ

1. **Джуринский К.Б.** Конструктивные и технологические особенности модулей СВЧ. Современная электроника. 2008. № 1. С. 22–27.

2. **Зуев С.А., Полещук Н.Н.** САПР на базе AutoCAD – как это делается. СПб.: БХВ-Санкт-Петербург, 2004.

3. **Полещук Н.Н.** AutoCAD: разработка приложений, настройка и адаптация. СПб.: БХВ-Санкт-Петербург, 2006.

4. **Джуринский К.Б.** Миниатюрные коаксиальные радиокомпоненты для микроэлектроники СВЧ. М.: Техносфера, 2006.

УДК 004.42:347.77+338.32.053.4

К.С. Солнушкин

СИСТЕМА УЧЕТА ДОСТУПА К ЛИЦЕНЗИОННОМУ ПРОГРАММНОМУ ОБЕСПЕЧЕНИЮ В ЦЕНТРАХ КОЛЛЕКТИВНОГО ПОЛЬЗОВАНИЯ

Современное прикладное программное обеспечение (ПО), в частности наукоемкое ПО для инженерного анализа, весьма часто распространяется по дорогостоящим коммерческим лицензиям. В этой связи возникает вопрос о его эффективном использовании в рамках учреждения.

Основной метод повышения эффективности использования имеющихся ресурсов – коллективное использование. На практике он реализуется путем создания центров коллективного пользования (ЦКП), в которых совместно используются дорогостоящие ресурсы: высокопроизводительные вычислительные системы и лицензионное программное обеспечение.

Так, в зарубежных университетах число наименований программного обеспечения достигает нескольких десятков (см., например, [5]), большая часть этого ПО требует покупки лицензии. В российских университетах суммы, выделяемые на закупку ПО, также постепенно растут. Например, Томский политехнический университет обладает лицензиями на ПО компании "Schlumberger" общей стоимость 20 млн долл. США [1].

Оптимизация расходов на лицензионное ПО производится за счет закупки требуемого числа лицензий. Недостаточное число лицензий ведет к задержкам в работе пользователей, так как

запуск лицензионного ПО, необходимого для работы, возможен только после освобождения лицензий, занятых другими пользователями. Излишнее число лицензий, напротив, ведет к нерациональному расходованию средств учреждения.

Таким образом, возникает задача определения оптимального числа лицензий. Эту задачу следует решать, анализируя потребности пользователей ЦКП в лицензионном ПО. Для этого необходимо учитывать количество лицензий, задействованных пользователями в каждый момент времени. Очевидно, что в крупных ЦКП это требует применения автоматизированной системы учета.

В ряде случаев также необходимо определить количественный вклад отдельных пользователей в потребление лицензий с целью назначения оплаты.

В тех ЦКП, где используются системы управления потоками задач (планировщики задач), возникает новый аспект использования лицензионного ПО: задача может начать выполняться только в том случае, если имеются все необходимые лицензии. В противном случае задача завершится с ошибкой, и потребуется ее повторная постановка в очередь. Эта проблема требует взаимодействия планировщика задач и серверов лицензий и до сих пор не получила удовлетворительного решения. Нами предлагается открытый расширяемый интерфейс опроса серверов лицензий.

В статье предложена архитектура программной системы, производящей учет доступа к лицензионному ПО в ЦКП, анализ собранных данных и представление результатов анализа лицам, принимающим решения по закупке лицензий. Прототип программной системы, реализующей предложенную архитектуру, разработал сотрудник радиофизического факультета СПбГПУ кандидат технических наук Д.И. Иванов.

Обзор имеющихся работ. Технические функции лицензирования не встроены в лицензируемое ПО или в операционную систему. Лицензирование ПО производится серверами лицензий – программно-аппаратными комплексами, которые принимают запрос от лицензируемого ПО и в зависимости от наличия свободных лицензий на те или иные модули ПО разрешают или запрещают работу этих модулей.

Например, для работы с программной системой MATLAB разным категориям пользователей могут потребоваться разные модули: цифровой обработки сигналов, математической статистики, параллельных вычислений и др. Работа каждого модуля лицензируется отдельно.

В условиях глобализации географически распределенные корпорации могут использовать серверы лицензий, находящиеся как под собственным, так и под внешним административным контролем, расположенные в различных местах земного шара и объединенные сетью Интернет. Это усложняет задачу учета доступа к лицензионному ПО.

Наиболее распространенным ПО для серверов лицензий является FLEXnet Publisher [4]; это ПО создает файл журнала, в который записываются данные о запросе и освобождении лицензий, однако формат этого файла – закрытый, а цена ПО для анализа этого файла неоправданно высока для отечественных ЦКП.

С другой стороны, имеется возможность регулярно опрашивать сервер лицензий на основе ПО FLEXnet Publisher и получать информацию о том, какие модули лицензионного ПО используются в данный момент. По каждому модулю можно получить информацию, сколько лицензий использовано и кем (с точностью до DNS-имен рабочих станций и идентификаторов пользователей на рабочей станции).

Существуют и альтернативные серверы лицензий, например Reprise License Manager (RLM) [10]. Этот сервер лицензий более современный и обладает большей функциональностью; он создает файл журнала в открытом формате.

Этот формат распознается несколькими утилитами, которые позволяют анализировать файл журнала и представлять результаты в удобном для человека виде. К числу таких утилит относятся License Tracker [6], License Monitor [8] и License Analyzer [7]. Эти утилиты являются закрытым, платным ПО частных компаний; кроме того, они не предназначены для взаимодействия с планировщиками задач.

Предлагаемая архитектура системы учета. Несмотря на указанные ограничения существующих серверов лицензий и систем учета, построение гибкой, модульной, наращиваемой системы учета возможно за счет открытого и расширяемого интерфейса взаимодействия между предлагаемой системой, серверами лицензий и планировщиками задач.

Для того чтобы определить, насколько эффективно используются лицензии организации, необходимо иметь данные о том, сколько лицен-

зий занято в каждый момент времени. Эти данные можно получать либо из файлов журналов серверов лицензий, либо путем их регулярного опроса. Однако у самого распространенного сервера лицензий [4] файл журнала имеет закрытый формат. Другие серверы лицензий, например [10], практически не используются для лицензирования современного наукоемкого ПО. Таким образом, универсальным способом получения данных об использовании лицензий в учреждении де-факто является регулярный опрос серверов лицензий.

Кроме того, для функционирования планировщиков задач с учетом доступности лицензий необходим именно опрос серверов лицензий в реальном масштабе времени.

Метод сбора данных с применением опросов хотя и универсален, имеет свои недостатки. Процедуры опроса проводятся через определенные интервалы времени. Если какая-либо лицензия запрашивается на короткий период времени, который укладывается в промежуток между двумя опросами, то система учета не зафиксирует это событие. Это может повлечь за собой определенное искажение в собранных данных. Указанная проблема решается увеличением частоты опросов, что, однако, увеличивает и интенсивность обработки собираемых данных.

Кроме того, в отличие от анализа файлов журналов, опрос серверов лицензий не может выявить ситуацию, при которой лицензируемое приложение не смогло продолжить работу из-за отсутствия свободных лицензий. Таким образом, невозможно выяснить, насколько часто появляется ситуация нехватки лицензий и появляется ли вообще. Налицо потеря информации, которая заставляет принимать более осторожные решения по итогам анализа данных.

Тем не менее универсальность метода опроса и его пригодность для обеспечения работы планировщиков задач обусловливают его использование в проектируемой системе учета.

В процессе работы системы учета все субъекты информационного взаимодействия – лицензируемое ПО, серверы лицензий, модули системы учета – могут размещаться на различных ЭВМ, связанных сетью Интернет. Систему учета следует проектировать исходя из этого обстоятельства; кроме того, желательно, чтобы отдельные модули системы также обладали возможностью потенциального размещения на отдельных ЭВМ и взаимодействовали друг с другом и с остальными субъектами через сеть Интернет. Это является необходимым условием расширяемости системы в будущем.

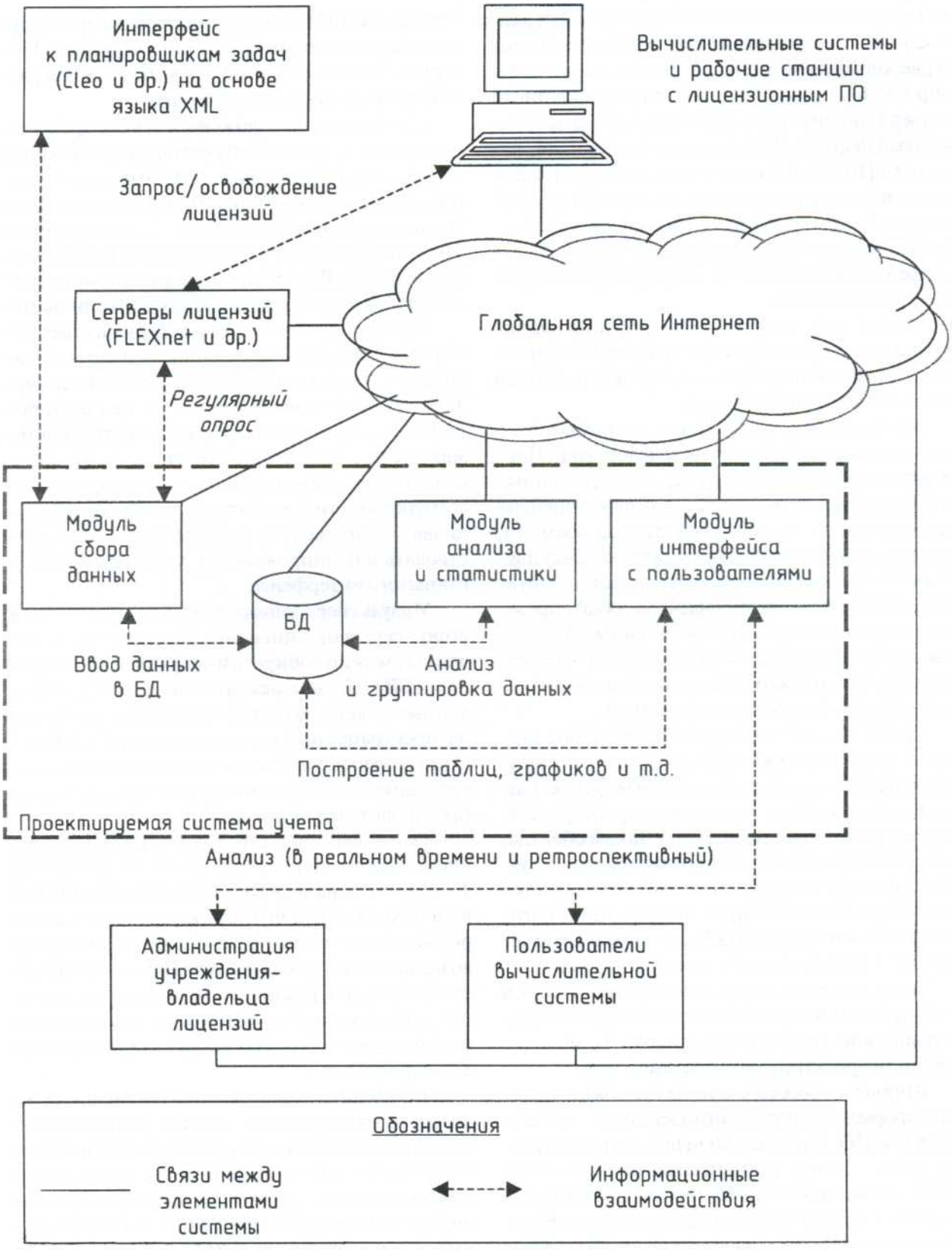
С учетом вышеизложенного нами предлагается архитектура системы учета, представленная на рисунке. В этой архитектуре выполнено требование, согласно которому все субъекты информационного взаимодействия связаны с сетью Интернет; это обозначено на рисунке сплошными линиями. Различные информационные взаимодействия изображены штриховыми линиями.

Проектируемая система состоит из трех основных модулей: сбора данных, анализа и статистики, а также интерфейса с пользователями. Такой способ деления на модули не нов и, по-видимому, применяется и в аналогичных системах учета с закрытой архитектурой. Новое – свойство расширяемости, являющееся результатом открытости архитектуры, а также отсутствующая у конкурентов способность взаимодействовать с планировщиком задач при помощи открытого интерфейса.

Модуль сбора данных отвечает за регулярный опрос серверов лицензий. Рассмотрим, какие данные можно собирать и сохранять в базе данных (БД). (Будем исходить из той структуры данных ответа, которую использует наиболее распространенный сервер лицензий FLEXnet.) Один сервер может быть использован для лицензирования нескольких видов ПО от различных производителей (англ. vendor).

Взаимодействие сервера и утилиты опроса происходит в закрытом формате по протоколу TCP; утилита направляет на стандартный вывод в удобочитаемом для человека виде ответ, который подлежит разбору. В ответ на запрос сервер возвращает: по каждому типу ПО – список модулей этого ПО, на которые приобретены лицензии, а по каждому модулю – сколько лицензий приобретено и сколько из них используется в данный момент.

По каждой использованной лицензии также возвращаются учетные данные, позволяющие сопоставить использование лицензий отдельным подразделениям и пользователям учреждения: DNS-имя вычислительной системы, запросившей лицензию (рабочая станция, вычислительный узел кластерной ЭВМ), идентификатор пользователя на вычислительной системе, а в случае удаленного доступа к вычислительной



Архитектура проектируемой системы

системе – также и DNS-имя рабочей станции пользователя.

Структура реляционной БД для хранения информации о том, лицензии на какое ПО и в каком количестве имелись у учреждения в каждый момент времени, представлена в табл. 1.

Таблица 1

Пример отношения реляционной БД для хранения информации о лицензиях, имеющихся у учреждения

Имя сервера лицензий	Наименование ПО / поставщик ПО	Наименование модуля ПО	Количество лицензий на модуль	Дата начала действия лицензии
ALICE	MATLAB	Parallel Toolbox	10	01 января 2008
ALICE	MATLAB	Signal Processing	5	01 января 2008
ALICE	ANSYS	Workbench	7	12 марта 2008
BOB	MSC	Adams View	12	24 апреля 2008
ALICE	MATLAB	Signal Processing	10	15 июня 2008
ALICE	MATLAB	Signal Processing	0	20 июля 2008
BOB	MATLAB	Signal Processing	10	23 июля 2008

В табл. 1 приведена ситуация, при которой учреждение разместило лицензии для двух программных продуктов – MATLAB и ANSYS – на сервере ALICE, а лицензию на ПО MSC – на сервере BOB. Далее количество лицензий на модуль Signal Processing увеличивается с пяти до десяти. Затем срок действия лицензий на этот модуль завершился, и после трехдневного перерыва новые лицензии стали обслуживаться уже сервером BOB.

В этом же отношении возможно с помощью специального признака фиксировать временную неработоспособность серверов лицензий.

За счет того, что в БД сохраняются только изменения по сравнению с предыдущим состоянием, удается в компактном формате хранить всю значимую информацию по наличию лицензий.

Аналогичная концепция хранения применяется и по отношению к учетным данным (табл. 2).

Схема построения этого отношения следующая. Задаем интервал T опроса сервера лицензий. Анализируем ответ сервера. Для каждой комбинации (*Идентификатор пользователя, DNS-имя ЭВМ, Наименование ПО, Наименование модуля ПО, Количество лицензий на модуль*) в ответе сервера пытаемся отыскать в отношении кортеж с таким же значением полей и, кроме того, у которого значение поля "Время окончания использования" соответствует текущему времени (с некоторым допуском).

Если такой кортеж не найден, то создаем его, фиксируя таким образом ситуацию начала использования лицензий.

Если кортеж найден, продлеваем значение в поле «Время окончания использования» на величину интервала опроса T, тем самым фиксируя продолжение использования лицензий.

Этот алгоритм гарантирует, что в БД всегда будет храниться самая полная информация об использовании лицензий, какую можно получить методом регулярного опроса.

Уменьшая интервал опроса T, мы увеличиваем детализацию представления данных. При этом БД обновляется чаще, тогда как размер ее не растет пропорционально интенсивности опросов. Впрочем, размер все же несколько увеличивается, так как при более частых опросах удается зафиксировать эффект коротких перерывов в использовании лицензий одним и тем же пользователем.

Для экономии размера БД текстовые значения в отношениях (см. табл. 1 и 2) целесообразно кодировать уникальными целочисленными идентификаторами (а соответствия хранить в других

Таблица 2

Пример отношения реляционной БД для хранения учетных данных

Идентификатор пользователя	DNS-имя ЭВМ	Наименование ПО / поставщик ПО	Наименование модуля ПО	Количество лицензий на модуль	Время начала использования	Время окончания использования
User1	Comp7	MATLAB	Parallel Toolbox	2	10:00	10:15
Stud2	Node5	MATLAB	Signal Processing	1	10:10	10:20
User1	Comp7	MATLAB	Parallel Toolbox	4	10:15	10:20

отношениях). В этом случае для задания одного кортежа в отношении (см. табл. 2) потребуется порядка 30 байт, а на одну рабочую станцию, активно использующую лицензии в течение дня, придется порядка десятков килобайт учетных данных.

Модуль анализа и статистики служит для того, чтобы анализировать накопленные учетные данные, группировать их, выявлять статистические закономерности и тенденции.

Группировка производится в целях представления учетных данных в удобных для восприятия человеком разрезах:

по наименованиям лицензионного ПО;

отдельным модулям ПО внутри одного наименования;

подразделениям учреждения;

отдельным пользователям;

временным интервалам и др.

Результаты группировки сохраняются в БД в целях кэширования при повторных обращениях; кроме того, это позволит удалять исходные учетные данные, чтобы сократить размер БД.

На модуль анализа также может быть возложена дополнительная функциональность, например, оповещать системных администраторов по электронной почте при определенных условиях: неработоспособности серверов лицензий, малом количестве доступных лицензий на отдельные модули и т. д. (подобная функциональность имеется и в конкурирующих решениях).

Анализируя учетные данные, можно подсчитать (с точностью до отдельного пользователя) потребление услуг по предоставлению лицензий подобно тому, как производится учет потребления электроэнергии, в единицах "лицензия·минута". Так, из табл. 2 следует, что пользователь с идентификатором User1 потребил услуги предоставления лицензий на модуль Parallel Toolbox в общей сложности на величину 50 лиц.·мин. Подобная информация может быть использована для выставления счетов.

Модуль интерфейса с пользователями служит для выдачи информации в форме, удобной для восприятия человеком – лицом, принимающим решения по закупке лицензий. Этот модуль выдает информацию в виде графиков и таблиц, в которых представлено усредненное и пиковое потребление лицензий за крупные интервалы времени – сутки, месяцы, годы – в разрезе по подразделениям и др. Это позволяет анализировать деятельность учреждения и прогнозировать закупку лицензий в долгосрочной перспективе.

Основное предназначение модуля – помочь управленческому персоналу принимать обоснованные решения по управлению фондом лицензий предприятия. С этой целью модуль позволяет быстро выявить лицензии, которые используются неэффективно, т. е. такие, для которых доля использованных лицензий ни в какой момент времени не превышала заданного порога, например 70 % (иными словами, 30 % лицензий никогда не были использованы). Также выявляются лицензии, которые находятся в дефиците, т. е. такие, для которых доля использованных лицензий большую часть времени находится выше определенного порога. Пороги задаются интерактивно в ходе работы с модулем.

Для удобства представления информации в отчетах можно сопоставлять идентификаторы пользователей их реальным именам. Для этого предлагается использовать службы каталогов с открытыми интерфейсами взаимодействия, например LDAP. Такая функциональность имеется в [7].

Интерфейс к планировщикам задач. В крупных ЦКП применяются системы управления очередями задач (планировщики задач). Задача пользователя может несколько часов находиться в очереди, ожидая доступа к ресурсам вычислительной системы. Если в момент запуска необходимые лицензии не будут доступны, задача завершится с ошибкой, и потребуется постановка в очередь заново.

Поэтому необходимо предоставлять системе управления очередями задач информацию о том, доступны ли лицензии и можно ли таким образом запускать задачу на счет. С этой точки зрения лицензии на ПО представляют собой информационный ресурс вычислительной системы [2], доступность которого является необходимым условием для запуска задачи.

Когда задача попадает в начало очереди, планировщик задач анализирует доступность аппаратных ресурсов для ее выполнения (например, свободных узлов кластерной ВС). В случае взаимодействия с предлагаемой программной системой планировщик задач также проверяет доступность необходимых лицензий; если оба вида ресурсов доступны, задача может быть отправлена на счет.

Так, планировщику задач для ОС Microsoft Windows HPC Server 2008 предложено дополнение, которое перед запуском задачи на счет производит опрос сервера лицензий [3]. Планировщику задач PBS Pro компании "Altair" можно статичес-

ки указать, лицензии какого типа и в каком количестве имеются, и перед запуском задачи планировщик будет проверять доступность нужного количества запрошенных лицензий в пуле; в этой схеме опрос серверов не применяется [9].

В разработанной программной системе в модуле сбора данных предусмотрен открытый интерфейс с планировщиками задач на основе языка XML, а в прототипе реализован интерфейс с российской системой "Cleo" (разработчик – кандидат физико-математических наук С.А. Жуматий, НИВЦ МГУ). Система "Cleo" обеспечивает прохождение задач на крупнейшей сегодня университетской кластерной ЭВМ в России и СНГ – суперкомпьютере "СКИФ-МГУ".

Вне зависимости от используемого планировщика задач при постановке задачи в очередь пользователь тем или иным образом указывает такие сведения, как наименование ПО, модулей ПО и количество лицензий, которые требуются для успешного выполнения задачи, а также к каким лицензионным серверам предпочтительно обратиться в первую очередь.

В свою очередь планировщик передает эти данные, представленные на языке XML, в предлагаемую программную систему. Модуль сбора данных опрашивает доступные серверы лицензий и выдает актуальную информацию о том, доступны ли запрошенные лицензии и на каком сервере.

Использование актуальной информации о доступности лицензий позволяет сократить вероятность неудачного запуска задач в том случае, если свободных лицензий нет. При необходимости планировщику задачи передается и расширенная диагностика, в текстовом формате, например, если лицензионный сервер недоступен.

Чтобы упростить работу пользователей вычислительной системы, предусмотрен интерфейс, позволяющий определить как общее количество лицензий на ПО определенного типа, имеющихся в наличии в учреждении, так и количество лицензий, свободных в текущий момент. Это даст пользователям возможность выбирать реалистичные значения числа параллельных потоков для запуска своих задач, исходя из числа доступных лицензий, так как параллельные вычисления лицензируются, как правило, именно по числу потоков.

Кроме того, уже в момент постановки задачи в очередь планировщик имеет возможность произвести проверку, не запросил ли пользователь больше лицензий, чем имеется в учреждении. В этом случае задачу можно не ставить в очередь.

В статье предложена модульная, наращиваемая архитектура распределенной программной системы, позволяющая организовать учет доступа к лицензионному ПО в масштабах предприятия и упрощающая принятие обоснованных решений по управлению фондом лицензий. Это повышает эффективность коллективного доступа к дорогостоящему оборудованию и программному обеспечению.

Предлагаемая система способна взаимодействовать с планировщиками задач, обеспечивая их информацией о доступных лицензиях. Это позволяет уменьшить вероятность неуспешного запуска задач из-за недоступности лицензий. Реализованный в соответствии с данной архитектурой прототип интегрирован с системой управления очередями задач "Cleo".

СПИСОК ЛИТЕРАТУРЫ

1. Лицензия на добычу. Web-сайт «ПОИСК – Газета российского научного сообщества». http://www.poisknews.ru/2009/07/15/licenzija_na_dobychu.html

2. **Солнушкин К.С.** О значении термина "ресурсы вычислительной системы" // Научно-технические ведомости СПбГТУ. 2006. № 4 (46). С. 44–46. (Информационные технологии в технике и экономике.)

3. C# Example: An Activation Filter that Checks for License Availability. http://technet.microsoft.com/ja-jp/library/dd346641(WS.10).aspx

4. FLEXnet Publisher. http://en.wikipedia.org/wiki/FLEXnet_Publisher

5. Indiana University. Software Available on High Performance Systems. http://rac.uits.iu.edu/hps/software.shtml

6. License Tracker Toolset. http://www.licensetracker.ca

7. LicenseAnalyzer by Open iT. http://www.openit.com

8. LicenseMonitor by Runtime Design Automation – a software license monitoring product. http://www.rtda.com

9. Matlab on Biowulf. Matlab resources (licenses). http://biowulf.nih.gov/apps/matlab.html

10. Reprise License Manager (RLM). http://www.reprisesoftware.com

УДК 004.514.62

А.А. Сухов

ОЦЕНКА ЭМОЦИОНАЛЬНОЙ СОСТАВЛЯЮЩЕЙ ИНТЕРФЕЙСОВ ТРЕНАЖЕРОВ

Разработка и реализация современных методов создания человеко-машинного интерфейса управления для персонала, принимающего решения и обслуживающего персонала в любых человеко-машинных системах, приобрела сегодня особенно актуальное значение, поскольку это – один из важнейших факторов, определяющих надежность и экологическую безопасность различных технологических объектов и управляющих ими людей.

Техногенные катастрофы национального и транснационального масштаба, ряд авиационных аварий вновь привлекли внимание мировой общественности к вопросу о регламентации систем подготовки и тренировки человека с целью повышения надежности его действий по управлению технологическими объектами в аварийных и экстремальных ситуациях.

Основное требование к интерфейсу автоматизированных систем управления технологическим процессом опасного производства заключается в максимально быстром распознавании опасной ситуации и принятии решения. Главная проблема в том, что часто создание интерфейса находится полностью на совести программиста и очень редко – дизайнера, так как чаще всего в создании автоматизированных систем последний участия не принимает.

Разработчики программных комплексов склонны рассматривать функциональность системы отдельно от её пользовательского интерфейса (ПИ). При этом предполагается, что ПИ – это своего рода дополнение к функциональности системы. Со своей стороны, пользователи программ, как правило, не разделяют функциональность и пользовательский интерфейс. Для пользователей именно ПИ является программой. Для них, если интерфейс хороший, стало быть, и сама программа хороша и удобна.

Метод оценки интерфейса на основе нейронных сетей

Пользовательский интерфейс часто понимают только как внешний вид программы. Однако на деле пользователь воспринимает через ПИ всю систему в целом, а значит, такое понимание ПИ слишком узко. В действительности ПИ включает в себя все аспекты дизайна, которые оказывают влияние на взаимодействие пользователя и системы. Это не только экран, который видит пользователь. Пользовательский интерфейс включает множество составляющих, такие, как:

набор задач пользователя, которые он решает при помощи системы;

используемая системой метафора (например, рабочий стол в MS Windows и т. п.);

элементы управления системой;

навигация между блоками системы;

визуальный (и не только) дизайн экранов программы.

Теперь необходимо определиться с тем, что подразумевать под удобным интерфейсом.

Практичность (Usability) (по ИСО/МЭК 9126-1:1998) – это способность ПО, обусловливающая легкость его понимания, изучения и использования, а также привлекательность для пользователя при использовании в указанных условиях.

Примечания: 1. Некоторые аспекты функциональных возможностей, надежности и эффективности будут влиять и на практичность, но для целей ИСО/МЭК 9126 не классифицированы как практичность.

2. В число пользователей могут быть включены операторы, конечные пользователи и косвенные пользователи, которые находятся под влиянием или зависят от использования ПО. В практичности следует учитывать всё разнообразие сред пользователей, на которые ПО может влиять, включая возможную подготовку к использованию и оценку результатов.

Понятность (Understandability) – способность ПП, обеспечивающая пользователю понимание, является ли ПО пригодным и как его можно использовать для конкретных задач и условий использования.

Примечание. Понятность зависит от документации и первичных впечатлений от ПО.

Изучаемость (Learnability) – способность ПП, обеспечивающая изучение пользователем его применения.

Примечание. Внутренние атрибуты, соответствующие пригодности для изучения, как определено в ИСО 9241-10.

Легкость использования (Operability) – способность ПП, обеспечивающая пользователю возможность его эксплуатировать и управлять им.

Примечания: 1. Аспекты изменяемости, адаптируемости и легкости установки могут быть предпосылками для легкости использования.

2. Легкость использования соответствует управляемости, устойчивости к ошибке и согласованности с ожиданиями пользователя, как определено в ИСО 9241-10.

3. Для системы, которая эксплуатируется пользователем, комбинация функциональных возможностей, надежности, практичности и эффективности можно измерить внешне через качество в использовании.

Привлекательность (Attractiveness) – способность ПП нравиться пользователю.

Примечание. Привлекательность связана с атрибутами ПО, обусловливающими большую привлекательность ПО для пользователя.

Согласованность (Compliance) – способность ПО придерживаться стандартов, соглашений, руководств по стилю или норм, связанных с практичностью.

Практичность есть совокупность свойств программного средства, характеризующая: усилия или легкость использования; управляемость и устойчивость к ошибке; согласованность с ожиданиями пользователя; привлекательность и способность нравиться пользователю.

Получается, что практичность – это некий набор показателей качества. Теперь необходимо рассмотреть методы оценки интерфейса, ГОСТ [3] предлагает несколько способов оценки программных средств, не регламентируя отдельно оценку интерфейса.

Методы оценки практичности по показателям качества

Методы определения показателей качества ПС различаются: по способам получения информации о ПС (измерительный, регистрационный, органолептический, расчетный); по источникам получения информации (традиционный, экспертный, социологический).

Измерительный метод основан на получении информации о свойствах и характеристиках ПС с использованием инструментальных средств. Например, метод применяют при определении объема ПС (число строк исходного текста программ и число строк – комментариев; число операторов и операндов; число исполненных операторов; число ветвей в программе; число точек входа (выхода); время выполнения ветви программы; время реакции и другие показатели).

Регистрационный метод основан на получении информации во время испытаний или функционирования ПС, когда регистрируются и подсчитываются определенные события, например, время и число сбоев и отказов, время передачи управления другим модулям, время начала и окончания работы.

Органолептический метод основан на использовании информации, получаемой в результате анализа восприятия органов чувств (зрения, слуха), и применяется для определения таких показателей, как удобство применения, эффективность и т. п.

Расчетный метод основан на использовании теоретических и эмпирических зависимостей (на ранних этапах разработки), статистических данных, накапливаемых при испытаниях, эксплуатации и сопровождении ПС. При помощи расчетного метода определяются длительность и точность вычислений, время реакции, необходимые ресурсы.

Экспертный метод. Определение значений показателей качества ПС экспертным методом осуществляется группой компетентных экспертов-специалистов на базе их опыта и интуиции. Метод применяется в случаях, когда задача не может быть решена никаким другим способом или другие способы значительно более трудоемки. Рекомендован при определении показателей наглядности, полноты и доступности программной документации, легкости освоения, структурности.

Социологические методы основаны на обработке специальных анкет-вопросников.

Автором предлагается использовать для оценки интерфейса обученную нейронную сеть, которая смогла бы оценить такие параметры интерфейса, как наглядность элементов управления (распознаваемость), а также соответствие цветовой насыщенности и гаммы интерфейса. То есть это будет гибридный метод, сочетающий в себе черты практически всех методов, установленных ГОСТом.

С этой целью придется использовать несколько нейронных сетей. Две сети, отвечающие за цветовую насыщенность и гамму, будут давать эмоциональную оценку всему интерфейсу, а третья сеть будет определять скорость распознавания элементов.

Оценка эмоциональной составляющей. Необходимо создать нейронную сеть, которая в процессе обучения сможет выдавать оценку эмоциональному окрасу интерфейса. Очевидно, что оценка эмоциональной составляющей может быть получена только экспертом. Процесс обучения нейронной сети оценке эмоциональной составляющей человеко-машинного интерфейса очень похож на процесс обучения распознавания образов. Поэтому автор использовал однослойную нейронную сеть (персептрон) (рис. 1), которая применяется в подобных задачах.

Алгоритм обучения персептрона

Персептрон обучают, подавая множество образов по одному на его вход и подстраивая веса до тех пор, пока для всех образов не будет достигнут требуемый выход. Допустим, что входные образы нанесены на демонстрационные чернобелые карты. Каждая карта разбита на квадраты и от каждого квадрата на персептрон подается вход. Если в квадрате имеется линия, то от него подается единица, в противном случае – ноль. Множество квадратов на карте задает, таким образом, множество нулей и единиц, которое и подается на входы персептрона. Цель состоит в том, чтобы научить персептрон включать индикатор при подаче на него множества входов, задающих нечетное число, и не включать в случае четного.

Такая персептронная конфигурация показана на рис. 2. Допустим, что вектор **X** – образ распознаваемой демонстрационной карты. Каждая компонента (квадрат) $X-(x_1, x_2, ..., x_n)$ – умножается на соответствующую компоненту вектора весов $W-(w_1, w_2, ..., w_n)$. Эти произведения суммируются. Если сумма превышает порог Θ, то выход нейрона Y равен единице (индикатор зажигается), в противном случае он – ноль. Эта операция компактно записывается в векторной форме как $Y=XW$, а после нее следует пороговая операция.

Для обучения сети образ **X** подается на вход и вычисляется выход Y. Если Y правилен, то ничего не меняется. Однако если выход неправилен, то веса, присоединенные ко входам, усиливающим ошибочный результат, модифицируются, чтобы уменьшить ошибку.

Чтобы увидеть, как это осуществляется, допустим, что демонстрационная карта с цифрой три подана на вход, и выход Y равен единице (показывая нечетность). Так как это правильный ответ, то веса не изменяются. Если, однако, на вход подается карта с номером четыре и выход Y равен единице (нечетный), то веса, присоединенные к единичным входам, должны быть уменьшены, так как они стремятся дать неверный результат. Аналогично, если карта с номером три дает нулевой выход, то веса, присоединенные к единичным входам, должны быть увеличены, чтобы скорректировать ошибку.

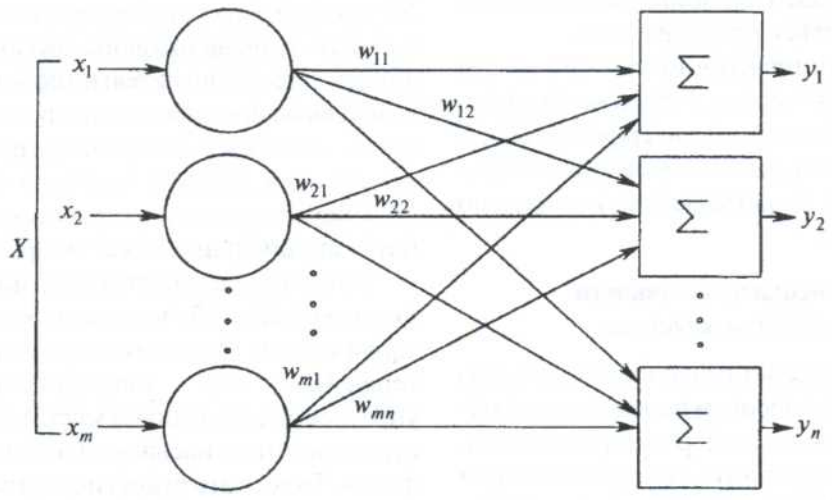


Рис. 1. Однослойная нейронная сеть (персептрон)

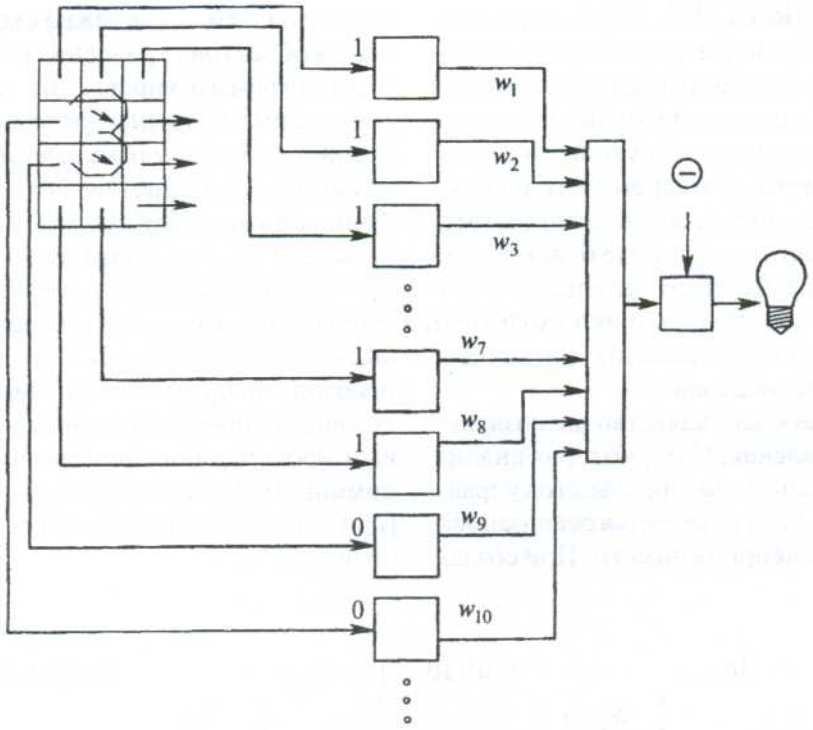


Рис. 2. Персептронная система распознавания изображений

Этот метод обучения может быть подытожен следующим образом:

1. Подать входной образ и вычислить Y.

2. а) Если выход правильный, то перейти на шаг 1; б) если выход неправильный и равен нулю, то добавить все входы к соответствующим им весам; в) если выход неправильный и равен единице, то вычесть каждый вход из соответствующего ему веса.

3. Перейти на шаг 1.

За конечное число шагов сеть научится разделять карты на четные и нечетные при условии, что множество цифр линейно разделимо. Это значит, что для всех нечетных карт выход будет больше порога, а для всех четных – меньше. Отметим, что это обучение глобально, т. е. сеть обучается на всем множестве карт. Возникает вопрос о том, как это множество должно предъявляться, чтобы минимизировать время обучения. Должны ли элементы множества предъявляться последовательно друг за другом или карты следует выбирать случайно? Несложная теория служит здесь путеводителем.

Дельта-правило. Важное обобщение алгоритма обучения персептрона, называемое дельта-правилом, переносит этот метод на непрерывные входы и выходы. Чтобы понять, как оно было получено, шаг 2 алгоритма обучения персептрона может быть сформулирован в обобщенной форме с помощью введения величины δ, которая равна разности между требуемым или целевым выходом T и реальным выходом Y: $\delta = (T - Y)$. Случай, когда $\delta = 0$, соответствует шагу 2а, когда выход правилен и в сети ничего не изменяется. Шаг 2б соответствует случаю $\delta > 0$, а шаг 2в – случаю $\delta < 0$.

В любом из этих случаев персептронный алгоритм обучения сохраняется, если δ умножается на величину каждого входа x_i и это произведение добавляется к соответствующему весу. С целью обобщения вводится коэффициент "скорости обучения" η, который умножается на δx_i, что позволяет управлять средней величиной изменения весов.

В алгебраической форме записи

$$\Delta_i = \eta \delta x_i;$$

$$w(n+1) = w(n) + \Delta_i,$$

где Δ_i – коррекция, связанная с i-м входом x_i; $w_i(n+1)$ – значение веса i после коррекции; $w_i(n)$ – значение веса i до коррекции.

Пример с черно-белыми картами был приведен для наглядности, на деле же в качестве входных сигналов будут подаваться значения различных

цветовых моделей (RGB, HSL, HSB). А на выходе будет числовое значение эмоционального окраса изображения. Различные цветовые модели выбраны потому, что представление яркости в модели HSL больше соответствует визуальному восприятию человека, в то время как нашему восприятию насыщенности более точно соответствует насыщенность в модели HSB. А это приводит к тому, что необходимо использовать две сети со сходной структурой, однако входные и выходные сигналы будут рассчитываться для различных цветовых моделей.

Сеть, анализирующая качество распознавания объектов управления. Сеть, которая анализирует качество распознавания объектов управления оператором, по сути является реализацией интерфейса в виде нейронной сети. При создании интерфейса, где важна скорость распознавания объектов управления (в первую очередь, это программы управления технологическим процессом), известны последовательность действий и временные рамки, и потому не составит труда сделать нейронную сеть, используя временные функции срабатывания. Последние показывали бы, на каком элементе управления произошла ошибка (т. е. время, затраченное на выполнение последовательности действий, превысило допустимое.). Для создания такой сети необходимо представить интерфейс и его работу в виде непрерывного автомата. Возможность имитировать работу непрерывного автомата при помощи нейронных сетей была рассмотрена в [2]. Общая структура такой сети представлена на рис. 3.

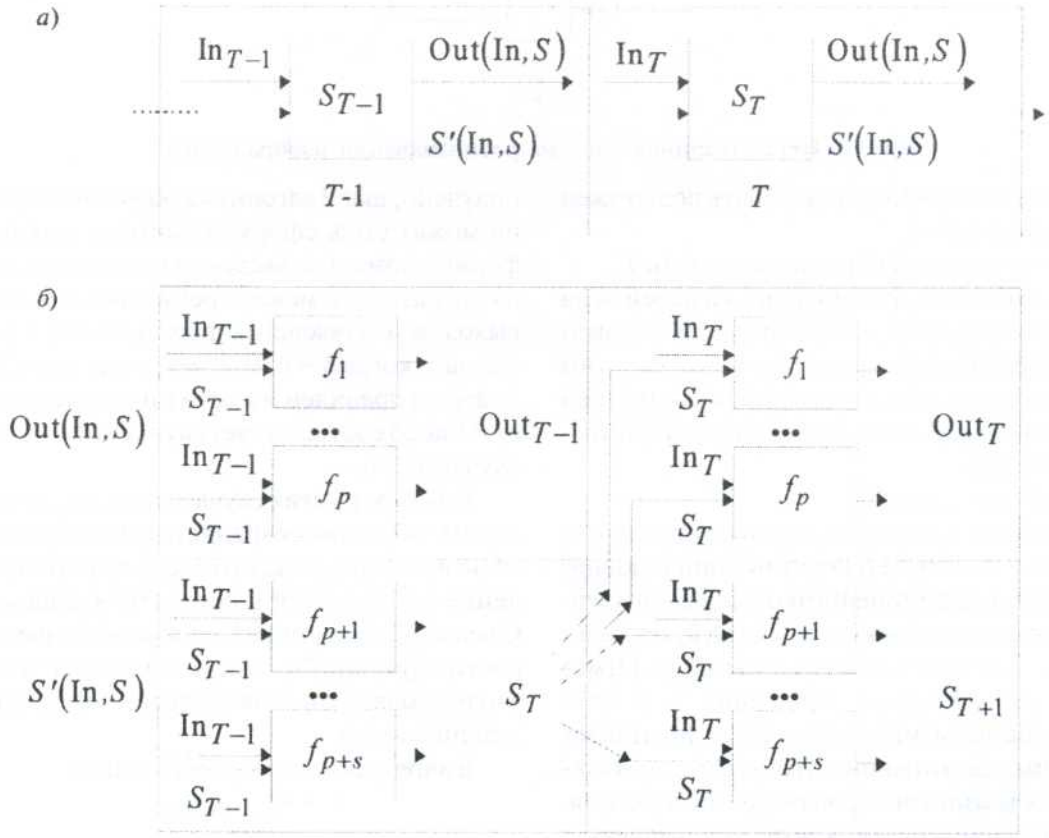


Рис. 3. Представление общего автомата с помощью модулей, вычисляющих функции многих переменных от входных сигналов: *а* – функционирование автомата; *б* – разложение его на отдельные модули

In, Out – входные и выходные сигналы; S – параметры состояния; T – дискретное время; Out (In, S) – зависимость выходных сигналов от значений входных и параметров состояния; S'(In, S) – зависимость состояния в следующий момент дискретного времени от входных сигналов и текущего состояния; $f_1 - f_p$ и $f_{p+1} - f_{p+s}$ – функции переменных (In, S) компонент вектора соответственно Out (In, S) и S'(In, S); T, $T \pm 1$ – индексы, соответствующие моментам времени

Сеть, предлагающая решения для изменения объектов управления. Данная сеть представляет собой однослойный персептрон с числом нейронов, равным числу способов изменения объектов управления интерфейса. На вход подается временная разность между эталонным временем выполнения задания на конкретном интерфейсе и реальным. Улучшение требуется, только если реально затраченное время больше эталонного. На выходе окажется вектор, который показывает преобразования, рекомендуемые для данного элемента. Для обучения сети будет использоваться метод обучения с подкреплением. В качестве среды, которая станет поощрять или наказывать, выступит нейронная сеть, имитирующая работу интерфейса. Таким образом, сеть будет постоянно обучаться. Если добавится несколько способов изменения объектов управления, то структура сети также должна измениться. Сигнал улучшения будет подаваться, если предложенное решение успешно, и напротив, если решение привело к замедлению времени распознавания, веса соответствующих нейронов будут уменьшены. Данный метод похож на метод обучения с учителем с той лишь разницей, что в качестве учителя выступает среда (нейронная сеть).

В статье рассмотрен метод оценки человеко-машинного интерфейса при помощи нескольких нейронных сетей, который позволяет более объективно оценить качество интерфейса.

СПИСОК ЛИТЕРАТУРЫ

1. Нейрокомпьютерная техника: Теория и практика / Ф. Уоссерман. / Пер. на рус. яз. Ю.А. Зуева, В.А. Точенова. 1992.

2. Нейроинформатика / А.Н. Горбань, В.Л. Дунин-Барковский, А.Н. Кирдин и др. Новосибирск: Наука; Сибирское предприятие РАН, 1998. 296 с.

3. ГОСТ Р ИСО/МЭК 9126-93. Информационная технология. Оценка программной продукции. Характеристики качества и руководства по их применению.

УДК 332.04:621

Д.Ю. Братин

ОСОБЕННОСТИ ЯЗЫКА FDL v. 1.0 ДЛЯ ОПИСАНИЯ ФРОНТЕНДА ВЕБ-ПРИЛОЖЕНИЙ

Веб-ресурс в течение своего жизненного цикла очень быстро видоизменяется и расширяется, а вместе с ним – и веб-приложение, обеспечивающее работоспособность ресурса. Также очень часто кроме программистов в работу включены верстальщики веб-страниц – специалисты по компоновке текстовых, графических и навигационных компонентов веб-страниц. Их работа заключается только в придании страницам ресурса нужного вида. Вследствие этого важнейшими задачами при разработке веб-приложения становится предоставление средств для отделения работы верстальщика от работы над программным кодом и обеспечение максимального повторного использования компонентов приложения при его расширении.

Проект FDI [1] представляет решение этих проблем путём разделения приложения на фронтенд и бэкенд. Бэкенд представляет собой фасад над реализацией бизнес-логики приложения и предоставляет доступ к его возможностям в виде команд, а фронтенд – интерпретатор для языка описания пользовательских интерфейсов, с которым работают верстальщики. В этой статье дано описание языка, используемого в FDI (FDL – Frontend Description Language).

FDL – это декларативный микроязык на базе XML. Он включает три подъязыка, предназначенных для описания составляющих частей фронтенда и интерфейса с бэкендом:

язык описания карты разделов ресурса;

язык описания шаблонов страниц;
язык описания поставщиков данных.

Навигация по ресурсу и его внешний вид формируются на базе документов, составленных на первых двух языках (карта сайта и шаблоны). На языке описания поставщиков данных даётся формальное описание API бэкенда, которое служит справочником для верстальщика и источником данных для контроля за корректностью взаимодействия фронтенда и бэкенда.

Схема поставщиков данных

Создание шаблонов страниц – основная работа верстальщика веб-страниц. По сути, она заключается в задании правил форматирования динамических данных. Другими словами – в задании правил подстановки значений динамических данных в значения элементов страницы. Очень редко данные, подлежащие подстановке в страницу, представляют собой простое строковое значение, чаще всего это целые таблицы значений. Чтобы отформатировать массив значений, верстальщику необходимо точное и, желательно, формальное описание его структуры.

Схема поставщиков данных (далее для краткости "схема ПД" или "схема поставщиков"), с одной стороны, – справочник для верстальщика по командам бэкенда и структуре данных, которые они возвращают, с другой – она декларирует программисту необходимый для реализации интерфейс. Также схема ПД снабжает интерпретатор данными для контроля за корректностью обращений к бэкенду и корректностью возвращаемых им данных.

В процесс шаблонизации (см. рисунок) вовлечены два интерпретатора: интерпретатор шаблонов страниц и интерпретатор схемы ПД. Стрелка от документа *Шаблон страницы* к документу *Схема поставщиков данных* означает, что в шаблоне страницы обращения к поставщикам данных оформляются в виде ссылок на сущности из схемы поставщиков.

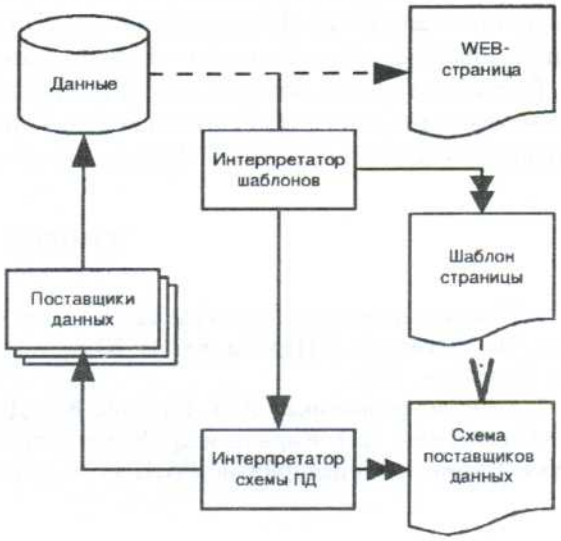


Схема процесса шаблонизации

Семантика и синтаксис. Документ схемы состоит из двух частей: описания моделей поставщиков данных и описания структур возвращаемых ими типов данных.

```
<dps:model name="news">
    <description>Новостная лента</description>
    <dps:request name="getList" uri="dp:cms/getList?type=news">
        <dps:parameter name="length" datatype="integer">5</dps:parameter>
        <dps:parameter name="page" datatype="integer">1</dps:parameter>
        <dps:parameter name="order" datatype="enum(asc,desc)">desc</dps:parameter>
        <dps:parameter name="order_field" datatype="date-time">date</dps:parameter>
        <dps:result datatype="newslist"/>
    </dps:request>
    <dps:request name="getItem" uri="dp:cms/getItem?type=news">
        <dps:parameter name="id" datatype="integer"/>
        <dps:result datatype="newsitem"/>
    </dps:request>
</dps:model>
```

Листинг 1: Пример модели в схеме ПД

Модель – условная структурная единица документа, объединяющая некоторые запросы в группу. Примером модели может быть группа запросов, предназначенная для работы с лентой новостей.

В листинге 1 приведён пример модели "Новостная лента", состоящей из двух запросов: "получить список новостей" (getList) и "получить новостную статью" (getItem). Для каждого запроса перечисляется набор его параметров с указанием типа и при необходимости – значения по умолчанию. Дополнительно указывается тип результата запроса с помощью элемента dps:result и его атрибута datatype:

```
<dps:data-type name="newslist">
        <dps:collection>
                <dps:object>
                        <dps:field name="id" datatype="integer"/>
                        <dps:field name="date" datatype="date-time"/>
                        <dps:field name="title" datatype="string"/>
                        <dps:field name="intro" datatype="string"/>
                </dps:object>
        </dps:collection>
</dps:data-type>
```

Листинг 2: Пример описания типа данных

В схеме ПД предусмотрены базовые типы данных, таких, как integer, float, string, datetime и пр. Сложные типы данных, которые возвращают запросы, описываются с помощью коллекций объектов и полей этих объектов. В листинге 2 приведён пример описания коллекции новостей, которые возвращает запрос getList из модели news. Запросы могут возвращать не только коллекции, но и одиночные объекты. В этом случае в описании типа не употребляется элемент dps:collection. Поля объектов также могут иметь сложный тип. В этом случае в атрибуте datatype указывается имя объявленного в этом же документе типа данных.

Адресация поставщиков данных и компонентов приложения

Все три типа документов, описывающих работу фронтенда (карта, шаблоны, схема ПД), содержат обращения к компонентам веб-приложения и к элементам друг друга. Например, шаблон содержит обращения к элементам схемы ПД. Как видно из примера описания модели в схеме поставщиков (см. листинг 1), запрос (элемент dps:request) в атрибуте uri содержит идентификатор компонента, который исполняет его. Данный идентификатор соответствует формату URI [2]. В этом формате задаются запросы ко всем компонентам веб-приложения.

Интерпретация URI-запросов происходит по схеме:
<имя компонента>:<запрос>

Схема интерпретации запроса не унифицирована и зависит от компонента. Для обращения из документов доступны следующие компоненты:

layout – вызов шаблона страницы
интерпретация: <файл шаблона>?<имя параметра>=<значение>&...;

env – доступ к переменным среды приложения
интерпретация: <пространство имён>[<имя переменной>];
доступные пространства имён:
request – переменные из HTTP запроса,
layout – переменные из запроса к шаблону,
section – атрибуты текущего раздела;

dp – доступ к поставщикам данных
интерпретация: <модель>/<команда>?<имя параметра>=<значение>&...;

dps – доступ к поставщикам данных посредством схемы поставщиков данных
интерпретация: <файл схемы ПД>?<модель>:-<запрос>;

index – доступ к индексу сущностей в текущем шаблоне
интерпретация:
<имя элемента>:<индекс>/<вл. элемент 1>/<вл. элемент 2>...;

context – доступ к родительскому элементу
интерпретация: ./<вл. элемент 1>/<вл. элемент 2>...;

nav – доступ к элементам карты ресурса интерпретация:
xpath:<выражение xpath>,
pxpath:<шаблон выражения xpath>;

file – доступ к файлам в специальной папке на локальном диске;

http – доступ к документам по HTTP-протоколу;

ftp – доступ к документам по FTP-протоколу.

Структура ресурса и шаблоны страниц

Основные задачи фронтенда – это интерпретация шаблонов страниц и обеспечение навигации по страницам веб-ресурса.

Навигация

```
<root name="Главная" id="1" layout="layout:default.xml?content=main">
  <main name="Главная" id="1" layout="layout:default.xml?content=main">
    <blog name="Блог разработчиков" id="2" layout="layout:default.xml?content=blog...">
      <article id="21" layout="layout:default.xml?content=article&amp;model=blog"/>
      <comment id="22" action="dps:cms.xml?cms:postComment" redirect="21?..."/>
    </blog>
    <news name="Новости проекта" id="3" layout="layout:default.xml?content=blog...">
      <article id="31" layout="layout:default.xml?content=articlelist..."/>
    </news>
  </main>
  <overview name="Концепция" id="4" layout="layout:default.xml?content=overview"/>
  <download name="Скачать" layout="layout:default.xml?content=download" id="7"/>
  <bagtrack name="Сообщить об ошибке" id="6" layout="layout:default.xml?content=bugtrack"/>

</root>
```

Листинг 3: Пример карты веб-ресурса

Навигация веб-ресурса формируется на базе карты – иерархической структуры разделов, описанной в виде XML-дерева. Каждый элемент карты (листинг 3) является разделом. Он имеет обязательные атрибуты id, name и layout. Id – это уникальный идентификатор раздела в документе карты. Name – имя раздела, используемое для формирования пунктов навигационных меню. Layout – ссылка на шаблон, по которому формируется страница раздела. В атрибуте layout чаще всего указывается запрос к интерпретатору шаблонов, как в примере. Но там также могут быть употреблены обращения к статичным XML/HTML файлам, доступным на локальном диске или по сети через http или ftp.

Разделы, указываемые в карте, не всегда служат для отображения страницы. Иногда они используются для выполнения какого-то интерактивного действия, например для добавления комментария в блоге. В этом случае вместо атрибута layout указываются два атрибута action, в который записывается запрос к схеме ПД, и redirect – переход в другой раздел после выполнения.

Использование шаблонов. Шаблоны являются описаниями правил подстановки динамических данных в страницу и их форматирования. Элемент, с помощью которого выполняется подстановка значений в документ, – это fdl:value. Он имеет один обязательный атрибут uri, в котором указывается URI идентификатор значения:
<title><fdl:value uri="env:section[name]"/></title>

В этом фрагменте шаблона в значение элемента title подставляется имя раздела, с которым связан шаблон.

```
<fdl:template applyto="index:request:mainmenu/collection/object">
<button class="menuitem">
<state><fdl:value uri="context:./state"/></state>
<link><fdl:link section="context:./id"/></link>
<caption><fdl:value uri="context:./name"/></caption>
</button>
</fdl:template>
```

Листинг 4: Обход коллекции объектов и форматирование их значений

```
<fdl:data-output>
<fdl:request index="mainmenu" uri="dps:website.xml?navigation:main"/>
<fdl:on-ok>
<fdl:template applyto="index:request:mainmenu/collection/object">
<button class="menuitem">
<state><fdl:value uri="context:./state"/></state>
<link><fdl:link section="context:./id"/></link>
<caption><fdl:value uri="context:./name"/></caption>
</button>
</fdl:template>
</fdl:on-ok>
<fdl:on-empty>
<header>Сайт одностраничный</header>
</fdl:on-empty>
<fdl:on-error>
<header>Произошла ошибка</header>
<p><fdl:value uri="index:request:sitemap/error/message"/></p>
</fdl:on-error>
</fdl:data-output>
```

Листинг 5: Пример формирования меню навигации

Для обхода коллекций объектов, которые могут возвращать команды бэкенда, предусмотрен элемент fdl:template. В его атрибуте applyto указывается ссылка на коллекцию объектов.

Коллекции объектов и объекты, содержащие динамические данные, поставляются в документ с помощью элемента fdl:request. В его обязательном атрибуте uri содержится ссылка на запрос, описываемый в схеме поставщиков данных. Также обязателен атрибут index, с его помощью fdl:template идентифицирует данные, подлежащие форматированию.

Для того чтобы можно было отслеживать результаты выполнения запроса к бэкенду, объявляемого элементом fdl:request, и решать, что подставить в страницу, существует элемент fld:data-output. Он имеет обязательные дочерние элементы fdl:request, fdl:ok-ok, fdl:on-empty, fdl:on-error.

Три элемента fdl:ok-ok, fdl:on-empty, fdl:on-error определяют фрагменты шаблона для случаев успешного выполнения, пустого результата и ошибки соответственно.

Чтобы обеспечить большую гибкость процессу отображения веб-страницы, её формирование производят в два этапа: строят промежуточную XML-страницу и преобразуют её посредством XSLT в HTML/XHTML страницу. Теги промежуточной страницы используют только для разметки блоков, на которые разбивается страница, и схематичного форматирования значений динамических данных.

В FDL нет спецификаций элементов промежуточной страницы. Их семантика и синтаксис полностью зависят от XSLT-шаблонов, которые индивидуальны для каждого проекта, поэтому формат промежуточной страницы оставлен открытым для разработки.

FDL по сути представляет собой язык описания пользовательских интерфейсов [3] наподобие UIML[4], XUL и т. д. Наиболее близким родственником является XML Sapience [5], так как он тоже ориентирован на создание интерфейсов веб-приложений. Главное отличие FDL от XML Sapience в том, что он оперирует в основном привычными для каждого веб-разработчика понятиями: карта сайта, шаблон страницы. Из новых понятий вводится только схема поставщиков данных, которое довольно просто для понимания. URL – вещь, знакомая всем пользователям Интернета, поэтому обращение к компонентам приложения в форме URI тоже не должно вызывать затруднений в освоении, так как URL – частный случай URI.

FDL обладает очень высоким уровнем визуализации устройства фронтенда, так как документы, управляющие его работой, для выполнения большинства возложенных на них задач ссылаются друг на друга: карта на шаблон, шаблон – на запросы из схемы ПД. Правила составления этих ссылок очень просты, мы можем убедиться в этом, посмотрев на список компонентов, доступных для обращения из XML документов.

Существует много исследовательских работ, посвящённых разработке пользовательских интерфейсов веб-приложений с помощью языков графического описания [6, 7]. В них приложение моделируется с помощью диаграмм состояний и переходов (класс диаграмм в языке UML). Преимущество такого подхода заключается в использовании широко распространённых стандартов моделирования и интуитивной понятности графической модели. Однако для него отсутствуют надёжные инструменты, позволяющие связать модель с реализацией и реализацию с моделью. В контексте разработки инструментов графического моделирования FDL можно рассматривать как связующее звено между графической моделью и реализацией. Графическая визуализация сущностей документов, связей между ними и возможность редактировать их даст нам инструмент для выполнения задач прямого и обратного моделирования.

При изучении FDL могут возникнуть проблемы с пониманием взаимодействия элементов fdl:request и fdl:template. К сожалению, это языковое средство ещё не очень хорошо проработано.

В целом, FDI сравнительно компактное и простое в изучении средство разработки визуальной части веб-ресурса.

СПИСОК ЛИТЕРАТУРЫ

1. Веб-сайт проекта FDI <http://fdi.bratin.ru>.
2. RFC 3986 Uniform Resource Identifier (URI) <http://tools.ietf.org/html/rfc3986>.
3. **Шейко Д.** Языки описания пользовательских интерфейсов. <http://www.codenet.ru/webmast/html/ui_langs.php>.
4. **Abrams S., Phanouriou C., Batongbacal A.L.** et al. UIML: An appliance independent XML user interface language // Proc. of the eighth Intern. world wide web conf. (WWW8). Toronto. May 1999.
5. Спецификации XML Sapience <http://xmlsapiens.org>.
6. **Горшкова Е.А., Новиков Б.А.** Использование диаграмм состояний и переходов для моделирования гипертекста // Программирование. 2004. № 1. С. 64–70.
7. **Turine M.A.S., de Oliveira M.C.F., Masiero P.C.** A statechart-based model for modeling hyper-media applications. ACM Transactions on Information Systems, 2001.

УДК 681.3

Т.М. Гильмуллин, И.В. Аникин

МОДЕЛИРОВАНИЕ ОБЪЕКТОВ ИНФОРМАЦИОННОЙ БЕЗОПАСНОСТИ ДЛЯ ЗАДАЧИ ОЦЕНКИ РИСКОВ

Обеспечение информационной безопасности (ИБ) информационных систем (ИС) часто осуществляется с позиции управления рисками ИБ [1].

Известен ряд специализированных методик оценки рисков ИБ, однако основной недостаток большинства их них – отсутствие формальных подходов, заложенных в их основу, что не позволяет исследовать их релевантность, оценить качество работы [5]. Это делает актуальным решение задачи разработки формальных моделей оценки рисков ИБ.

Постановка данной задачи заключается в разработке концептуальной модели предметной области (ПО) "Риск информационной безопасности", формальных моделей оценки рисков ИБ, а также моделирования тех объектов ИБ, которые участвуют в решении задачи оценки рисков. При решении задачи необходимо учесть некоторые особенности [1]:

задача оценки рисков ИБ трудно формализуема, а формирование оценок риска на практике основывается на экспертных оценках [3, 4];

знания эксперта о свойствах ИС и ее компонентах часто неполные и неточные (например, уровень критичности актива, уровень возможности реализации угроз). Формируемые модели должны быть "терпимы" к таким неопределенностям. В связи с этим актуальность приобретает использование при моделировании объектов ИС элементов теории нечетких множеств [2].

В настоящей статье предлагаются формальные модели для задачи оценки рисков ИБ, учитывающие данные особенности. В качестве подхода к формированию моделей использована теория алгебраических систем [6]. Кроме этого использованы элементы математического аппарата теории категорий и функторов, а также теории нечетких множеств и нечеткой логики [2, 3, 6].

Концептуальная модель предметной области

Каждый из объектов ПО имеет определенный тип $t \in Type$, где Type – множество всех типов. Отнесение объекта к некоторому типу может осуществляться экспертами при первом упоминании объекта, устанавливаться процедурно по его свойствам либо вследствие изменения состояния объекта.

Определив на Type множество отношений и операций над элементами, можно получить модель в виде формальной алгебраической системы, называемой концептуальной моделью предметной области. Впервые подход к синтезу концептуальной модели произвольной ПО был предложен П. Ченом при помощи построения так называемых ER-диаграмм типов (Entity Relationship Diagram) [6].

Введенная концептуальная модель ПО представлена на рисунке.

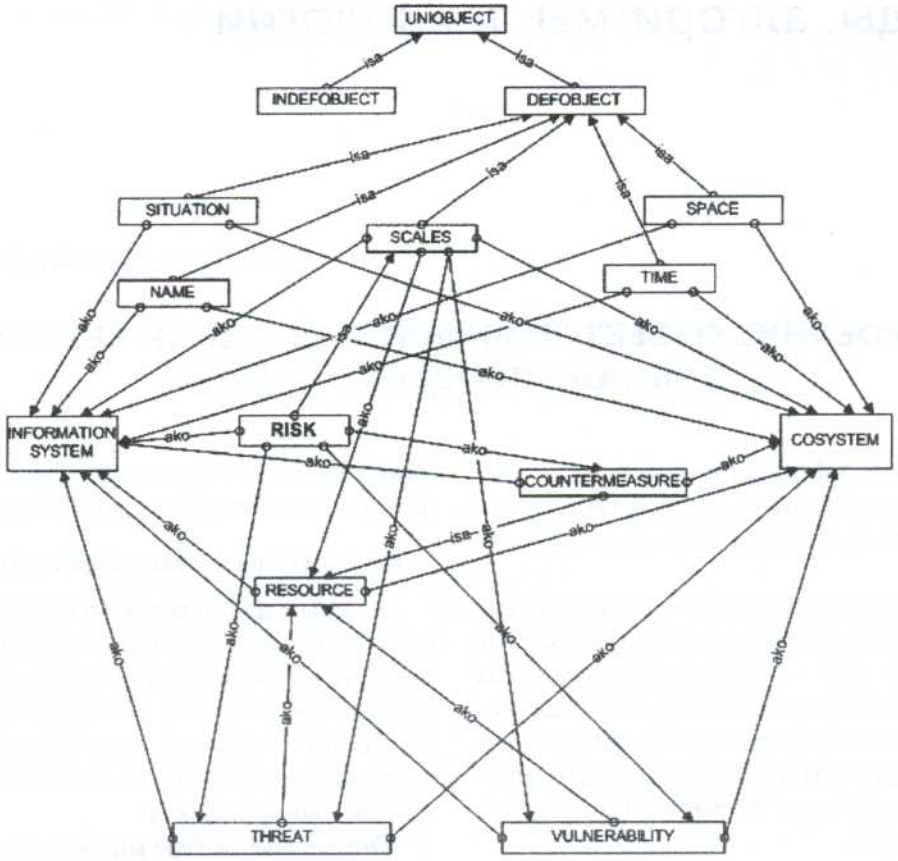
Введены следующие типы объектов $t \in Type$:

Базовый. t_U = UNIOBJECT – универсальный тип для любого объекта, который позволяет отличать в концептуальной модели ПО факт его существования.

Дополнительные типы. t_{Indef} = INDEFOBJECT – отнесение некоторого объекта ПО к типу t_{Indef} говорит о том, что в ПО ИБ "известно" о существовании объекта, но неизвестны его свойства.

t_{Def} = DEFOBJECT – отнесение некоторого объекта ПО к типу t_{Def} свидетельствует о том, что в ПО "известно" о существовании объекта, его свойствах и их значениях.

t_{Sit} = SITUATION – объекты ПО, которые выражают некоторые отношения между остальными



Концептуальная модель ПО "Риск информационной безопасности"

объектами в ПО, или операции над объектами. Именно с помощью объектов типа t_{Sit} выражаются состояния ПО.

t_{Name} = NAME – лингвистический тип t_{Name} позволяет описывать ПО ИБ на некотором естественном либо формальном языке.

t_{Time} = TIME, t_{Space} = SPACE – эти типы позволяют моделировать пространственно-временные отношения между объектами ПО.

Специальные типы. t_{Scale} = SCALE – измерительные шкалы, на которых осуществляется оценка базовых свойств объектов.

t_{IS} = INFORMATION SYSTEM – информационные системы.

t_{CoSys} = COSYSTEM – тип объектов ПО, дополнительный для t_{IS}: $t_{CoSys} = t_U \backslash t_{IS}$. Множество объектов COIS $\subseteq (t_U)$, COIS $= (t_U)\backslash\{IS\}$ называется косистемой для IS.

t_{Risk} = RISK – объекты ПО, являющиеся элементами пространства рисков RS = $[0, \infty)$.

t_{Res} = RESOURCE – ресурсы ИС.

t_T = THREAT – угрозы ИБ для ИС.

t_V = VULNERABILITY – уязвимости в ИС.

t_{CM} = COUNTERMEASURE – контрмеры для ИС.

К базовым отношениям на множества типов Type относят: =, isa, ako. Они – носители специальных отношений в ПО.

Отношение равенства (=) на множестве типов Type – это бинарное отношение, определенное через равенство множеств объектов данных типов:

$$\forall t_1, t_2 \in \text{Type}\ t_1 = t_2 \Leftrightarrow (t_1) = (t_2).$$

Отношение isa (is a – "относится к") на множестве типов Type – это отношение, интерпретируемое следующим образом: "Для любых типов t_1, $t_2 \in$ Type, t_1 isa t_2, если в любой момент времени каждый объект типа t_1 является объектом типа t_2":

$$\forall t_1, t_2 \in \text{Type}\ t_1 \text{ isa } t_2 \Leftrightarrow (t_1) \subseteq (t_2).$$

Отношение ako (a kind of – "является разновидностью") на множестве типов Type – это от-

ношение, интерпретируемое следующим образом: "Для любых типов $t_1, t_2 \in$ Type, t_1 ako t_2, если в определенный момент времени некоторые объекты типа t_2 могут состоять из некоторых объектов типа t_1":

$$\forall t_1, t_2 \in \text{Type}\ t_1 \text{ ako } t_2 \Leftrightarrow \exists o \in (t_2)\ \exists o_1, ..., o_k \in (t_1)\ o = \{o_1, ..., o_k\}.$$

Обозначим множество всех объектов ПО определенного типа t в конкретный момент времени через (t).

Каждый из представленных типов объектов имеет свойства, которые без ограничения общности будем считать нечеткими. Формальная модель каждого объекта obj в ИС представляется в виде

$$obj = \{Name(obj), fp(obj, bmf(obj, fp(obj)))\}, \quad (1)$$

где Name – наименование объекта; fp(obj) – вектор нечетких базовых характеристик объекта; bmf(obj, fp(obj)) – функция принадлежности, используемая для оценки свойств объекта.

Моделирование базовых объектов информационной безопасности

Тип INFORMATION SYSTEM. Модель объекта данного типа предлагается представить в виде

$$IS(E) = \{R, T, V, C, Risk, Eff, AllScales, Rel\}, \quad (2)$$

где $R \subset (t_{Res})$ – множество ресурсов ИС; $T \subset (t_T)$ – множество угроз для ИС; $V \subset (t_V)$ – множество уязвимостей ИС; $C \subset (t_{CM})$ – множество контрмер для ИС; Risk – функция оценки риска; Eff – функция оценки эффективности мер безопасности; AllScales – множество шкал для оценки свойств элементов ИС; Rel – отношения между элементами ИС.

Для работы с информационными системами определены операции равенства, объединения, пересечения и дополнения ИС и некоторые другие.

Тип RESOURCE. Ресурсом $t_{Res} \in$ Type называется тип объектов ИС, удовлетворяющий условию t_{Res} ako $t_{IS} \wedge t_{Res}$ ako $t_{CoSys} \wedge t_{Scale}$ ako t_{Res}. Каждый ресурс обладает некоторыми свойствами, определяемыми вектором нечетких базовых характеристик fp, которые измеряются на шкалах из множества AllScales.

Для оценки базовых характеристик элементов ИС введены следующие нечеткие измерительные шкалы типа t_{Scale}:

1. Уровень конфиденциальности (Confidential Level – L_C).
2. Уровень целостности (Integrity Level – L_I).
3. Уровень доступности (Availability Level – L_A).
4. Уровень критичности ресурса (Critical Level – CL).
5. Уровень воздействия элемента (Affect Level – AL).
6. Уровень простоты использования уязвимости (Easy Level – EL).
7. Уровень риска (Risk Level – RL).
8. Уровень ущерба для системы (Damage Level – DL).
9. Уровень возможности реализации угрозы (Opportunity Level – OL).
10. Уровень эффективности контрмеры (Effect Level – EffL).

Результат оценки свойств объекта формируется на нечеткой шкале базовых характеристик с помощью базовой функции принадлежности.

Нечеткой шкалой базовых характеристик (Fuzzy Scale of Basic Properties) для типа $t \in$ Type, $t \neq t_{Scale}$ называется упорядоченное нечеткое множество $FSBP(t) \in t_{Scale}$ векторов fp, координатами которых являются нечеткие базовые характеристики свойств – лингвистические нечеткие значения, характерные для объектов (t):

$$\forall t \in \text{Type}\ \ t \neq t_{Scale};$$
$$FSBP(t) \overset{def}{=} \{fp = (fp_i) \mid fp_i \in FP_i\ \ FP_i \in t_{Scale}\} \in t_{Scale}. \quad (3)$$

Базовой функцией принадлежности (Base Membership Function – bmf) для объектов e типа $t \in$ Type, $t \neq t_{Scale}$, называется функция принадлежности объектов $e \in (t)$ к векторам шкалы FSBP(t):

$$bmf: (t) \times FSBP(t) \rightarrow [0, 1]. \quad (4)$$

Тип VULNERABILITY. Уязвимостью $t_V \in$ Type называется тип объектов с измеримыми свойствами "Уровень воздействия элемента" и "Уровень простоты использования уязвимости", удовлетворяющие условиям:

$$t_V \text{ ako } t_{Res} \wedge t_{Scale} \text{ ako } t_V;$$
$$\forall th \in (t_T)\ \forall v \in (t_V)\ Risk(IS \cup (th, v)) > Risk(IS),$$

где (th, v) определяет пару, состоящую из угрозы th и реализующей ее уязвимости v.

Тип THREAT. Угрозой $t_T \in$ Type называется тип объектов с измеримыми свойствами "Потенциальная стоимость реализации угрозы на ресурс", "Предпочтение выбора угрозы по стоимости", "Уровень воздействия элемента", "Уровень

возможности реализации угрозы", удовлетворяющие условиям:

$$t_T \text{ ako } t_{Res} \wedge t_{Scale} \text{ ako } t_T;$$

$$\forall th \in (t_T)\ Risk(IS \cup th) \geq Risk(IS).$$

Тип RISK. Риском $t_{Risk} \in Type$ называется тип объектов, измеримый на одной из следующих шкал RS, где $RS = [0, \infty)$ – пространство рисков:

на непрерывной шкале рисков

$$RS_C \overset{def}{=} \{r \in (t_{Risk}) \mid \exists r_{high} \in RS \quad r \leq r_{high} < \infty\} \subset RS; \quad (5)$$

дискретной шкале рисков

$$RS_D \overset{def}{=} \{r \in (t_{Risk}) \mid r \in N \cup \{0\} \wedge \exists r_{high} \in N \cup \{0\} \quad r \leq r_{high} < \infty\} \subset RS; \quad (6)$$

на нечеткой шкале рисков

$$RS_F \overset{def}{=} \{r \in (t_{Risk}) \mid r \in [0,1]\} \subset RS. \quad (7)$$

В качестве отношений Rel между элементами ИС выбраны следующие:

1. Нечеткое отношение критичности ресурса (Resource Criticity Relationship – RCR) – $rcr \subset R \times IS$. Принадлежность пар $(r, IS) \in R \times IS$ к этому нечеткому подмножеству rcr задается на лингвистической шкале FP_{CR}: $\mu_{rcr} : R \times IS \to FP_{CR}$. Для определения критичности ресурса используется функция дефазификации $Critical(r) = Defuz(\mu_{rcr}(r, IS))$.

2. Нечеткое отношение индуцирования ущерба (Damage Induction Relationship – DIR) – $dir \subset R \times R$. Принадлежность пар $(r_i, r_j) \in R \times R$ к этому нечеткому подмножеству dir задается на лингвистической шкале FP_{Affect}: $\mu_{dir} : R \times R \to FP_{Affect}$. Для определения степени воздействия элемента на ИС используется функция дефазификации $Affect(r_i, r_j) = Defuz(\mu_{dir}(r_i, r_j))$. Через данное нечеткое отношение фактически задается трехдольный граф (уязвимость – угроза – ресурс), определяющий стандартную модель угроз ИС.

3. Нечеткое отношение простоты использования уязвимости (Vulnerability Easy Using Relationship – VEUR) – $veur \subset V \times T \times R$. Принадлежность троек $(v_i, t_j, r_k) \in V \times T \times R$ к этому нечеткому подмножеству veur задается на лингвистической шкале FP_{Easy}: $\mu_{veur} : V \times T \times R \to FP_{Easy}$. Для определения степени простоты использования уязвимости используется функция дефазификации $Easy(v_i, t_j, r_k) = Defuz(\mu_{veur}(v_i, t_j, r_k))$.

Оценка рисков ИБ

Функцией оценки риска информационной безопасности (Information Safety Risk Function) названо отображение, сопоставляющее любым объектам ПО ИБ типа t_{IS} уровень шкалы рисков $RS_F \in t_{Scale} : Risk : (t_{IS}) \to RS_F$.

Уровень риска $r \in RS_F : r = Risk(IS(E)) \in RS_F$ называется уровнем риска информационной безопасности (Information Safety Risk) для заданной ИС, или риском ИБ для ИС.

Число $s \in RS_F$, такое, что $s = 1 - r \in RS_F$, где $r \in RS_F$ – вычисленный уровень риска для некоторой IS(E), называется уровнем информационной безопасности (Information Safety Level) для заданной ИС.

Поставим вопрос о "близости" двух ИС в смысле близости их рисков.

Назовем квазиметрикой риска действительную функцию $ModIS : (t_{IS}) \times (t_{IS}) \to RS_F : \forall IS_1, IS_2 \in (t_{IS})$ $ModIS(IS_1, IS_2) \overset{def}{=} |Risk(IS_1) - Risk(IS_2)|$.

Рассмотрим квазиметрическое пространство информационных систем: $<(t_{IS}), ModIS>$ и поставим для него вопрос о полноте. Положим, что

$$\lim_{n \to \infty} IS_n(E) = IS \Leftrightarrow \forall \varepsilon > 0\ \exists N = N(\varepsilon) \forall n > N$$
$$ModIS(IS_n, IS) < \varepsilon.$$

Назовем последовательность информационных систем по квазиметрике риска фундаментальной, если выполняется условие

$$\forall \varepsilon > 0\ \exists N = N(\varepsilon)\ \forall n > N\ \forall m > N\ ModIS(IS_n, IS_m) < \varepsilon.$$

Число информационных систем IS_k, а также их рисков r_k конечно и линейно упорядочено. Если для некоторой IS $Risk(IS) = r$, и $r_k \leq r < r_{k+1}$, то будем считать, что $IS \cong IS_k$. При этом допущении на множестве (t_{IS}) можно сохранить аналог теоремы о полноте метрических пространств.

Доказаны следующие теоремы.

Теорема 1. Квазиметрическое пространство $<(t_{IS}), ModIS>$ является полным.

Теорема 2. Всякое сжимающее отображение $f : (t_{IS}) \to (t_{IS})$, определенное в полном квазиметрическом пространстве $<(t_{IS}), ModIS>$, имеет неподвижную точку IS, такую, что $f(IS) = IS$.

Следствие. Неподвижная точка для сжимающего отображения, заданного в полном квазиметрическом пространстве $<(t_{IS}), ModIS>$, не обязательно единственна.

Произвольную интерпретацию алгебраической системы < M; O; R >, где M – множество всех информационных систем и рисков, в общем случае $M = (t_{IS}) \cup (t_{Risk})$, O – множество операций на M, R – множество отношений на M, назовем моделью состояния информационной безопасности заданной информационной системы.

Моделирование контрмер в задаче оценки рисков ИБ

Пусть задана $IS(E_1) \in (t_{IS})$. Контрмерой (Countermeasure) для $IS(E_1)$ называется такая ИС $cm = IS(E_2) \in (t_{IS})$, объединение которой с $IS(E_1)$ уменьшает риск ИБ $IS(E_1)$, т. е. $cm \in (t_{CM}) \Leftrightarrow Risk(IS(E_1) \cup cm) < Risk(IS(E_1))$.

В рамках разработанных моделей доказаны следующие теоремы.

Теорема 3. Для любой ИС $IS(E) \in (t_{IS})$ существует единственный уровень риска $r_{min} \in SR_F$, такой, что для любой контрмеры $C \in CM(IS(E))$ уровень риска ИБ для ИС с введенной контрмерой не меньше r_{min}, т. е.

$$\forall IS(E) \in (t_{IS})\ \exists!\, r_{min} \in SR_F\ \forall C \in CM(IS(E))\quad Risk(IS(E) \cup C) \geq r_{min}. \qquad (8)$$

Из данной теоремы вытекает следствие о существовании контрмеры, минимизирующей риск.

Следствие. Для любой ИС $IS(E) \in (t_{IS})$ существует контрмера $C \in CM(IS(E))$, такая, что уровень риска ИБ для ИС с введенной контрмерой равен r_{min}:

$$\exists C \in CM(IS(E))\ Risk(IS(E) \cup C) = r_{min}. \qquad (9)$$

Для оценки эффективности контрмер, заданной ИС $IS(E) \in (t_{IS})$, введены следующие функции:

абсолютная оценка эффективности контрмеры, заданная через уменьшение уровня риска ИБ, определяемая в виде (10):

$$Eff_1(cm, IS) = \begin{cases} r_{old} - r_{new}, & r_{old} > r_{new} \\ 0, & r_{old} \leq r_{new} \end{cases}; \qquad (10)$$

относительная оценка эффективности, определяемая через изменение риска относительно старого уровня в виде (11):

$$Eff_2(cm, IS) = \begin{cases} \dfrac{r_{old} - r_{new}}{r_{old}}, & r_{old} > r_{new} \\ 0, & r_{old} \leq r_{new} \end{cases}. \qquad (11)$$

Оценка эффективности контрмер с точки зрения риска ИБ позволяет ставить и решать оптимизационные задачи по целенаправленному снижению уровня риска ИБ.

В статье предложен подход к моделированию объектов информационной безопасности для задачи оценки рисков ИБ.

Разработка формальных моделей в рамках теории алгебраических систем дала возможность доказать ряд фундаментальных теорем, позволяющих формально решать задачи управления рисками информационной безопасности.

Разработанные модели можно использовать для создания формальных методик оценки и управления рисками ИБ, в рамках которых будут предложены четко обоснованные шаги для снижения рисков.

СПИСОК ЛИТЕРАТУРЫ

1. **Аникин И.В.** Управление внутренними рисками информационной безопасности корпоративных информационных сетей // Научно-технические ведомости СПбГПУ. Сер. Информатика, телекоммуникации, управление. 2009. № 3 (80). С. 35–39.

2. **Аникин И.В., Аджели М.А., Глова В.И.** Мягкие вычисления и их приложения: Учеб. пособие. Казань: Изд-во Казан. гос. техн. ун-та, 2000. 64 с.

3. **Гильмуллин Т.М.** К вопросу о моделировании предметной области "Оценка, анализ и управление рисками в ИБ" // Проблемы техники и технологий телекоммуникаций ПТиТТ-2008: Тез. докл. 9-й междунар. науч.-техн. конф. Казань, 25–27 ноября 2008. Казань: Изд-во Казан. гос. техн. ун-та, 2008. С. 453–454.

4. **Гильмуллин Т.М.** Развитие методики оценки рисков от Digital Security // XVI Туполевские чтения. Междунар. молодежная науч. конф. 28–29 мая 2008: Тр. конф. Т. III. Казань: Изд-во Казан. гос. техн. ун-та, 2008. С. 89–90.

5. **Петренко С.А., Симонов С.В.** Управление информационными рисками. Экономически оправданная безопасность. М.: Компания АйТи; ДМК Пресс, 2004. 384 с.

6. **Цаленко М.Ш.** Моделирование семантики в базах данных. М.: Наука. Гл. ред. физ-мат. лит., 1989. 288 с.

УДК 65.001.575

А.С. Зраенко, К.А. Аксенов, Ван Кай

КОАЛИЦИОННАЯ МОДЕЛЬ МУЛЬТИАГЕНТНОГО ПРОЦЕССА ПРЕОБРАЗОВАНИЯ РЕСУРСОВ

Системы поддержки принятия решений

На многих современных предприятиях и в организациях (производственных, образовательных, IT-сферы и др.) существует необходимость проведения быстрого анализа ситуации на рынке и принятия решения в короткие сроки. Чтобы минимизировать риски, связанные с неточностью и необоснованностью решений, разрабатываются системы поддержки принятия решений (СППР).

Основная функция СППР – генерация различных альтернатив решений для определенной ситуации, возникшей в конкретном процессе преобразования ресурсов (ППР). Под ППР будем понимать некоторую совокупность действий (операций), потребляющих ресурсы и приводящих к значащему результату – продукту процесса [1].

Трудность понимания причинно-следственных зависимостей в сложных системах приводит к их неэффективной организации, ошибкам в проектировании, большим затратам на устранение ошибок. Поэтому использование средств моделирования часто необходимо при поддержке принятия ответственных управленческих решений [8].

Мультиагентный подход в системах поддержки принятия решений

Среди методов моделирования наиболее полно проблемной области ППР соответствует мультиагентное моделирование. Используется оно для исследования децентрализованных систем, динамика функционирования которых определяется не глобальными правилами и законами, а наоборот, эти глобальные правила и законы являются результатом индивидуальной активности агентов системы [2].

Понятие агент соответствует аппаратно или программно реализованной сущности, которая способна действовать в интересах достижения целей, поставленных перед ним владельцем и/или пользователем [3]. Под коалицией будем понимать агентную структуру, представляющую собой объединение некоторого числа агентов в сообщество, основанное на сонаправленности их доминирующих целей. В качестве мультиагентного процесса преобразования ресурсов (МППР) будем рассматривать ППР в мультиагентной системе (МАС), управляемый агентом или коалицией [7]. Структура МППР может использоваться при описании большинства реально существующих процессов: производственных, бизнес-процессов, организационно-технических, процессов логистики и др. Элемент (компонент) ППР или весь процесс можно представить в виде структуры, включающей: вход (ресурсы), средства (необходимые для преобразования), преобразование, выход (результаты преобразования) [5]. Агент или коалиция может управлять преобразованием в МППР и определять алгоритмы использования ресурсов и средств в зависимости от специфики конкретного процесса и текущей ситуации.

Математическая модель МППР

Основой для создания коалиционной модели МППР стала математическая модель, разработанная К.А. Аксеновым и Н.В. Гончаровой [5]. Для реализации коалиционной модели МППР необходимо дополнить модель [5] такими объектами, как: коалиция (K), база знаний коалиции (KB_K), цель коалиции (G_K), общая база знаний МАС (GKB), действие агента (D_A) и коалиции (D_K), план действий агента (P_A) и коалиции (P_K), жизненный цикл агента (L_A) и коалиции (L_K). В новой модели необходимо функционирование следующих механизмов: обмена сообщениями между агентами и коалициями, формирования коалиций, формирования планов действий агента и коалиции, проведения аукционов, согласования решений, организации досок с объявлениями. Основные объекты коалиционной модели МППР представлены на рис. 1.

В коалиционной модели МППР функционируют процессы, представленные на рис. 2.

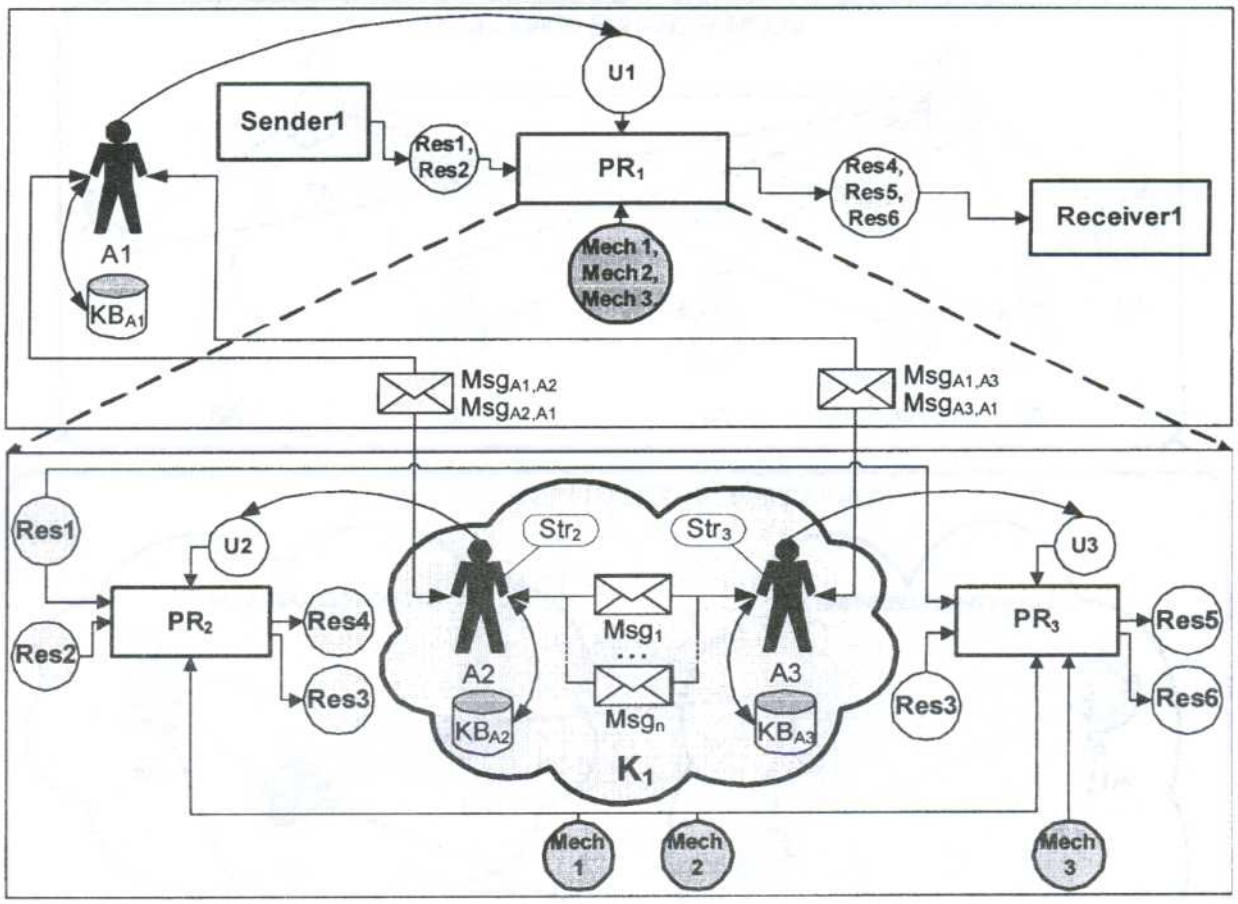


Рис. 1. Объекты коалиционной модели МППР

Для описания коалиционной модели МППР введены некоторые основные понятия. Коалиция агентов имеет следующую структуру:

$$K = <Name_K, \{A_1, ..., A_m\}, G_K, \{Str_1, ..., Str_v\}, KB_K>, (1)$$

где $Name_K$ – имя коалиции; $\{A_1, ..., A_m\}$ – множество агентов, входящих в коалицию; G_K – цель коалиции; $\{Str_1, ..., Str_v\}$ – множество допустимых стратегий поведения коалиции; KB_K – база знаний коалиции.

Плану действий P_K коалиции K соответствует структура

$$P_K = < \{D_{K1}, ..., D_{Kp}\}, G_K, KB_K, \{AD_1, ..., AD_h\} >, (2)$$

где $\{D_{K1}, ..., D_{Kp}\}$ – действия коалиции; $\{AD_1, ..., AD_h\}$ – множество агентов коалиции K, участвующих в данном плане действий.

Алгоритм мультиагентного моделирования, разработанный для практической реализации коалиционной модели МППР, представлен на рис. 3. В качестве его основы использован алгоритм, описанный в [5], включающий следующие основные этапы: определение текущего момента времени; диагностирование возникших ситуаций, выработка команд управления, формирование очереди правил преобразования; выполнение правил преобразования и изменение состояния рабочей памяти (данных по загрузке ресурсов и средств).

Данный алгоритм дополнен возможностью формирования коалиций, механизмом составления планов действий активных объектов (агентов и коалиций) и расширенными возможностями коммуникаций (для реализации взаимодействия активных объектов и формирования совместных решений агентов в коалиции). При выполнении 5-го и 9-го блоков алгоритма происходит формирование перечня правил и выполняемых действий над объектами рабочей памяти, а также состояния характеристик объектов. При выполнении 11-го блока выполняется расчет целей, производных и консолидированных показателей для объектов модели.

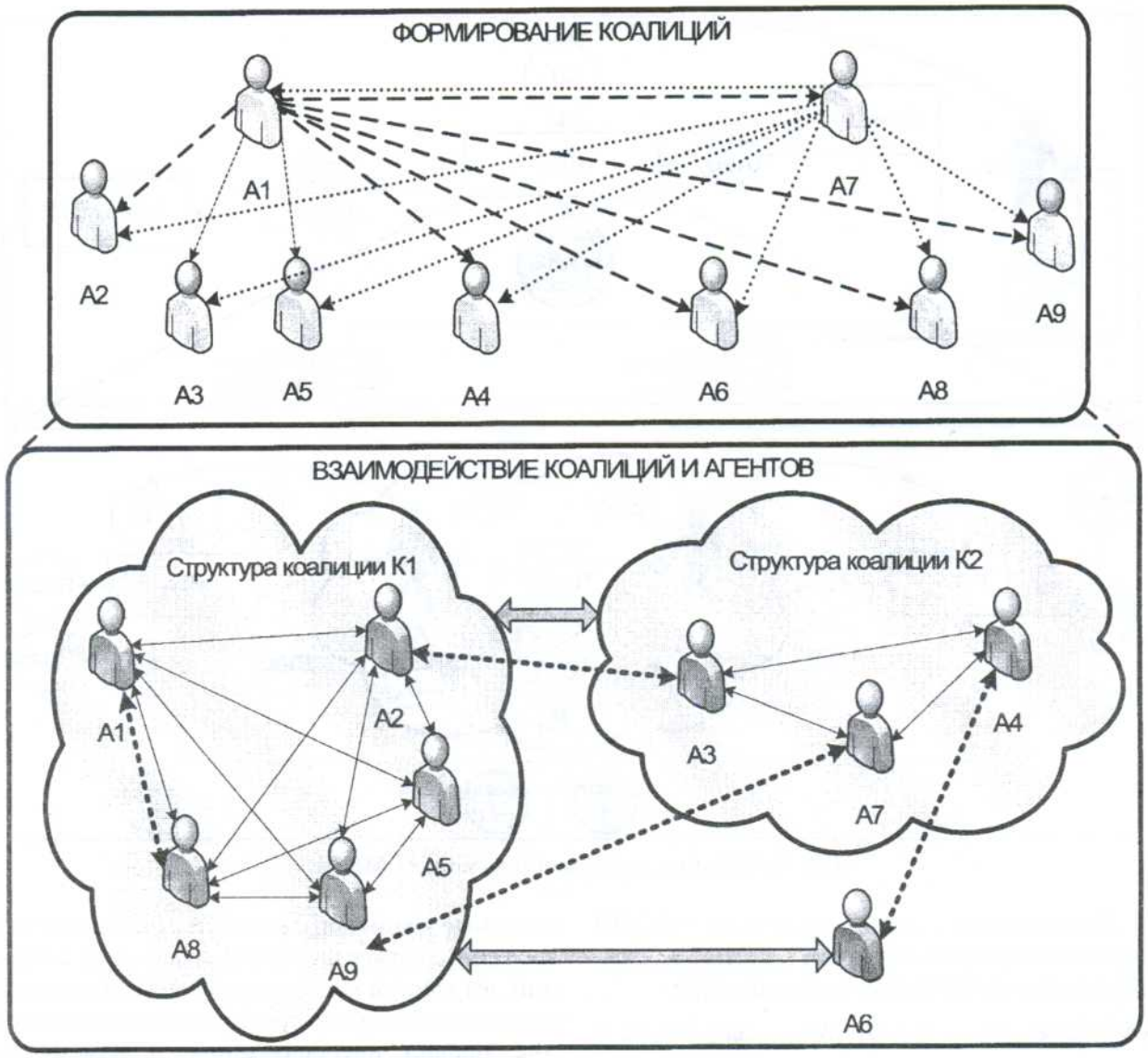


Рис. 2. Схемы процессов в коалиционной модели МППР

Процессы формирования коалиции К1 (◆——→) и коалиции К2 (◆·····→); процессы взаимодействия агентов внутри коалиции, направленные на достижение их общей цели (←——→); процессы взаимодействия коалиций между собой и с отдельными агентами (⟺); процессы взаимодействия агентов, направленные на достижение их второстепенных целей (◆·····→)

Программный комплекс поддержки принятия решений

Одним из проектов, разработанных для создания мультиагентных моделей процессов преобразования ресурсов [4], стала система динамического моделирования ситуаций (СДМС) BPsim.MAS. Библиотека коммуникаций агентов BPsim.LIB [6] разработана для расширения возможностей BPsim.MAS и представляет собой программный модуль, практически реализующий коалиционную модель МППР.

Использование библиотеки BPsim.LIB обеспечивает возможности формирования сообществ агентов (коалиций), планирования действий агентов (и коалиций), применения различных сценариев взаимодействий (на основе использования стратегий поведения и проведения различного вида аукционов) и формирования решений агентов (на основе механизма согласования решений и организации досок с объявлениями) при моделировании. Графический интерфейс библиотеки представлен на рис. 4 на примере

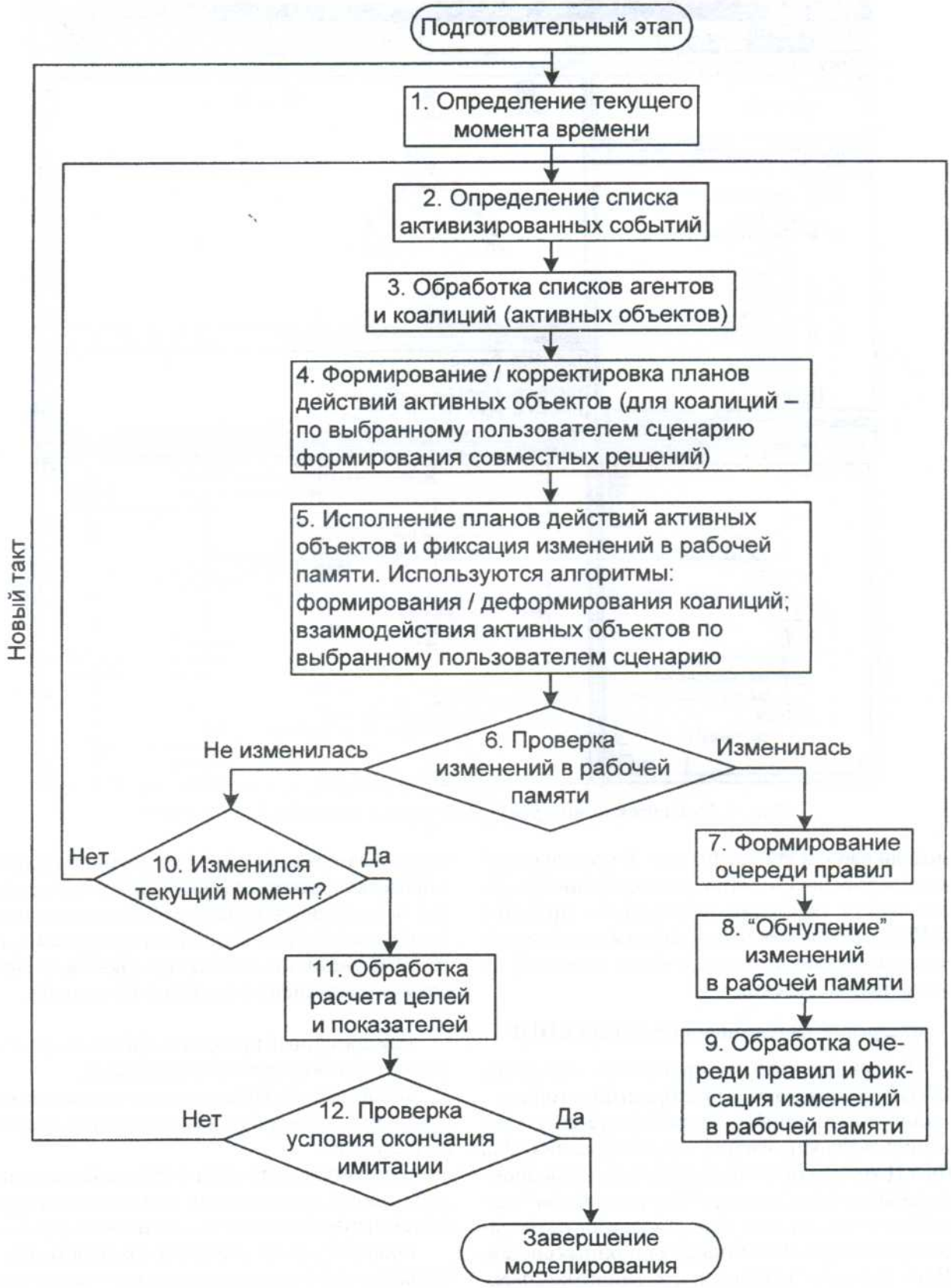


Рис. 3. Алгоритм мультиагентного моделирования

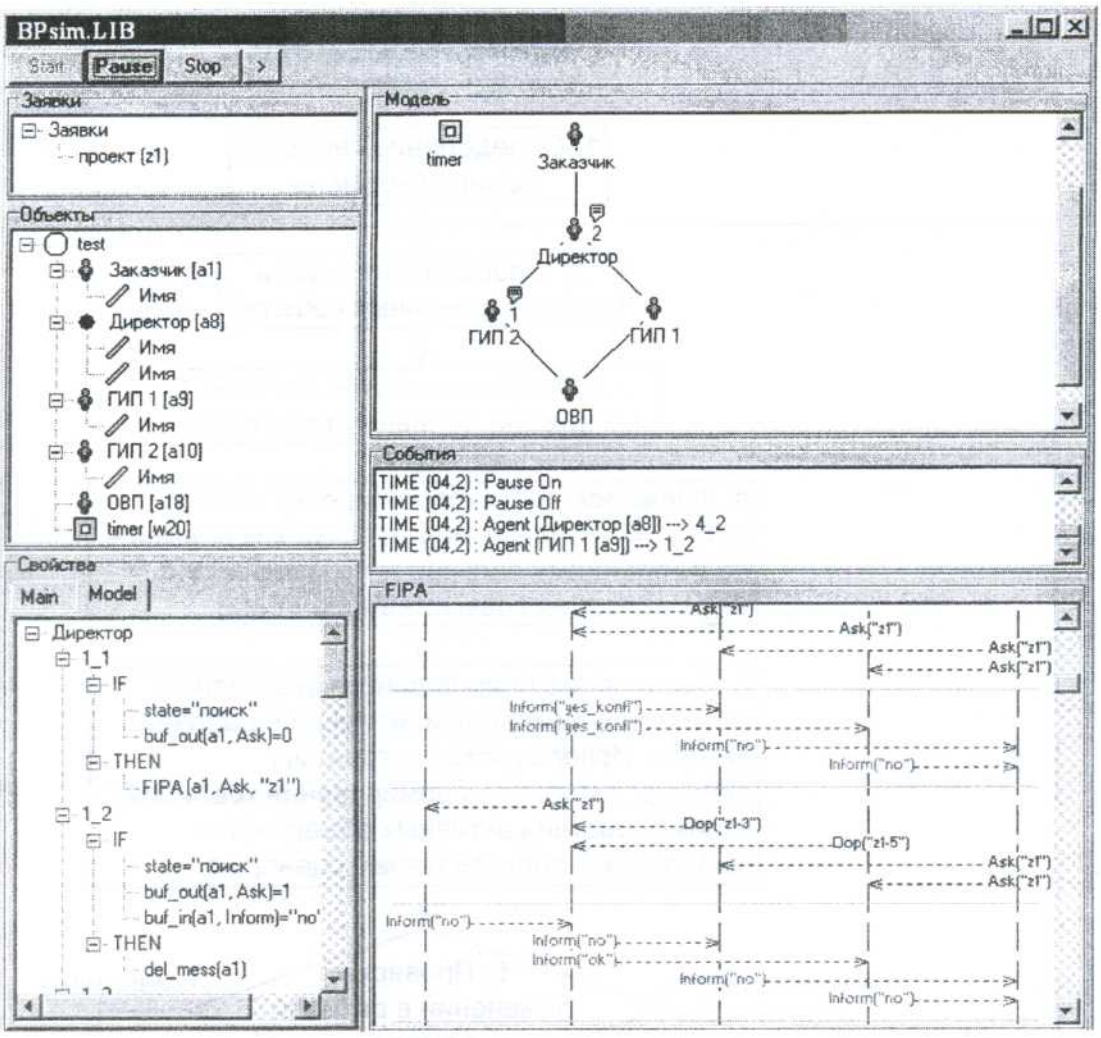


Рис. 4. Графический интерфейс библиотеки коммуникаций агентов

модели работы ИТ-компании. В модели решаются задачи оптимизации распределения ресурсов между главными инженерами проектов (ГИП) с использованием различных механизмов взаимодействия и планирования действий по выполнению проекта.

Внедрения программного комплекса СППР

Разработанный программный комплекс СППР внедрен на ряде предприятий и в организациях г. Екатеринбурга и используется для решения задач управления и планирования. Для ФГУП Уральский региональный информационно-аналитический центр "Уралгеоинформ" разработана модель поступления заказов на разработку и внедрение геоинформационных систем. Использование данного программного комплекса позволило: определить динамику поступления заказов, издержек и прибыли предприятия; сформировать стратегию главного инженера проектов, позволяющую наиболее оптимально управлять имеющимися кадровыми ресурсами; на 20–25 % увеличить количество заказов, выполняемых на предприятии в единицу времени.

Разработанный программный комплекс поддержки принятия решений позволяет:

вырабатывать эффективные управленческие решения на предприятиях и в организациях различного типа;

создавать модели процессов преобразования ресурсов с использованием возможности разработки сценариев поведения агентов и коалиций;

проводить имитационные эксперименты с использованием механизма составления планов действий агентов и коалиций.

СПИСОК ЛИТЕРАТУРЫ

1. **Аксенов К.А.** Развитие аппарата прерываний процессов преобразования ресурсов // На передовых рубежах науки и инженерного творчества: Тр. третьей междунар. науч.-практ. конф. Регионального Уральского отделения Академии инженерных наук им. А.М. Прохорова. Вестник УГТУ – УПИ / ГОУ ВПО УГТУ – УПИ. Екатеринбург, 2004. № 15 (45). Ч. 1. С. 187–190.

2. **Карпов Ю.Г.** Имитационное моделирование систем. Введение в моделирование с AnyLogic 5. СПб.: БХВ-Петербург, 2005. 400 с.: ил.

3. **Андрейчиков А.В., Андрейчикова О.Н.** Интеллектуальные информационные системы: Учебник. М.: Финансы и статистика, 2004. 424 с.: ил.

4. **Аксенов К.А., Клебанов Б.И.** Принципы построения системы имитационного моделирования процессов преобразования ресурсов BPsim2 // Матер. первой Всерос. науч.-практ. конф. "Опыт практического применения языков и программных систем имитационного моделирования в промышленности и прикладных разработках": Сб. докл. / ФГУП ЦНИИ технологии судостроения. СПб., 2003. Т. 1. С. 36–40.

5. **Аксенов К.А., Гончарова Н.В.** Динамическое моделирование мультиагентных процессов преобразования ресурсов: Монография / ГОУ ВПО УГТУ – УПИ. Екатеринбург, 2006. 311 с.

6. **Аксёнов К.А., Зраенко А.С.** Разработка языка коммуникации агентов для мультиагентной системы моделирования процессов преобразования ресурсов // Журнал научных публикаций аспирантов и докторантов. Курск: ООО "Редакция журнала научных публикаций аспирантов и докторантов", 2008. № 5. С. 134–136.

7. **Зраенко А.С., Аксенов К.А.** Конфликтные ситуации в интеллектуальных мультиагентных процессах преобразования ресурсов // Исследование, разработка и применение высоких технологий в промышленности: Сб. тр. второй междунар. науч.-практ. конф. СПб.: Изд-во СПбГПУ, 2006. № 5. С. 24–26.

8. **Зраенко А.С., Коновалова С.А.** Применение мультиагентного подхода в системах поддержки принятия решений // Науч. тр. междунар. науч.-практ. конф. "СВЯЗЬ-ПРОМ 2007" в рамках 4-го Евро-Азиатского форума "СВЯЗЬ-ПРОМЭКСПО 2007". Екатеринбург: ЗАО "Компания Реал-Медиа", 2007. С. 214–217.

УДК 33:681.324

М.Ю. Арзуманян

МОДЕЛИ ОПИСАНИЯ ПРЕДПРИЯТИЙ И ИХ МОНИТОРИРУЕМЫЕ СВОЙСТВА

Главный структурный компонент нового информационного общества – сетевая информационная экономика, представляющая глобальную сеть рынков, основанных на информационных технологиях. Сегодня трудно говорить об экономике в отрыве от сетевой инфраструктуры, телекоммуникаций и информационных технологий. Основным субъектом экономики является предприятие. По тенденции функционирования и развития предприятий во времени можно судить о том, что предприятия и организации различных секторов экономики в своей деятельности все больше используют инфокоммуникационные системы. Многие предоставляют информацию о себе посредством электронных ресурсов, другие поддерживают значительную часть своей деятельности за счет телекоммуникационных систем и информационных технологий, а иные имеют уже целый ряд технологических и бизнес-процессов в области "электронного бизнеса", протекающих частично или полностью в электронном виде. Существует несколько общепризнанных причин, заставляющих компании развивать электронный бизнес. Прежде всего это возможность предоставления таких новых услуг, которые в традиционном бизнесе организовать технологически невозможно или нецелесообразно из-за высокой стоимости. Вторая причина – желание расширить рынок и привлечь новых клиентов. Третья причина – возможность динамичного оперативного контроля реакции потребителей на предоставленные ему товары и услуги.

В настоящее время опубликовано большое количество работ, в которых предпринимаются

попытки сформулировать главные особенности влияния сетевых интернет-технологий на экономику [1].

В отчете Европейской комиссии по телеработе дается следующее определение сетевой экономики [1]: "Сетевая экономика представляет собой сетевую системно организованную пространственную структуру взаимоотношений между экономическими субъектами, включающую в себя индустрию создания и использования новых информационных технологий и продуктов, телекоммуникационные услуги, электронный бизнес, электронную коммерцию, электронные рынки и другие составляющие компоненты".

Объект исследования данной работы – модели преимущественно электронного предприятия, функционирующего в сетевой среде и использующего инфокоммуникационные технологии.

Предприятие относится к разряду сложных (больших) систем, которые можно рассматривать под различными углами зрения. Такие системы не поддаются полному, исчерпывающему описанию и проявляют свойства, не вытекающие непосредственно из свойств, составляющих систему элементов. Кроме того, предприятия относятся к классу целенаправленных систем, стремящихся к целесообразному поведению и достижению наиболее предпочтительных состояний. Существуют разные подходы и методики для исследования предприятий. Их можно условно разделить на классы по объекту, на котором они фокусируются. Такими объектами внутри предприятия традиционно являются организационная структура предприятия, функциональная структура предприятия и протекающие в нем бизнес-процессы (рис. 1, 2). Каждому из вариантов описания и представления предприятия присущи свои особенности и свойства.

Структурный подход. Описание организационной структуры предприятия – это полностью статическое представление, ограниченное описанием ее элементов, таких, как отделы предприятия, и указанием вида их взаимодействия. В основе подобного представления лежит принцип Адама Смита, изложенный им в работе "Достояние народа" еще в конце XVIII века. Такое описание не дает возможности проследить взаимодействия, происходящие как внутри предприятия, так и с внешними системами.

Функциональный подход. Другой подход к описанию предприятия связан с его функциональностью. Функция системы проявляется в ее поведении, при функциональном подходе изучается не то, из чего состоит система, а то, как она себя ведет. Описание предприятия через его функциональную структуру предполагает разделение предприятия на функциональные составляющие. Долгое время модель, основанная на этом описании, использовалась для управления предприятием. Подобный подход позволяет декомпозировать деятельность по функциональному принципу и, обеспечив соответствующими стандартами согласование между функциями, осуществлять осознанное управле-

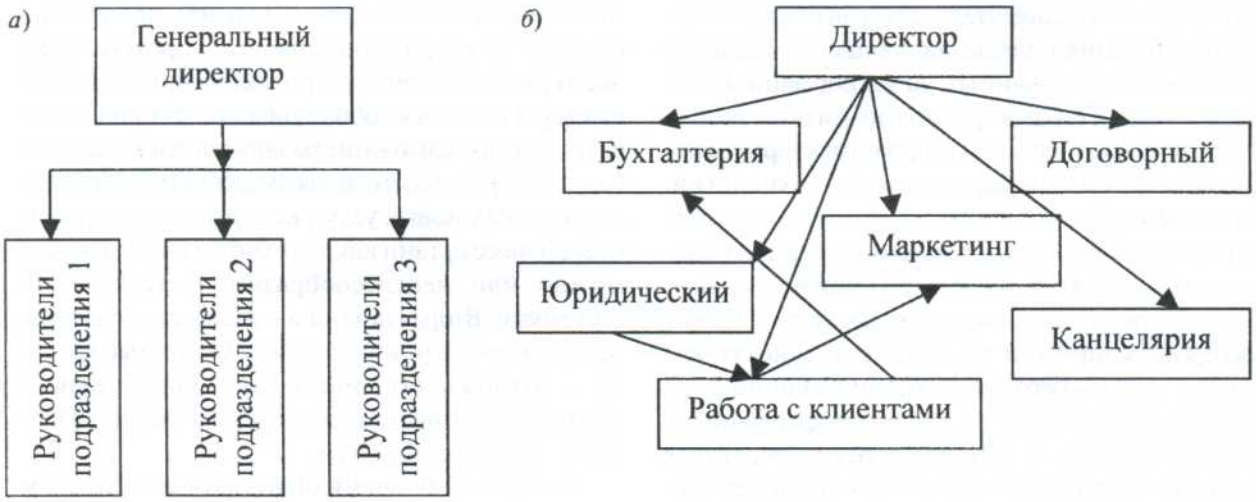


Рис. 1. Структурный (*а*) и функциональный (*б*) подходы

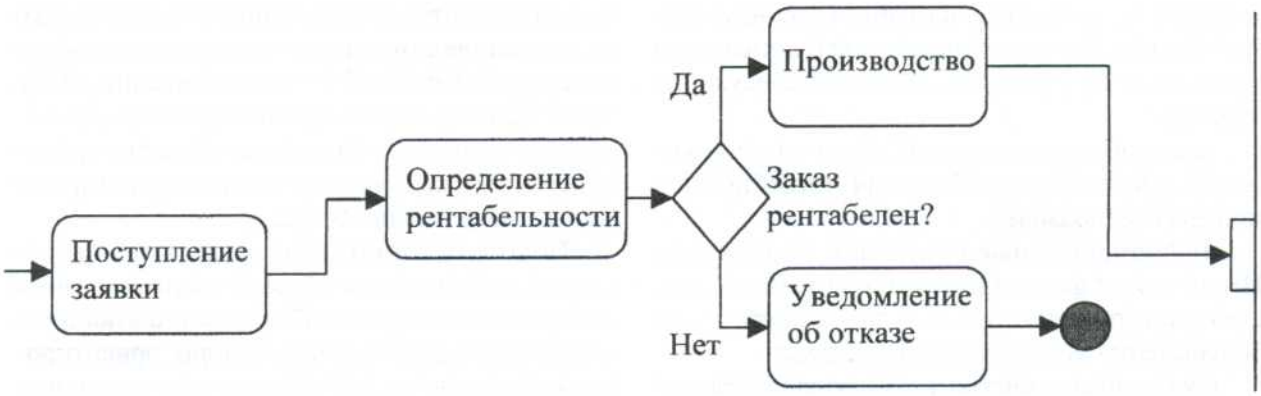


Рис. 2. Процессный подход

ние. Десятки лет назад, когда вычислительные средства поддержки информационной деятельности не были доступны, существование функционально-ориентированного подхода к управлению было не только оправданным, но и единственным возможным решением в управлении сложными системами. Эффективная реализация фунционально-ориентированного подхода требует постоянного и усиленного контроля и управления персоналом, поскольку реальные рабочие процессы предприятия состоят из множества функций и выходят за рамки отдельных подразделений. Рост межфункционального, технологически не обоснованного информационного обмена приводит к неоправданному увеличению сроков выработки управленческих решений и в конечном итоге к большим накладным расходам.

Бизнес-процессы. Реальная деятельность предприятия не осуществляется в рамках линейно-функционального описания, она связывает реализации функций в целостный бизнес-процесс. Подход, основанный на описании бизнес-процессов, приобретает все большую актуальность. Одна из основных причин, диктующая необходимость перехода на процессную ориентацию управления и описания предприятия, заключается в изменившихся способах и методах поддержки принятия решения и доступности информационных средств поддержки деятельности. При использовании информационных средств возникает возможность охватить всю систему целиком, со всеми процессами, и посмотреть на нее как на единое целое, введя контур исполнительного управления на передний план. Переход на процессную ориентацию, по мнению профессора А.В. Шера [2], это та неизбежность, с которой придется столкнуться практически каждому среднему или крупному предприятию, если оно хочет выжить и успешно развиваться. Именно этот подход позволяет оптимизировать деятельность предприятия при помощи современных инфокоммуникационных средств и технологий. Процессный подход позволяет осуществлять анализ, реинжиниринг, оптимизацию и автоматизацию бизнес-процессов, ведущую к увеличению эффективности деятельности предприятия (рис. 2).

Межфункциональные ключевые процессы определяют конкурентоспособность предприятия и взаимодействуют как с клиентами, так и с поставщиками. Организация бизнес-процессов предполагает тщательный анализ синергетических взаимодействий функций в рамках единой процедуры. Процессы можно условно делить на продуктивные (основные) и вспомогательные. Продуктивные бизнес-процессы направлены на изготовление продукции (услуги), имеющей ценность и представляющей интерес для покупателя (потребителя).

Анализ процессов предприятия позволяет глубже познать механизмы производства товара/услуги, выявить зависимости параметров на выходе от внутренних характеристик. Знание структуры процессов обеспечивает возможность управления путем воздействия на параметры процессов в соответствии с целевой функцией либо иной тактической задачей. Такие возможности недоступны в модели организационно-функционального представления.

Изменение моделей происходило вследствие появления новых требований, диктуемых ситуацией

в обществе, экономике, на рынке и в конкурентной борьбе. На эволюцию моделей описания и представления предприятий влияли следующие факторы:

масштаб предприятия. С увеличением масштаба появлялась все большая потребность в процессном подходе;

информационные средства и технологии. Именно этот фактор сделал возможным и, как следствие, привлекательным для конкурентной борьбы переход на процессный подход;

необходимость четкого контроля и управления временными характеристиками;

требование отображения динамических характеристик, вызванное необходимостью адаптации к быстрым изменениям внешней среды путем перестройки, пересмотра и реорганизации деятельности предприятия.

Каждая последующая из приведенных моделей становится все более адаптивной для наилучшего отображения именно динамических свойств и характеристик предприятий, продиктованных внешней ситуацией. При описании БП предприятие рассматривается не как статическая конструкция, а как процесс.

Еще одна важная особенность бизнес-процессов – это то, что они есть по сути наборы взаимодействий. Понятие взаимодействия – одно из ключевых, именно оно будет рассматриваться далее более подробно.

Описание и представление предприятия через бизнес-процессы – не отдельный, самостоятельный, ракурс. Этот вариант описания включает в себя предыдущий опыт организационно-структурного и функционального представления. Перечисленные типы описания отражают три различных взгляда на одно целое, они взаимосвязаны, так как элементы организационной структуры взаимодействуют в рамках бизнес-процессов с целью реализации определенных функций и общего назначения предприятия. Такое взаимодействие, потенциально отражающее все происходящее на предприятии, предлагается выразить и описать при помощи понятия "сервис", что открывает новые возможности для анализа процессов информатизации через систему мониторинга сервисов.

Сервисное описание

Сервис в ИС. Понятие сервиса пришло из области проектирования и построения информационных систем, где оно наиболее четко отображено в концепции сервис-ориентированной архитектуры (Service-Oriented Architecture, SOA). Идеи SOA не новы и не революционны, они отражают и вмещают парадигмы объектно-ориентированного программирования и проектирования на ином уровне абстракции.

Подход, основанный на представлении процессов, протекающих в бизнес-системах, в виде сервисов используется в IT-отрасли для построения информационной архитектуры, ориентированной на сервисы. SOA [6] связывает различные функциональные модули приложений, называемые сервисами, посредством четко определенных интерфейсов и соглашений между ними. Наблюдается устойчивый рост интереса к концепции сервис-ориентированной архитектуры.

Понятие сервиса. Сервис не строго техническое понятие. Под сервисом следует понимать любое результативное взаимодействие, обладающее определенной полезностью. Предприятие – структура, где зарождаются, предоставляются и потребляются сервисы.

Через понятие сервиса выражаются все функциональные и информационные взаимодействия предприятия. Следует отметить, что именно понятие "сервис" наиболее близко к понятию "взаимодействие", т. е. к наиболее точному его отражению, что важно для обоснования актуальности существования данной теории в условиях движения к информационной сетевой экономике, основанной на информационных взаимодействиях.

Сервис предприятия. Сервисный подход позволяет более детально и тщательно рассмотреть структурные элементы самого бизнес-процесса и дать новое описание предприятия как совокупности сервисов, отражающее взаимодействия внутри бизнес-процессов. Сервисная модель – это более углубленное описание, которое не отвергает процессный подход, а напротив, основываясь на нем, продвигает и расширяет его.

Предлагается использовать SOA подход на ином уровне абстракции, для описания предприятия со стороны его информационно-функционального взаимодействия, как внутреннего, так и внешнего. В этих целях под сервисом можно понимать как услуги сторонних организаций, так и взаимодействие между различными подразделениями внутри предприятия или любое результативное взаимодействие между сотрудниками.

Актуальной задачей видится разработка общей методологии и конкретных методик анализа состояния бизнес-систем, включая количественные оценки сервисов предприятий, а также способов описания структур на основе сервис-ориентированной модели. Описав предприятия при помощи сервисов, т. е. составив "сервисную модель" или сервисный портрет предприятия, получаешь возможность анализировать и сопоставлять модели различных предприятий. Это достигается благодаря предлагаемой классификации сервисов.

У каждой модели описания есть свои характеристики и свойства. Характеристики модели описания зависят и определяются свойствами того элемента или сущности, которые лежат в основе каждой модели:

Структурная модель → Организационная единица
Функциональная модель → Функция предприятия (внутренняя или внешняя)
Процессная модель → Бизнес-процесс
Сервисная модель → Сервис

Исходя из общих целей – оценки эффективности электронного предприятия как элемента сетевой экономики в период перехода к ИО, следует использовать модель, которая обладает наибольшим количеством актуальных мониторируемых свойств, – сервисную модель. Сегодня для общества и экономики особенно важны вопросы эффективного перехода, т. е. наиболее быстрого и устойчивого [3]. Как известно из теории управления, оценка эффективности нормальной деятельности в стабильной системе методологически отличается от оценки эффективности в переходный период.

При расчетах эффективности необходимо иметь аппарат для сравнения различных состояний, т. е. количественные оценки для измерения динамических характеристик, которые в свою очередь определяются той или иной моделью. Следовательно, необходима модель, позволяющая выделить наибольшее число количественных параметров для оценки динамических показателей современного предприятия.

Выделим следующие параметры и группы параметров, представляющие возможный интерес.

- Группа временных параметров:
  среднее время выполнения каждой операции;
  суммарное время выполнения процесса;
  максимальное/минимальное время выполнения процесса.
- Группа стоимостных параметров:
  средняя стоимость выполнения каждой операции;
  суммарная стоимость выполнения процесса;
  максимальное/минимальное время выполнения процесса.
- Группа ресурсных параметров:
  количество ресурсов, соотнесенное с процессом;
  общее количество ресурсов, задействованных в процессе;
  степень возможного дробления ресурсов.
- Характер взаимодействия между объектами в процессах предприятия.
- Ответственные роли.
- Параметры структурной сложности:
  степень связанности элементов;
  максимальное/минимальное количество элементов.

Группа временных параметров присуща только процессной и сервисной моделям. В структурных и функциональных моделях временные характеристики не являются основными и естественными и порой отсутствуют. Стоимостные параметры есть у процессной и сервисной моделей, могут добавляться к функциональной как временные, но будут носить менее детальный, описательный, характер. Зато параметры структурной сложности – вполне универсальны. Их можно выделить в любой модели, но смысловое, содержательное, наполнение параметров в разных моделях будет отличаться. Это объясняется тем, что способы их вычисления будут неодинаковы в разных моделях.

Ресурсные параметры присутствуют в полном объеме у процессной и сервисной моделей. У структурной и функциональной они могут косвенно присутствовать, т. е. при развитии модели появляться в некоторой форме. В процессной и сервисной моделях они все же более естественны, осязаемы и поддаются количественной оценке. Таким образом, с точки зрения ресурсных параметров эти модели отличаются по степени четкости и очевидности их отражения.

Другой актуальный и важный параметр – "характер взаимодействия" между объектами в процессах предприятия. Этот параметр особенно актуален для мониторинга переходного состояния. С его помощью можно определять процент "физических", "нематериальных", "информационных",

"когнитивных" и "комбинированных" сервисов в общей совокупности на предприятии. Такой показатель определяет место предприятия и позволяет отслеживать его движение в переходный период. Данный параметр доступен для сервисной модели, точнее, для такого ее представления, когда она наиболее близка к процессному.

Рассмотрим более подробно сервисную модель и ее представления. Рис. 3 иллюстрирует сущностное отличие сервисной модели от процессной. Обе модели могут отражать одни и те же явления, но иметь различное представление. В модели БП все представляется в естественной последовательности. Если процесс включает задачу A, содержащую действия X и Z, и задачу B, содержащую действия X и Y, то процессная и сервисная модели будут выглядеть, как представлено на рис. 3.

Суть сервисной модели заключается в выделении объектов (сервисов разного уровня) для повторного использования. Методики выделения сервисов в данном случае могут соответствовать методикам выделения объектов при объектно-ориентированном проектировании информационных систем. Одна из таких методик – выделение глаголов и существительных в вербальном описании прецедентов [4].

Сервисная модель может иметь разные представления, что позволяет использовать ее для выявления множества разных параметров, не доступных для других описаний. Приведенный пример описания через сервисную модель – это лишь один из возможных вариантов описания. Различные представления сервисной модели, статические и динамические, можно выразить при помощи логики и семантики построения диаграмм UML (Unified Modeling Language). UML – открытый стандарт, использующий графические обозначения для создания абстрактной модели системы. UML применяют для моделирования бизнес-процессов, системного проектирования и отображения организационных структур [4].

Другие примеры сервисной модели представлены на рис. 4.

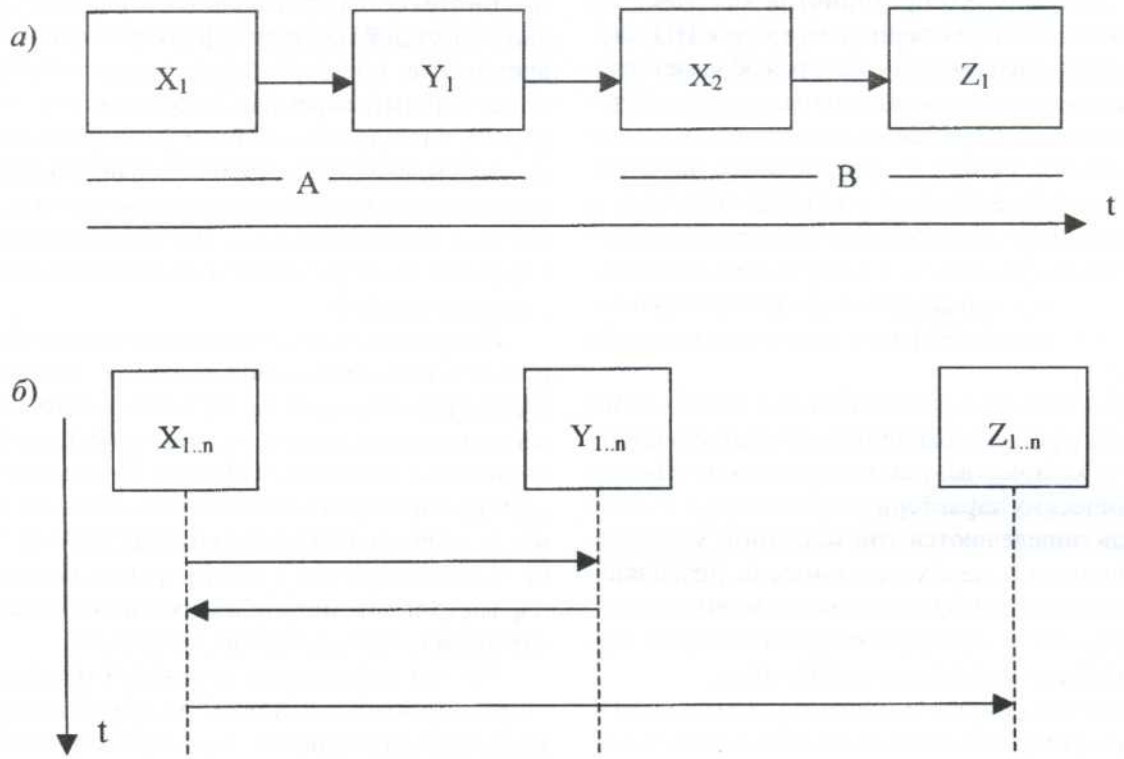


Рис. 3. Сервисная модель: процессное (*а*); и сервисное (*б*) представление

Отметим, что диаграмма классов может быть разной степени описательной детализации.

Могут быть и другие описания, иллюстрирующие, например, "включения" одних сервисов в другие, подразделения их на группы и пр. Это можно использовать как раз для классификации сервисов на "физические", "нематериальные" и "комбинированные".

Параметры, исключительные для сервисной модели:

• Количество сервисов в соотношении с количеством задач. Данный параметр выявляется во всех представлениях. В предельном случае одна задача равна одному взаимодействию, в другой крайности – все наоборот. За счет этого параметра можно определять соответствие построенной системы процесса выбранному критерию.

• Структурная сложность (для некоторых представлений). Данный показатель перекликается с предыдущим, тут проявляется и добавляется параметр структурной сложности, который вычисляется на основе графовой модели [5]. В случае, когда одна задача равна одному взаимодействию (сервису), он минимален. Таким образом вычисляется загруженность одной задачи сервисами.

• Количество одинаковых методов и атрибутов в разных сервисах. Здесь определяется степень похожести сервисов, что позволит их структурировать, создавать классы, делить на родовые, последователи и пр.

• Взаимодействие объектов. Сервисная модель позволяет отследить все взаимосвязи между объектами путем простого выделения всех "соединений". Показатели взаимодействия объектов можно использовать при построении матрицы инцидентности для определения структурной сложности, что впоследствии можно сравнивать с параметром структурной сложности графа БП. Также этот параметр помогает уяснить структуру функционирования процессов и возможное построение ИС для автоматизации процесса.

Вычисляя данные показатели, можно сравнивать их как с самими собой для определения динамики (мониторинг), так и с другими объектами (анализ-сравнение).

Таким образом, сравнение различных моделей описания электронных предприятий позволяет сделать выводы о практической целесообразности применения всех моделей в различных задачах исследования. Для задач, связанных с анализом динамических характеристик и свойств, более предпочтительны модели, основанные на сервисном подходе. Укрупненные количественные параметры и характеристики дают представление о поведенческих свойствах объектов, описываемых сервисной моделью – целесообразной и продуктивной для описания деятельности предприятий как основных экономических субъектов, особенно в период перехода к информационному обществу и информационной экономике.

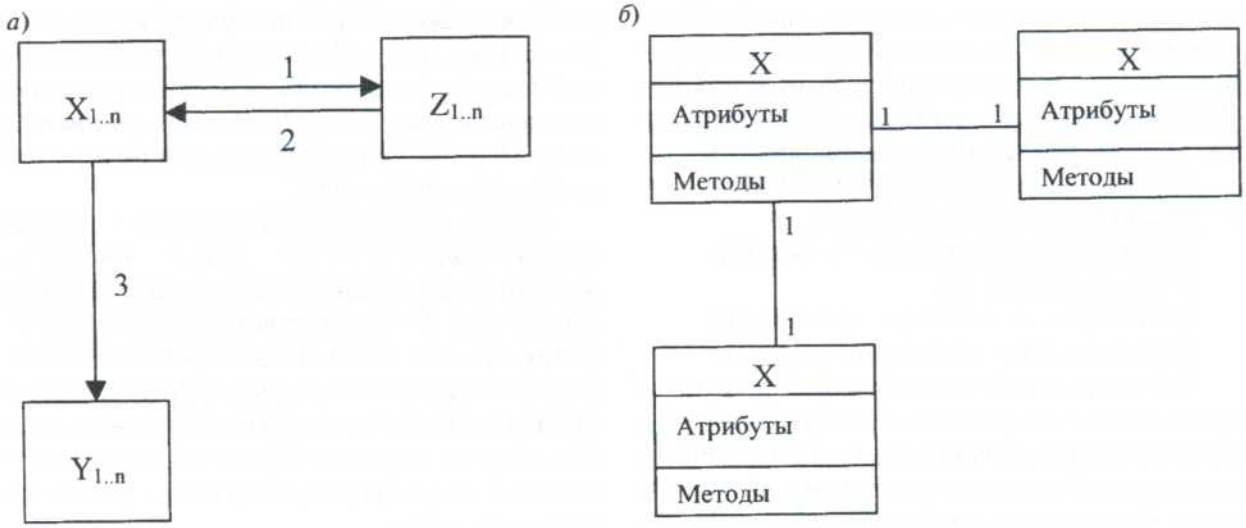


Рис. 4. Сервисное представление:
а – диаграмма взаимодействия; *б* – диаграмма классов

СПИСОК ЛИТЕРАТУРЫ

1. Status Report on European Telework. (Telework 1997, European Commission Report).
2. **Шеер А.-В.** Бизнес-процессы. Основные понятия. Теория. Методы. М.: Весть-МетаТехнология, 1999.
3. **Арзуманян М.Ю.** Мониторинг "информатизации" предприятий в процессе перехода к информационной экономике // Вестник ИНЖЭКОНа. 2008. № 6. С. 262.
4. **Ларман К.** Применение UML 2.0 и шаблонов проектирования. М.: ИД "Вильямс", 2009.
5. **Кристофидес Н.** Теория графов: алгоритмический подход. М.: Мир, 1978.

УДК 004.032.34

А.П. Тейхриб, Л.Г. Доросинский

СРАВНИТЕЛЬНЫЙ ОБЗОР МОДЕЛЕЙ ЛОГИЧЕСКОГО ВРЕМЕНИ

Одним из методов, применяемых в асинхронных распределенных системах для синхронизации процессов, является упорядочивание событий по логическим временным отметкам. Существует большое количество моделей логических часов, которые были разработаны на протяжении последних тридцати лет, начиная с фундаментальной работы Лампорта [1], определившей отношения между событиями в распределенной системе, и заканчивая, например, [2]. В русскоязычной литературе практически отсутствуют работы на данную тему, за исключением статьи [3], в которой вводится русская терминология для логического времени, приводится соответствующий формализм (близкий к [4]), а также описываются несколько моделей логических часов и способы их оптимизации. Однако многим моделям не было уделено внимания. Цель данной статьи – обзор различных моделей логических часов, не вошедших в [3] (за исключением векторных логических часов, которые приведены тут для общности).

Рассматриваются следующие модели:

векторные логические часы [5, 6];

адаптивные правдоподобные часы [2];

разностные часы [7];

часы на основе динамических цепочек.

Статья содержит следующие разделы.

Формализм логических часов. Приводится краткое описание математической модели ([3]), которая будет использована в качестве "общего знаменателя" для рассматриваемых логических часов. Использование общей модели необходимо, так как позволяет взглянуть на все модели с единой точки зрения.

Модели логических часов. Приводится описание перечисленных моделей логических часов. Строится оно по принципу: модель, свойства, область применения.

Результаты сравнения. В этом разделе производится сравнение по возможностям масштабирования, а также наличию свойства изоморфности. Последнее – важный показатель качества модели часов (свойство описано ниже). Сравниваются также области применения.

ФОРМАЛИЗМ ЛОГИЧЕСКИХ ЧАСОВ

В публикациях на тему логических часов используются различные формализмы для описания модели логического времени, что приводит к определенным трудностям. Как упоминалось выше, в русскоязычной литературе имеет место только одна работа [3], поэтому для формирования традиции в данной статье будет использоваться тот же формализм (особенно с учетом того, что он близок к предложенному в [4]). Введем несколько определений.

Модель распределенной системы – направленный граф $D = < P, C >$, где P – множество вершин графа, представляющих собой процессы системы, C – множество ребер, соответствующих каналам связи. Также предполагаем определенный алфавит сообщений M, и используем *null* для задания отсутствия сообщения в канале. Для каждого процесса P_i будем обозначать out_i – множество исходящих ребер и in_i – множество входящих ребер.

Процесс $P_i \in P$ – совокупность $P_i = < \boldsymbol{S}_i, \boldsymbol{S}_i^0, Msg_i, Trans_i >$, где $\boldsymbol{S}_i$ – множество состояний,

$S_i^0 \in S_i$ – множество начальных состояний, Msg_i – функция, отображающая $S_i \times out_i$ в элементы $M \cup \{null\}$, которые означают отсутствие сообщения. $Trans_i$ – функция переходов, отображающая S_i и векторы (индексируемые по in_i) сообщений из $M \cup \{null\}$ в S_i.

Событие – переход процесса из одного состояния в другое. Все события в распределенной системе можно разделить на внутренние и коммуникационные, которые в свою очередь делятся на события отправки и приема. e_j^i – j-е событие процесса P_i.

Вычислительная история $h_i = (e_0^i, e_1^i, ..., e_j^i, ...)$ – упорядоченная последовательность событий процесса P_i.

Множество событий системы $E = \cup_{i=1,n} E_i$, где $E_i = \{e_j^i, j = 0, 1, ...\}$ – упорядоченная последовательность событий процесса P_i.

Введем отношения между событиями системы:

Отношение совпадения $a \equiv b$, где $a \in E$, $b \in E$ – тривиальное отношение эквивалентности на множестве E.

Отношение элементарного условия a – элементарное условие b, если не выполняется $a \equiv b$ и при этом: a и b – события одного и того же процесса, причем a предшествует b в вычислительной истории этого процесса, или a и b – соответственно события передачи и приема одного и того же сообщения.

Отношение обусловленности на множестве E (обозначается $a \to b$) есть транзитивное замыкание отношения "является элементарным условием". Два события a и b называются взаимно необусловленными ($a \| b$), если не выполняется ни $a \to b$, ни $b \to a$.

Логическое время – совокупность правил и соглашений, позволяющих каждому процессу асинхронной распределенной системы ставить в соответствие произвольному событию e своей вычислительной истории некоторую величину $C(e)$ из заданного множества T. T – множество значений логического времени. $C(e)$ – значение логического времени.

Формально система логического времени задается пятеркой

$$SLT = < T, G, C, M, R >,$$

где T – множество значений логического времени, одно из которых выбирается в качестве начального; G – множество значений временных маркеров (временной маркер – составная часть передаваемого сообщения в системе, функция от временной отметки события отправки); C – функция логического времени, $C: E \to T$; M – функция датировки сообщения, т. е. присвоения временной метки сообщениям на основе временной метки события их отправки; R – функция сравнения логических времен, $R: T \times T \to O$, $O = \{EQ, GT, LT, NC\}$, где EQ – равенство временных меток, GT – отношение больше, LT – отношение меньше, NC – отношение несопоставимости.

Свойства логических часов. Перечислим свойства, которыми может обладать система логического времени.

Монотонность (состоятельность в широком смысле). Система обладает свойством монотонности, если

$$a \to b \Rightarrow R(C(A), C(b)) = LT.$$

Строгая состоятельность. Система обладает свойством строгой состоятельности, если

$$a \to b \Leftrightarrow R(C(a), C(b)) = LT.$$

Правдоподобность. Система обладает свойством правдоподобности, если

$$a \to b \Rightarrow R(C(a), C(b)) = LT;$$

$$a \equiv b \Leftrightarrow R(C(a), C(b)) = EQ.$$

Изоморфность. Система обладает свойством изоморфности, если

$$a \to b \Leftrightarrow R(C(a), C(b)) = LT;$$

$$a \equiv b \Leftrightarrow R(C(a), C(b)) = EQ;$$

$$a \| b \Leftrightarrow R(C(a), C(b)) = NC.$$

Данная математическая модель позволяет формально подойти к описанию и сравнению различных моделей логического времени.

МОДЕЛИ ЛОГИЧЕСКИХ ЧАСОВ

Классические векторные логические часы

Описание. Классические векторные логические часы были предложены в [5, 6]. Основная идея заключается в том, что на каждом узле поддерживается N-мерный вектор, где N – количество процессов в системе, т. е. каждому процессу P_i, приписывается вектор V_i, причём $V_i[i]$ – число событий, которые произошли с процессом P_i к настоящему времени. Другими словами, если $V_i[j] = k$, то процесс P_i знает, что с процессом P_j произошло k событий.

Каждому вновь создаваемому событию (приход сообщения с другого узла также является новым событием) приписывается векторная отметка, равная векторной отметке на процессе, в котором создается данное событие на момент создания этого события, при этом $V_i[i]$ увеличивается на некоторую величину d (обычно равную единице).

Модель. Множество значений логического времени представляет собой множество N-мерных векторов с целыми неотрицательными компонентами. Множество временных маркеров совпадает с множеством значений логического времени.

Функция логического времени C для процесса P_i работает следующим образом:

изначально $\forall k \in [1..N] : V_i[k] = 0$;

в случае локального события, или события отправки, $V_i[i] = V_i[i] + 1$;

в случае события получения сообщения с векторной отметкой V_r происходит $\forall k \in [1..N] : V_i[k] = \max\{V_i[k], V_r[k]\}$, затем $V_i[r] = V_i[r] + 1$.

Функция датировки M присваивает каждому исходящему сообщению значение функции C для соответствующего события отправки.

Рассмотрим работу функции сравнения. Предположим, что есть две векторные отметки V_1, V_2:

$$R(V_1, V_2) = EQ \Leftrightarrow \forall k \in [1...N] : V_1[k] = V_2[k];$$

$$R(V_1, V_2) = LT \Leftrightarrow \forall k \in [1...N] : V_1[k] \le V_2[k] \wedge \exists_j : V_1[j] < V_2[j];$$

$$R(V_1, V_2) = GT \Leftrightarrow \forall k \in [1...N] : V_1[k] \ge V_2[k] \wedge \exists_j : V_1[j] > V_2[j];$$

$$R(V_1, V_2) = NC \Leftrightarrow \exists i : V_1[i] > V_2[i] \wedge \exists_j : V_1[j] < V_2[j].$$

Данные часы обладают свойством изоморфности.

Область применения. Классические векторные логические часы находят применение в системах, где требуется изоморфность. Однако основной недостаток таких часов – потребность в N компонентах для поддержания изоморфности, что препятствует масштабируемости.

Адаптивные правдоподобные часы

Описание. Адаптивные правдоподобные часы были предложены в [2] и являются развитием R-компонентных правдоподобных часов, описанных в [4]. Цель данного подхода в снижении стоимости использования классических векторных логических часов [5, 6]. Данные часы также используют векторы, но в несколько ином виде. Особенность адаптивных правдоподобных часов заключается в возможности настройки объема передаваемой информации (например, увеличение размера отметки времени до размера векторной отметки и получение векторных логических часов). Стоимость стандартных векторных часов оценивается в $O(N)$, где N – число процессов в распределенной системе. Данный механизм имеет проблемы с масштабируемостью, хотя и позволяет точно упорядочивать причинно-связанные события. Для повышения масштабируемости предлагаются механизмы, обладающие меньшей точностью, но с фиксированным размером отметок. Главное отличие от обычных правдоподобных часов (например, [4]) заключается в том, что каждый процесс может сам выбрать способ связи идентификатора процесса и компоненты векторной отметки, причем способ связи может меняться динамически во время работы распределенной системы.

Модель. Множество значений адаптивных правдоподобных часов и совпадающее с ним множество значений временных маркеров представляют собой множество наборов $< i, V_i, f_i >$:

$$T = G = \{(i, V_i, f) : i \in \mathbf{Z}_0, V \in \mathbf{Z}_0{}^L, f \in \mathbf{F}\},$$

где L – заранее определенное количество компонент вектора; F – множество функций преобразования идентификаторов процессов в компоненты векторов; i – уникальный идентификатор процесса; V – вектор размерности L; f – функция преобразования идентификатора процесса в номер позиции вектора.

Функция логического времени работает следующим образом:

изначально: i = идентификатор процесса; f_i некоторая функция преобразования; $\forall r \in \{1...L\} : V_i[l] = 0$;

при событии отправки либо локальном событии создается новая временная отметка

$$< i, V_i^+, f_i^+ >,$$

где f_i^+ = обновленная функция преобразования;

$$V_i^+[l] = \max\{V_i[f_i(j)] : \forall j \cdot f_i^+(j) = l\} + own(l),$$

где $own(l) = \begin{cases} 1, \text{если } f_i^+(i) = l \\ 0 \text{ иначе} \end{cases}$;

при событии прихода сообщения с временной отметкой $< i_s, V_s, f_s >$ создается новая временная отметка $< i, V_i^+, f_i^+ >$:

f_i^+ = обновленная функция преобразования;

$V_i^+[l] = \max\{V_i[f_i(j)], V_s[f_s(j)] : \forall j \cdot f_i^+(j) = l\} + own(l)$.

Функция датировки M присваивает каждому событию значение функции C. С каждым отправляемым сообщением передается временная отметка, присвоенная событию отправки этого сообщения.

Данные логические часы не определяют функцию преобразования, в качестве которых могут использоваться различные функции.

Можно сказать, что данный вид логических часов есть обобщение классических векторных отметок, а также R-компонентных векторных отметок, описанных в [4], так как достаточно выбрать соответствующие f и L.

Рассмотрим работу функции сравнения R. Пусть есть две отметки $ts_i = < i, V_i, f_i >$ и $ts_j = < j, V_j, f_j >$, тогда:

$R(ts_i, ts_j) = EQ \Leftrightarrow i = j \wedge V_i[f_i(i)] = V_j[f_j(i)]$;

$$R(ts_i, ts_j) = LT \Leftrightarrow \begin{cases} i = j \wedge V_i[f_i(j)] < V_j[f_j(j)], \\ i \neq j \wedge (\forall k. V_i[f_i(k)] \leq \\ \leq V_j[f_j(k)]) \wedge (V_i[f_i(j)] < \\ < V_j[f_j(j)]); \end{cases}$$

$$R(ts_i, ts_j) = GT \Leftrightarrow \begin{cases} i = j \wedge V_i[f_i(j)] > V_j[f_j(j)], \\ i \neq j \wedge (\forall k. V_i[f_i(k)] \geq \\ \geq V_j[f_j(k)]) \wedge (V_i[f_i(j)] > \\ > V_j[f_j(j)]); \end{cases}$$

$R(ts_i, ts_j) = NQ \Leftrightarrow (V_i[f_i(i)] > V_j[f_j(i)]) \wedge (V_i[f_i(j)] < V_j[f_j(j)])$.

Данные часы являются правдоподобными, доказательство приводится автором в [4].

Область применения. Основной упор автор делает на возможность изменения функции преобразования и количества компонент вектора, что позволяет добиться компромисса между ошибками, которые допускают логические часы в упорядочивании событий, и возможностью масштабирования. Область применения для данного вида часов – системы с масштабируемой широковещательной рассылкой, а также приложения совместной разработки в системах с соединением точка-точка.

Разностные логические часы

Описание. Разностные часы были предложены в [7] как альтернатива векторным часам для систем с внутрипроцессным параллелизмом. Основная идея заключается в использовании различных логических часов для локальных и межпроцессных событий.

Модель. На каждом P_i узле вводятся два счетчика: lc_i – счетчик локальных событий в виде битового вектора переменной длины, и gc_i – счетчик глобальных событий, который представляет собой неотрицательное целое.

Множество значений функций логического времени и совпадающее с ним множество значений временных маркеров представляет собой набор из двух компонент $< B^{Z_0}, Z^4 >$. Первая компонента – это битовый вектор переменной длины, который отражает локальные зависимости между событиями и не передается вместе с сообщениями, а вторая компонента – 4-мерный вектор с целыми компонентами – предназначена для отражения зависимостей между событиями различных процессов. Рассмотрим составляющие второй компоненты:

идентификатор события, присваивающий значение счетчика глобальных событий узла, на котором создано событие;

идентификатор процесса, формируемый поразному для различного типа событий. Для событий отправки он равен идентификатору процесса, которому адресовано сообщение, для событий приема он равен идентификатору процесса, который отправил сообщение. Для внутренних событий данное поле равно *null*;

поле разности используется только для событий приема и равно разности между идентификатором события отправки соответствующего сообщения и данным событием приема;

тип события: отправка, прием, внутреннее.

Локальное влияние. Дополнительно вводится отношение локального влияния. Оно определяется для двух локальных событий, находящихся в отношении элементарного условия, и дополнительно определенного условия (решение отдается разработчику, например, можно явным образом выбирать такие события). Данное отношение транзитивно.

Счетчики работают следующим образом:

изначально оба равны нулю;

любое событие: $gc_i = gc_i + 1$, а размерность lc_i увеличивается на единицу и $lc_i[gc_i] = 1$, при этом для каждого k-го события, которое локально влияет на данное событие, производится операция побитового ИЛИ для всех компонент lc. В итоге получаем, что для каждого такого события $lc_i[k] = 1$, для остальных l событий $lc_i[l] = 0$.

Функция логического времени на процессе P_i работает следующим образом:

на начальном этапе все компоненты равны нулю;

при локальном событии $C = <lc_i, < gc_i, null, null, Local >>$;

при событии отправки процессу P_j, $C = < lc_i, < gc_i, j, null, Send >>$;

при событии получения от процесса P_j.

Функция датировки сообщения каждому отправляемому сообщению процесса P_i присваивает пару $< i, gc_i >$.

Для сравнения между собой временных меток они сначала приводятся к битовым матричным отметкам. Битовые матричные отметки – это множество N-мерных векторов (N – количество узлов в системе), каждая компонента которых представляет собой битовый вектор переменной длины (данный вектор строится аналогично первой компоненте разностных логических часов). j-я компонента данного вектора некоторого события на i-м процессе отражает зависимость данного события от событий j-го процесса. Затем по данным отметок однозначно восстанавливается зависимость событий. Однако восстановление матричных битовых отметок по разностным отметкам возможно только в случае снятия глобального состояния со всех процессов. Это накладывает ограничение на область применения данных часов.

Стандартные свойства логических часов тут неприменимы, так как иначе определяется отношение обусловленности.

Область применения. Основная область применения данных логических часов – это отладка распределенных приложений, так как здесь получение глобального состояния не является фактором, снижающим производительность системы. Глобальное состояние, как утверждалось выше, есть необходимое условие для получения полного графа зависимостей между событиями.

Часы на основе динамических цепочек

Описание. Данная модель логических часов была предложена в [8]. Цель данной модели – использование меньшего количества компонент, чем в классических векторных отметках, а также динамическое изменение состава процессов в системе. Данный механизм построен на использовании разделяемых структур данных и поэтому применим в системах с разделяемой памятью. Предполагается, что в распределенной системе количество событий, между которыми должна сохраняться причинно-следственная связь, меньше, чем общее количество событий (авторы называют 1 % зависимых событий от общего числа событий в системе). Векторные часы есть частный случай данных логических часов при условии, что все события в системе зависимы друг от друга. Основная идея заключается в том, что вычислительная история может быть представлена в виде набора цепочек из зависимых событий, связи между цепочками представляют собой обмен сообщениями между процессами. В данных логических часах вычислительная история разбивается на связанные цепочки, и векторные часы используются не для процессов, как в классическом варианте, а для цепочек. При этом количество компонент определяется количеством цепочек. По мере выполнения вычислений количество компонент в векторе может увеличиваться.

Модель. Множество значений функций логического времени и совпадающее с ним множество значений временных маркеров представляют собой вектор переменной длины, каждая компонента которого соответствует определенной цепочке в вычислительном процессе. Данный вектор поддерживается на каждом процессе.

Работа функции логического времени похожа на работу этой функции в классических векторных отметках; отличие заключается в выборе компоненты, – она должна измениться при появлении событий. Автор предлагает использовать абстрактную функцию GetIndex для выбора изменяемой компоненты, которая на входе принимает локальный вектор процесса, а также событие, для которого необходимо сформировать временную отметку. При этом возвращенный индекс события может быть больше числа компонент – тогда увеличивается число компонент и все они инициализируются нулем. Работа функции GetIndex различна для разных видов часов на базе цепочек. Общий алгоритм работы функции логического времени выглядит следующим образом:

изначально $C = \{0\}$ – вектор нулевой размерности;

при событии получения сообщения с отметкой m: $C = \max\{C, m\}$;

если событие e принадлежит цепочке связанных событий, то $index = GetIndex(C, e)$, $C[index] = C[index] + 1$.

Функция датировки присваивает каждому отправляемому сообщению значение логических часов для соответствующего события отправки.

В часах на базе динамической цепочки количество компонент может изменяться. В этой модели используется глобальная переменная – вектор Z, который соответствует глобальному состоянию системы в терминах цепочек по всей

системе. Изначально $Z = null$. Функция $GetIndex(v, e)$ работает следующим образом:

- Если $\exists i : Z[i] = V[i]$, то $index = i$ (если таких компонент несколько, то выбирается любая из них).
- В противном случае увеличиваем размер вектора Z, тогда $index = Z.size$.
- Увеличиваем компоненту $Z[index]$ на единицу.
- Возвращаем значение переменной $index$.

В статье автор показывает, что такой алгоритм позволяет достичь изоморфности логических часов.

Функция сравнения работает аналогично классическим векторным отметкам.

Область применения. Основная область применения данной модели логических часов – распределенные системы с разделяемой памятью, в которых достаточно мало связанных друг с другом событий (зависимые события составляют примерно 1 % от общего числа событий). При этом основная цель применения данных часов – получение информации о глобальном состоянии в распределенной системе и определение по этому глобальному состоянию выполнения каких-либо условий (Predicate detection). Основное преимущество заключается в сокращении количества компонент, используемых в векторной отметке. Для того чтобы смягчить влияние необходимости использования глобальных переменных, автор предлагает разделить хранение глобальных переменных по нескольким серверам, однако в этом случае возникают проблемы с синхронизацией этих глобальных переменных.

РЕЗУЛЬТАТЫ СРАВНЕНИЯ

Сравним рассмотренные логические времена по свойствам, которыми они обладают, и масштабируемости, так как именно эти параметры являются определяющими при выборе модели логических часов для решения некоторой задачи. Результаты сравнения представлены в таблице. Как видим, только одни логические часы обладают свойством изоморфности и масштабируемости одновременно, однако они требуют использования данных, к которым обеспечен общий доступ. Другой интересный вид логических часов – адаптивные правдоподобные, которые представляют собой обобщение векторных логических часов и могут быть настроены для обеспечения заданного уровня правдоподобия (вплоть до изоморфности, но потеряв при этом масштабируемость).

Сравнительная таблица логических временных отметок

Логические часы	Свойство	Масштабируемость
Векторные	Изоморфны	Нет
Адаптивные правдоподобные	Правдоподобны	Да
Разностные	Неприменимо	Да
На основе динамических цепочек	Изоморфны	Да (в системах с малым числом связанных событий)

В данной статье приведен обзор различных временных отметок в едином формализме, рассмотрены их свойства, а также приведены области применения, которые наиболее характерны для каждого вида логических часов. Построение обзора таким образом позволяет выбрать наиболее подходящий алгоритм логических временных отметок, максимально подходящий для конкретной задачи.

СПИСОК ЛИТЕРАТУРЫ

1. **Lamport L.** Time, clocks and the ordering of events in a distributed system // Commun. ACM. 1978. Vol. 21, № 7. P. 558–565.

2. **Gidenstam A., Papatrianta M.** Adaptive plausible clocks // ICDCS '04: Proc. of the 24th Intern. conf. on Distributed Computing Systems (ICDCS'04). IEEE, 2004. P. 86–93.

3. **Жуклинец И.А., Хотимский Д.А.** Логическое время в распределенных программных системах // Программирование. 2000. № 6. С. 3–19.

4. **Torres-Rojas F.J., Ahamad M.** Plausible clocks: constant size logical clocks for distributed systems // Distrib. Comput. 1999. Vol. 12, № 4. P. 179–195.

5. **Mattern F.** Virtual time and global states of distributed systems // Parallel and Distributed Algorithms. North-Holland, 1989. P. 215–226.

6. **Fidge C.** Logical time in distributed computing systems // Computer. 1991. Vol. 24, № 8. P. 28–33.

7. **Vaidehi S., Ram D.J., Shukla A.** Difference clocks: a new scheme for logical time in distributed systems // Computers and Digital Techniques, IEEE Proceedings. 1996. Vol. 143, № 6. P. 426–430.

8. **Agarwal A., Garg V.K.** Efficient dependency tracking for relevant events in shared-memory systems // PODC '05: Proceedings of the twenty-fourth annual ACM symposium on Principles of distributed computing. New York, NY, USA: ACM, 2005. P. 19–28.

УДК 519.876.5

И.И. Икрянов

МОДЕЛИ ДВИЖЕНИЯ ИНДИКАТОРОВ КУРСА И ГЛИССАДЫ С МАЯТНИКОВОЙ СИСТЕМОЙ СТАБИЛИЗАЦИИ ОПТИЧЕСКИХ ОСЕЙ

Использование вертолетов на кораблях различных типов существенно повышает их эксплуатационные возможности. Вместе с тем использование вертолетов палубного базирования предъявляет существенные требования к обеспечению безопасности взлета и особенно посадки на палубу корабля.

Одним из важных аспектов безопасной посадки вертолета на взлетно-посадочную палубу (ВППл) корабля является обеспечение пилота вертолета курсо-глиссадной информацией. Для этих целей в разрабатываемых в настоящий момент отечественных системах визуальной посадки вертолета на ВППл используются оптические индикаторы курса (ИК) и глиссады (ИГ) с пассивной (маятниковой) системой стабилизации оптических осей относительно крена и дифферента корабля.

Конструктивно ИК/ИГ представляют собой блоки огней, которые объединены в две сборки, одна из которых используется для индикации курса/глиссады вдоль продольной оси корабля, т. е. по диаметральной плоскости в сторону кормы, а вторая под углом 10 град от ДП в сторону кормы. Обе сборки ИК/ИГ устанавливаются в общую раму, которая крепится к корпусу с помощью карданового подвеса, который в свою очередь обеспечивает свободное движение блоков по крену в диапазоне углов ±15 град и дифференту в диапазоне ±6 град. При отклонении от заданной глиссады вверх летчик наблюдает желтые, а вниз – красные огни ИК/ИГ. При отклонении от заданного курса влево летчик наблюдает синие огни, мигающие с одной частотой, а вправо – с другой. Примерно на расстояния до 100 м летчик видит огни раздельно в виде точек. При меньшем расстоянии он видит все огни одновременно, поэтому полет с использованием ИК и ИГ возможен в диапазоне расстояний по дальности от 2 км до 100–150 м.

Качка корабля вызывает погрешность стабилизации ИК/ИГ относительно истинной вертикали, которая выражается в колебаниях оптической оси индикаторов курса и глиссады. Колебания оптической оси индикаторов, в свою очередь, вызывают искажение сигнала, передаваемого летчику. Погрешность стабилизации индикаторов курса и глиссады относительно истинной вертикали зависит от размещения ИК/ИГ на ВППл (расстояния точки подвеса ИК/ИГ от кормы и борта корабля), конструктивных особенностей ИК/ИГ, от амплитуды и частоты качки корабля по крену и дифференту.

Величина погрешности стабилизации ИК/ИГ относительно истинной вертикали – фактор, определяющий возможность эксплуатации ИК/ИГ. Целью автора работы была разработка модели, которая позволит вычислять погрешность стабилизации ИК/ИГ относительно истинной вертикали.

Модель движения ИК/ИГ при воздействии качки корабля на регулярном и нерегулярном волнении

Система координат для описания движения ИК/ИГ показана на рис. 1. Вертикальные перемещения $h(t)$ центра масс корабля вдоль оси O_0Y_0 обусловлены возникновением восстанавливающей силы и характеризуют вертикальную качку корабля. Поворот на углы $\psi_к$ и $\psi_д$ характеризует качку корабля по крену и дифференту соответственно. Будем полагать, что корабль не совершает маневров по курсу и угол рысканья всегда равен нулю.

Угловое положение маятника в подвижной системе координат $OXYZ$ определяется углом φ.

Координаты точки T в системе $O_0X_0Y_0Z_0$ равны:

$$\begin{cases} x_r = R_{к/д}\cos(\psi_{к/д}) - D\sin(\psi_{к/д}) + \\ + L\sin(\varphi_{к/д} + \psi_{к/д}); \\ y_r = R_{к/д}\sin(\psi_{к/д}) - D\cos(\psi_{к/д}) + \\ + L\cos(\varphi_{к/д} + \psi_{к/д}) + h(t). \end{cases} \quad (1)$$

Уравнения движения ИК/ИГ в обобщенных координатах можно записать в виде уравнений Лагранжа 2-го рода [3]:

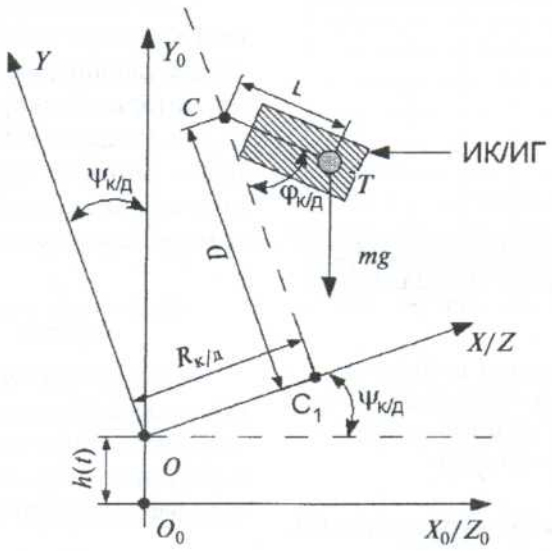


Рис. 1. Система координат
для описания плоского движения ИК/ИГ

$$\frac{d}{dt}\left(\frac{\partial K}{\partial \dot{q}_i}\right)-\frac{\partial K}{\partial q_i}=Q_i,(i=1,2,...,n), \qquad (2)$$

где $q_i, \dot{q}_i$ – обобщенные координаты и скорости; Q_i – обобщенные силы; n – число степеней свободы системы.

Будем полагать, что

$$Q_i=-\frac{\partial \Pi}{\partial q_i}-B_i q_i, \quad (i=1,2,\ ...,\ n)$$

где B_i – абсолютный коэффициент демпфирования.

Дифференцируя (1) и подставляя в (2), пренебрегая членами второго и большего порядка малости, получаем:

$$\ddot{\phi}\ +2C\sqrt{\frac{g}{L}}\cdot\dot{\phi}+\frac{1}{L}(g+\ddot{h}(t)\ -D\dot{\psi}^2_{к/д}\ +R_{к/д}\ddot{\psi}_{к/д})\,\phi= \qquad (3)$$
$$=\frac{1}{L}((D-L)\ddot{\psi}_{к/д}-(g+\ddot{h}(t)\ \psi_{к/д}+R_{к/д}\dot{\psi}^2_{к/д}),$$

где $C=\dfrac{B}{2gmL^2\sqrt{\dfrac{g}{L}}}$.

Уравнение (3) описывает плоское движение ИК/ИГ. Основная интересующая нас характеристика движения ИК/ИГ – это абсолютная величина отклонения ИК/ИГ от истинной вертикали. Данная величина определяется как сумма входного воздействия – угла наклона палубы корабля и выходного сигнала системы – угла наклона ИК/ИГ:

$$U=|\varphi+\psi_{к/д}|. \qquad (4)$$

Существует несколько способов описания качки корабля на морском волнении. В способе описания качки корабля на морском волнении, разработанном А.Н. Крыловым, принимается, что морское волнение периодично по времени, профиль взволнованной поверхности синусоидальный, гребни волн параллельны, волны распространяются в одном направлении. Реакция корабля на такое воздействие также может описываться синусоидой. Данная модель хорошо описывает мертвую зыбь и может быть использована для моделирования установившегося ветрового волнения при не очень больших значениях скорости ветра.

Более точные способы рассматривают морское волнение, а следовательно, и качку корабля как случайный процесс. Одной из характеристик таких процессов является спектральная плотность случайного процесса на выходе линейной системы, которую удобно использовать для моделирования качки корабля в условиях нерегулярного волнения.

В соответствии с моделью качки корабля, предложенной А.Н. Крыловым, будем полагать

$$h(t)=0,\ \psi_{к/д}=\psi_{max}\sin(\omega t). \qquad (5)$$

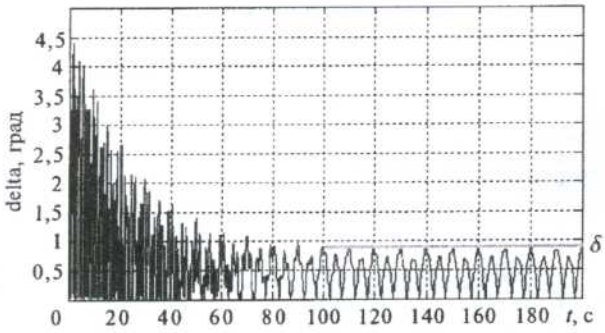


Рис. 2. Пример определения величины отклонения ИК/ИГ от истинной вертикали с помощью модели ИК/ИГ при качке корабля на регулярном волнении

Подставляя выражения (5) в уравнение (3), получаем:

$$\ddot{\phi} + 2C\omega_0\phi + \omega_0^2 \times$$
$$\times\left(1-\frac{1}{2}\frac{D}{L}\psi_{\max}^2 \;\eta_2-\frac{R}{L}\psi_{\max}\eta^2\sin(\omega t)\right)\phi = \quad (6)$$
$$=-\omega_0^2\,\psi_{\max}\left(\left(1+\frac{D-L}{L}\eta^2\right)\sin(\omega t)-\frac{1}{2}\frac{R}{L}\psi_{\max}\eta^2\right),$$

где $\eta=\frac{\omega}{\omega_0}, \omega_0=\sqrt{\frac{g}{L}}$.

Подставляя численное решение уравнения (6) и соотношения (5) в (4), получаем значение отклонения ИК/ИГ от истинной вертикали (рис. 2). В численном решении уравнения (6) период времени, на протяжении которого идет затухание собственных колебаний маятника, отбрасывается.

С целью моделирования нерегулярного морского волнения и соответственно качки корабля на нем удобно использовать методы Монте-Карло, которые базируются на спектральном описании процесса качки. Для использования данного метода плотность спектра качки, полученная экспериментальным путем, аппроксимируется выражением вида

$$S(w)=\frac{A}{\sqrt{2\pi D}}e^{\frac{(\omega-\omega_m)^2}{2D}}+\frac{A_1}{\sqrt{2\pi D_1}}e^{\frac{(\omega-\omega_{m1})^2}{2D_1}}.$$

Коэффициенты A, A_1, D, D_1, ω_1, ω_{m1} подбираются таким образом, чтобы наилучшим образом аппроксимировать полученные экспериментальные данные. После чего предполагается, что качка корабля на волнении задается выражением

$$\psi_{к/д}=\Sigma_i A_i\sin(\omega_i t+f_i). \quad (7)$$

Для генерации выборок A_i и ω_i используется метод усечения, суть которого состоит в следующем.

График функции $S(\omega)$ вписывается в прямоугольник (рис. 3). На ось Y подают случайное равномерно распределенное число из ГСЧ (генератор случайных чисел). На ось X подают случайное равномерно распределенное число из ГСЧ. Если точка в пересечении этих двух координат лежит ниже кривой спектральной плотности качки, то событие X произошло – гармоника с соответствующей частотой добавляется в выборку, иначе – нет. После получения таким образом выборки $\{\omega_i\}$ производится ее сортировка по возрастанию значений ω_i. Амплитуда каждой отдельной гармоники определяется из соотношения Винера–Хинчина:

$$A_i=\sqrt{2S(\omega_i)(\omega_{i+1}-\omega_i)}.$$

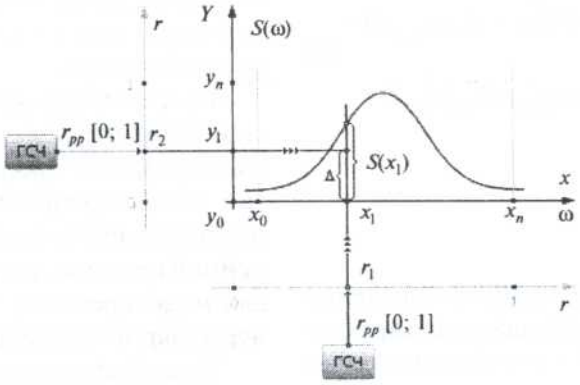


Рис. 3. Генерации выборок A_i и ω_i для выражения (7) с помощью метода усечения

Последняя гармоника в выборке отбрасывается приравниванием ее амплитуды к нулю. Фазы каждой из гармоник получаются с помощью ГСЧ как равномерно распределенное число на интервале 0–2π.

Определив, таким образом, векторы $\{A_i\}$, $\{\omega_i\}$ и $\{f_i\}$ и подставив (7) в выражение (3), получим:

$$\ddot{\phi} + 2C\sqrt{\frac{g}{L}}\cdot\dot{\phi} + \frac{1}{L}\left(g + D\left(\sum_i A_i\omega_i \cos(\omega_i t + f_i)\right)^2 - R_{к/д}\left(\sum_i A_i\omega_i^2 \sin(\omega_i t + f_i)\right)\right)\phi = \frac{1}{L}\left(-(D-L)\sum_i A_i\omega_i^2 \sin(\omega_i t + f_i) - g\sum_i A_i \sin(\omega_i t + f_i) + R_{к/д}\left(\sum_i A_i\omega_i \cos(\omega_i t + f_i)\right)\right). \quad (8)$$

Выражение (8) описывает движение ИК/ИГ при воздействии на него нерегулярной качки корабля. Решая данное уравнение численно и подставляя решение в выражение (4), получаем зависимость величины отклонения ИК/ИГ от истинной вертикали при воздействии нерегулярной качки судна от времени $\delta(t)$.

Чтобы получить конечный результат, произведем выборку значений из полученной зависимости. Первые 100 с, на протяжении которых происходит затухание собственных колебаний ИК/ИГ, отбрасываются. Оставшийся временной интервал разобьем на отрезки длительностью по 60 с. На каждом отрезке определим максимальное значение величины отклонения ИК/ИГ от истинной вертикали и будем считать его результатом испытания, произведенного на данном отрезке времени. В итоге получится выборка значений величины отклонения ИК/ИГ от истинной вертикали при воздействии нерегулярной качки судна $\{\delta_k\}$, которую можно рассматривать как результат определенного количества испытаний, при этом $\{\delta_k\} \in \delta(t)$. Для повышения точности расчетов рекомендуется получить несколько выборок.

Далее необходимо на основании одной или нескольких выборок определить доверительный интервал (δ_{min}, δ_{max}) величины отклонения ИК/ИГ от истинной вертикали по известным формулам [1].

Результаты расчета величины отклонения ИК/ИГ от истинной вертикали, выполненные с помощью различных моделей, представлены в таблице.

В качестве исходных данных для расчета по модели движения ИК/ИГ в условиях нерегулярной качки корабля использовались плотности спектров качки корабля, по дифференту полученные с помощью метода, описанного в [2]. Период собственных колебаний корабля по дифференту предполагался равным 5 с. Плотность спектра волновых ординат бралась из ОСТ 5.1003-80 для случая развитого морского волнения. При этом предполагается, что сравнение указанных моделей адекватно лишь при $h_{3\,\%} \le 2$ м. При больших значениях $h_{3\,\%}$ качку корабля нельзя рассматривать как регулярную и соответственно некорректно сравнивать указанные модели.

Для интегрирования уравнений (6) и (8) использовался метод Рунге–Кутта 4-го порядка с фиксированным шагом [4]. Шаг интегрирования равнялся $\Delta t = 0{,}01$ с. Для подготовки выборок $\{A_i\}$, $\{\omega_i\}$ и $\{f_i\}$ в уравнении (8) использовался штатный генератор случайных чисел системы Matlab2008b. Количество элементов в каждой выборке равнялось 1000.

Из полученных результатов видно, что при волнении $h_{3\,\%} \le 2$ м результаты, полученные с помощью обеих моделей, хорошо согласуются друг с другом.

Расчет величины отклонения ИК/ИГ от истинной вертикали

Амплитуда качки ψ_{max} (3 %)	Параметры качки			Модель нерегулярного волнения δ (95 %)	Модель регулярного волнения δ
	Частота максимума волнения ω_m, $с^{-1}$	Высота волн $h_{3\,\%}$, м	Частота максимума энергии качки ω, $с^{-1}$		
0,01	12,5	0,25	11,4	(0,08, 0,24)	0,07
0,61	1,5	1	1,36	(0,58, 0,86)	0,5
0,99	1,39	1,5	1,3	(0,85, 1,1)	0,75
1,74	1,06		1,03	(0,9, 1,25)	0,85

СПИСОК ЛИТЕРАТУРЫ

1. **Вентцель Е.С.** Теория вероятностей. М.: Наука, 1964. 576 с.

2. **Маков Ю.Л.** Качка судов: Учеб. пособие. Калининград: Изд-во ФГОУ ВПО "КГТУ", 2007. 321 с.

3. **Парс Л.А.** Аналитическая динамика / Пер. с англ. К.А. Лаурье. М.: Наука, 1971.

4. **Ильина А.В., Силаев П.К.** Численные методы для физиков-теоретиков / Институт компьютерных исследований. М.; Ижевск, 2003. 132 с.

5. **Крылов Ю.М.** Спектральные методы исследования и расчета ветровых волн. Л.: Гидрометеоиздат, 1966. 255 с.

6. **Крылов А.Н.** Качка корабля // Собрание трудов. Т. XI. М.; Л.: Изд-во АН СССР, 1951. 496 с.

УДК 658.012.011.55:658.512

В.С. Нагорный

БАЗОВЫЕ РЕЦЕПТУРЫ НАНОСТРУКТУРИРОВАННЫХ СЕРЕБРОМ РАБОЧИХ ЖИДКОСТЕЙ ДЛЯ ЭЛЕКТРОКАПЛЕСТРУЙНЫХ ТЕХНОЛОГИЙ

В [1, 2] предложен инновационный электрокаплеструйный метод управляемого формирования наноструктурированных проводящих топологий с заданными параметрами на диэлектрических подложках с управляемым получением монодисперсного потока капель из предварительно созданной (с использованием высокого давления 0,1–0,6 МПа) струи, структурированной наночастицами серебра рабочей жидкости.

Предложенные в [1, 2] электрокаплеструйные оборудование и технологии базируются на создании (с использованием микропроцессоров или микроЭВМ) линейной последовательности потока монодисперсных капель рабочей жидкости (таких капель в секунду может быть сформировано до ста тысяч на одном сопле) одинакового диаметра (десятки, сотни микрометров), находящихся на одинаковом расстоянии друг от друга, сообщении (при необходимости) любой выбранной капле рабочей жидкости (из этой последовательности) заданной величины электрического заряда того или иного знака и отклонении в электрическом поле данной капли в заданную точку, определяемую технологическим процессом. Синхронизация дробления струи рабочей жидкости на монодисперсные капли производится от пьезоэлектрического преобразователя (ПЭП), который, меняя свои геометрические параметры за счет обратного пьезоэффекта под действием гармонического управляющего электрического сигнала, вносит вынужденные возмущения в вытекающую струю рабочей жидкости. При этом расстояние между каплями в линейной последовательности капель равно длине волны возмущения $\lambda = v/f$ от ПЭП, где v – средняя скорость струи рабочей жидкости, f – частота электрического напряжения на электродах ПЭП. Частота каплеобразования равна частоте f напряжения на электродах ПЭП.

Объект разработок в настоящей статье – наноструктурированные наночастицами серебра (серебро обладает наибольшей электропроводностью) рабочие жидкости электрокаплеструйных технологий управляемого от микроЭВМ нанесения проводящих топологий на различные диэлектрические носители.

К выбираемым компонентам создаваемых рабочих жидкостей предъявляются многообразные (иногда взаимно противоречивые [1, 2]) требования. Рабочие жидкости должны обеспечить: требуемые электрофизические свойства (вязкость, поверхностное натяжение, электропроводность); совмещение каждого из компонентов рабочей жидкости друг с другом; снижение испарения компонентов рабочих жидкостей из сопла и струи; необходимое pH; биозащищенность; исключение развития микроорганизмов, выпадания наночастиц металла и коагуляцию коллоидов; хорошую адгезию к материалу подложки, на который они наносятся (стекло, полимер, керамика и т. п.); высокое качество отпечатков; предотвращение фоторазрушения отпечатка на подложках из различных материалов.

Проведен патентный поиск и библиографический анализ (с учетом результатов [3–5]) используемых в электрокаплеструйных устройствах рабочих жидкостей (РЖ) с позиций сформулированных выше требований к основным параметрам рабочих жидкостей и методов их обеспечения.

Рабочие жидкости делятся на две группы – на водной и органической основе. Основными компонентами РЖ являются растворитель, краситель и полимер. В РЖ используются полярные

органические растворители, обладающие ярко выраженной (в той или иной степени) растворяющей способностью в отношении поливинилхлорида (ПВХ). Растворитель составляет около 80 % от общей массы в РЖ, поэтому цена растворителя сильно отражается на РЖ. Более качественный растворитель – многокомпонентный и его можно получить путем комбинации нескольких растворителей, гораздо более дешевых. Это обусловлено тем, что некоторые жидкости способны при смешивании образовывать азеотропные смеси. Азеотропные смеси – это смеси, полученные путем смешивания нескольких различных жидкостей, обладающие собственными физико-химическими свойствами, и их невозможно разделить обычными путями на составные компоненты, при этом в жидкости не происходит химической реакции и не образуется новых веществ. Лучшие результаты получаются у 3–4-компонентных растворителей. Хотя в подавляющем большинстве производители ограничиваются двумя компонентами. Выбором компонентов растворителя обеспечивается и пожаровзрывобезопасность РЖ. С учетом большой токсичности растворителей обычно их предельно допустимая концентрация в РЖ составляет 5–1000 мг/м^3. С целью увеличения безопасности здоровья людей и окружающей среды используются низкотоксичные составы растворителей, которые обладают несколько худшими свойствами и более высокой ценой. Для компенсации потерянных свойств в них вводят 5–20 % сильных растворителей, хотя и обладающих повышенной токсичностью, но в целом такие РЖ гораздо менее вредны, нежели приготовленные на основе только сильных растворителей.

Для электрокаплеструйных головок (ЭКСГ) с большой частотой каплеобразования больше подходит летучий растворитель с повышенной скоростью испарения, чтобы до высыхания он не смог разрушить структуру носителя и ухудшить тем самым его механические свойства. В качестве компонентов растворителей используют: кетоны – наиболее агрессивные (т. е. обладающие наивысшей растворяющей способностью), средние по цене и наиболее опасные как в плане токсичности, так и пажаровзрывоопасности (например, метилэтилкетон, циклогексанон, производные пирролидона); простые эфиры – наиболее дешевые растворители со слабой растворимостью, средней токсичностью и пожаровзрывоопасностью (этилацетат, бутилацетат и т. п.); сложные эфиры – наиболее дорогие, средние по растворимости, наименее токсичные и среднепожаровзрывоопасные (эфиры этиленгликоля, пропиленгликоля, молочной кислоты, дигликоли и др.).

Выбор красителя определяется выбором цвета, а также степенью его растворимости в воде или органических растворителях.

В РЖ одним из важнейших компонентов является полимер, который обусловливает большинство их важнейших физических свойств: поведение в ЭКСГ и на подложках, прочность покрытия и адгезионные свойства, термостабильность, срок годности, условия хранения и эксплуатации, морозостойкость, теплостойкость, вязкость, поверхностное натяжение РЖ.

Часто требуется уменьшить силу поверхностного натяжения, чтобы увеличить смачивание подложки жидкостью. Обычно это достигается добавлением другой жидкости с более низким поверхностным натяжением либо при помощи добавок, которые называются смачивающими агентами, или ПАВ (поверхностно активные вещества), иногда их еще называют детергентами. От поверхностного натяжения зависит поведение РЖ на носителе. Этот параметр в большей степени регулируется специальными добавками – ПАВ (детергентами), однако начальный уровень задается растворителем и полимером.

РЖ должны обладать нормальной смачиваемостью носителя. Если поверхностное натяжение очень низкое, т. е. смачивание у РЖ высокое, то она будет растекаться на поверхности носителя. Точки растра изображения будут сливаться друг с другом, что скажется на четкости мелких элементов изображения. Если поверхностное натяжение РЖ станет слишком высоким, т. е. РЖ обладает низкой смачиваемостью, то капли РЖ начнут собираться на поверхности носителя, пространство между пикселами будет больше, увеличится время высыхания изображения.

Также следует иметь в виду и свойства поверхности материала носителя. Некоторые материалы сами по себе обладают гидрофильными или гидрофобными свойствами, что приводит к увеличению или уменьшению эффекта смачиваемости поверхности. Тогда нормальные по свойствам РЖ будут вести себя на таких материалах хуже, чем на других. Если свойства материала и РЖ совпадают, то это даст наихудший

результат, в таком случае говорят о несовместимости материала для печати и РЖ.

Поверхностное натяжение и вязкость оказывают существенное влияние не только на поведение капель РЖ на носителе, но и на формирование капли в соплах печатающих головок.

Вязкость РЖ в большей степени обусловливается используемым полимером, но чем ниже вязкость самого растворителя, тем больше полимера можно положить в конечный продукт, что хорошо для системы (РЖ) в целом.

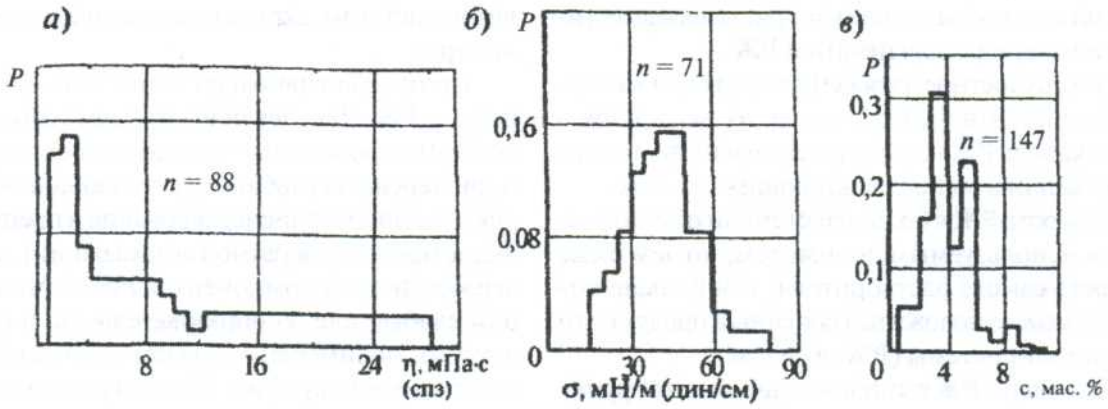
Полимер в РЖ выполняет две важные функции: связующего и пленкообразователя. Полимер как связующее (еще такие полимеры называют сшивателями) держит систему компонентов РЖ, не дает ей разваливаться, задает номинальную вязкость РЖ, обеспечивает термостабильность вязкости от изменения температуры. Чем меньше меняется вязкость от изменения температуры, тем лучше для ЭКСГ. Полимер как пленкообразователь обеспечивает: адгезию на поверхности носителя, на котором формируется изображение; механическую стойкость изображения; скорость высыхания РЖ, которая помимо летучести растворителя зависит от скорости полимеризации пленкообразователя; светостойкость изображения – помимо светостойкости самого пигмента полимер способен защитить изображение от УФ-излучения. Чем больше полимера, тем лучше для РЖ в целом как с точки зрения связующего, так и пленкообразователя. Но полимер – основной источник вязкости РЖ, а номинальная вязкость РЖ диктуется требованиями нормального функционирования ЭКСГ. При этом возникают противоречивые требования между количеством полимера и пигмента, который также оказывает существенное влияние на вязкость РЖ. Этот баланс между нормой полимера и пигмента очень сложно поймать. Но все-таки можно, если учитывать тот факт, что один и тот же полимер может иметь различный молекулярный вес, а чем он меньше, тем меньшее влияние на вязкость он оказывает в системе РЖ. Поэтому желательно использовать низкомолекулярные полимеры.

Некоторые полимеры разных групп способны образовывать соединения с иными свойствами, удовлетворяющими свойствам каждого из них в отдельности. Такие соединения называются сополимерами. Например, сейчас в РЖ активно используется сополимер винилхлорида с винилацетатом, акрилатные и стирольные сополимеры.

Сегодня активно идет поиск новых полимеров для РЖ. Перспективна группа полимеров на основе капролактана – они хорошие связующие и дисперсанты (добавки, улучшающие и ускоряющие процесс диспергирования и препятствующие процессу обратной агломерации), хорошо держат систему компонентов РЖ. Сами по себе они слабые пленкообразователи, но в смеси с другими полимерами создают очень стойкие и эластичные покрытия. Поэтому полимеры на основе капролактана используются в РЖ в смеси с другими полимерами или сополимерами. К основным достоинствам этого полимера можно отнести низкий молекулярный вес, что дает возможность поддерживать нормальный уровень вязкости РЖ при повышенном содержании полимера. В этом случае можно уменьшить уровень полимера до стандартных норм и повысить уровень пигментации либо сбалансировать содержание и полимера и пигмента, что позволяет получить более качественные РЖ по цвету, свойствам пленкообразования на носителе и по другим свойствам, отражающимся на стабильности печати и жизнеспособности РЖ.

Наряду с полимерами в РЖ вводят так называемые "модификаторы полимеров", чтобы облегчить совмещение полимера с растворителем и остальными компонентами РЖ. Растворителем РЖ на водной основе является вода, как правило, с добавками низших спиртов (метилового, этилового, пропилового), а для РЖ на основе органических растворителей – смесь спиртов, кетонов и других органических веществ (ароматический спирт, ароматический эфир, бензоловый спирт, бензоловый эфир, диметил сульфоксад и др.).

Статистический анализ значительной выборки зарубежных патентов (см. рисунок) показал, что динамический коэффициент вязкости большинства РЖ (61 %) не превышает 10 мПа·с, а почти у трети (28 %) она менее 3 мПа·с. У подавляющей части РЖ (83 %) величина коэффициента поверхностного натяжения варьируется от 25 до 60 мН/м, а более чем у половины (58 %) попадает в еще более узкий интервал от 30 до 50 мН/м. Концентрация красящего компонента (красителя или пигмента) в большинстве случаев (84 %) не превышает 5 мас. %, концентрация полимерных или олигомерных связующих близка



Результаты статистического анализа параметров из выборки зарубежных патентов объемом n:
$а$ – динамический коэффициент вязкости; $б$ – коэффициент поверхностного натяжения; $в$ – концентрация красителя (p – относительная частота)

к предельной концентрации существования истинного раствора этих соединений (0,5–10 мас. %). Это накладывает жесткие требования на качество используемых компонентов, так как в тонком слое полученного отпечатка (0,05–1 мкм) необходимо достичь высокой оптической плотности, адгезии и химической стойкости.

Для обеспечения управляемой электризации капель требуется удельное объемное сопротивление менее 2,5 кОм·см (чаще 25–1000 Ом·см). Поэтому при использовании растворителей, а также деминерализованной воды необходимо вводить добавки электропроводности. Такими добавками в количестве 0,5–2,5 мас. % являются неорганические и органические ионизирующиеся соединения: карбонат и гидрокарбонат натрия, нитрат и хлорид лития, тиоцианат калия или натрия, тиосульфат натрия, соли четвертичного аммония и квартеризованных аминов, в том числе мочевины и ее производных, дигидроацетат натрия.

Испарение компонентов РЖ может повлечь изменение pH, возрастание вязкости и поверхностного натяжения, подтолкнуть развитие микроорганизмов, выпадание проводящих добавок и коагуляцию коллоидов. Для снижения испарения из сопла и струи вводят замедлители испарения, т. е. вещества с низким значением давления насыщенных паров: гликоли (этилен- и диэтиленгликоль, глицерин) и их эфиры; низкомолекулярные полимерные гликоли и их эфиры; N-метил-2-пирролидон; сульфолан. Эти вещества регулируют вязкость, поверхностное натяжение и электропроводность, т. е. выступают многофункциональными компонентами.

Для стабилизации показателей вязкости, поверхностного натяжения, электропроводности, испарения и других факторов рекомендуется использовать соли лития, имеющие наибольшую растворимость по сравнению с щелочными металлами.

Характерное проявление начавшейся коагуляции молекулярных растворов – забивание сопла или изменение направления струи. Поэтому молекулярные растворы фильтруются через полимерные фильтры тонкой фильтрации с эффективным диаметром пор 0,2–1 мкм (размер тонких дисперсий и бактерий). Дисперсии фильтруют через более грубые фильтры или используют фугование, обеспечивающее требуемую дисперсность. Главная причина образования коллоидных частиц (1–200 нм) заключается в присутствии катионов поливалентных металлов (кальций, магний, железо), кремния, фосфат-, хлорид- и сульфат-анионов, а иногда даже катионов натрия. Недопустимыми считаются следующие концентрации ионов: хлоридов и сульфатов – 0,003–0,3 %, фосфата и магния – не более $4 \cdot 10^{-4}$ %, натрия – 10^{-2} %, железа – $(4–100) \cdot 10^{-4}$ %, кремния – $(5–60) \cdot 10^{-4}$ %, кальция – $(2–30) \cdot 10^{-5}$ %, а суммарное содержание поливалентных катионов – $(2–10) \cdot 10^{-3}$ %.

Фильтрование раствора загрязненного красителя через обратноосмотические мембраны с диаметром пор 0,1–1 нм позволяет разделить молекулы по их величине и получить чистый ра-

створ красителя. Для улучшения стабильности РЖ целесообразно заместить катионы щелочных металлов сульфогрупп красителя (чаще калия или натрия) ионами лития или квартеризованными аминами путем катионного обмена или обратного осмоса.

Биозащита РЖ осуществляется повышением или понижением pH среды путем введения кислоты (как правило, уксусной или соляной) либо щелочи (гидроокиси лития или аммония). Обычно pH водных композиций составляет 1,6–4,2 или 8,0–10,5. Однако лучше использовать щелочные среды, так как в них по сравнению с кислотными может развиваться меньшее количество микроорганизмов.

Возрастание требований к качеству отпечатков заставляет разрабатывать специальные акцепторные слои на носителях с целью формирования пористого слоя регулируемой шероховатости и с требуемой высокой поверхностной энергией. Хорошее качество дают силикатные адгезионные слои, которые обеспечивают микроструктуру типа "вертикальных" пор, способствующих впитыванию, а не растеканию. В качестве адгезионного слоя на полярных поверхностях применяют сополимер.

Для усиления фиксации отпечатка обеспечивают не только сорбционное, но и химическое взаимодействие компонентов РЖ с компонентами акцепторного слоя. В адгезионном слое непористой поверхности может содержаться низкомолекулярный сшиватель входящих в РЖ олигомера и полимера или полимер, сшиваемый входящим в состав РЖ сшивателем, или реакционноспособные группы не до конца отвержденного акцепторного слоя. Особенно большие трудности возникают при закреплении отпечатка на прозрачных полимерных пленках (полиэфирных, полиамидных). Непористая поверхность высококристаллических полимерных тел с низкой поверхностной энергией требует закрепления на них слоев с высокой поверхностной энергией.

Введение в РЖ высокомолекулярных пленкообразующих веществ или олигомеров, способных к сшиванию под действием низкомолекулярных веществ или физических факторов, существенно улучшает адгезионные и другие эксплуатационные характеристики отпечатков. В качестве пленкообразующих чаще всего используют сополимеры акрилатов и метакрилатов (бутиловые эфиры акриловой и метакриловой кислоты, метил- и этилметакрилат, ионизирующиеся соли акриловой и метакриловой кислоты). В последнем случае дополнительная прочность достигается последующим сшиванием полимерных кислот при повышенной температуре. В качестве гидрофобного сомономера с акрилатами наиболее часто используют стирол и акрилонитрил.

Для улучшения прочностных характеристик отпечатка после его нанесения на носитель можно увеличить молекулярную массу входящих в него компонентов путем сшивания: радикальной полимеризацией низкомолекулярных акрилатов при комнатной либо повышенной температуре, или поликонденсацией функциональных олигомеров (в основном эпоксидных и фенолформальдегидных), или взаимодействием кислотных функциональных групп высокомолекулярных полимеров. Полимерный компонент может сильно ухудшить вязкостные свойства РЖ, если он вводится в виде раствора, поэтому часто полимерное связующее вводят в виде стабилизированной дисперсии с учетом функциональных требований композиции. В пигментированных композициях в качестве связующего используют многофункциональные полимеры, которые также стабилизируют систему.

Большой класс задач охватывает запечатывание неабсорбирующих поверхностей без нанесения специального акцепторного слоя. Как правило, замаркировать стекло, керамику и металлы можно РЖ, включающими полимерные связующие.

Кроме роли связующего полимеры иногда выполняют и функции неионогенных ПАВ-композиций. Чаще всего используют полиэтиленгликоль (полиоксиэтиленгликоль, карбовакс) с молекулярной массой 200–3000 Да и его эфиры с муравьиной или жирными кислотами, и наиболее часто – сорбитали (твины). Сорбитали представляют собой частный случай алкилирования полиэтиленгликолей различными ангидридо-сахарами: арабинозой, глюкозой, фруктозой, сахарозой. Эти производные меньше, чем исходный полигликоль, понижают давление паров воды в РЖ, что сокращает время сушки отпечатка.

Для предотвращения отпечатка от фоторазрушения в состав РЖ вводят антиоксиданты и УФ-абсорберы.

Традиционна следующая совокупность свойств, например, маркировочной РЖ [4]: вязкость – до 20 мПа·с; поверхностное натяжение – не ниже 20 мН/м; объемное электрическое сопротивление – не выше 1000 Ом·см; дисперсность – наличие частиц дисперсной фазы размером на порядок меньше, чем диаметр используемого соплового элемента; стабильность – возможность хранения (статическая стабильность) и рециркуляции (динамическая стабильность) без потери свойств, необходимое для потребителя время.

Обычно суммарная составляющая всех добавок (пластификаторы, ПАВ, пеногасители, дегазаторы, консерванты, различные реологические добавки) не превышает 10 %, при этом они составляют вместе около 20 % стоимости РЖ.

Характерная тенденция при разработке в последние годы РЖ – переход на новые безопасные виды РЖ, а при разработке добавок – стремление создать универсальные добавки, регулирующие все свойства РЖ посредством добавления одной.

На основании полученных выше результатов исследований выбрана базовая рецептура создаваемых новых рабочих жидкостей электрокаплеструйных технологий на основе наночастиц серебра. Причем при ее разработке учтен отечественный и зарубежный опыт создания электрокаплеструйных композиций с колорантами для других задач и технологий.

Все колоранты делятся на растворимые и нерастворимые. Первые принято называть (см. выше) красителями, вторые – пигментами. Для нас при решении поставленных задач более интересны нерастворимые колоранты, т. е. пигменты. Пигменты – нерастворимые вещества, способные при определенных физических условиях образовывать взвесь частиц в растворе (коллоидный раствор).

В первом приближении можно считать, что наночастицы серебра в создаваемых новых рабочих жидкостях как бы эквивалентны частицам нерастворимого пигмента в существующих жидкостях. При этом понимаем, что процессы их взаимодействия с другими компонентами РЖ могут быть существенно различны. Но такой подход дает стартовые преимущества при разработке и создании новой наноструктурированной рабочей жидкости с наночастицами серебра и позволяет с максимальной пользой использовать накопленный опыт.

Разработан новый способ получения наноструктурированных серебром рабочих жидкостей с использованием процессов диспергирования на основе дезагрегации дисперсантами наночастиц серебра в системе органический растворитель (растворители) – полимер. Основными компонентами разрабатываемых рабочих жидкостей являются растворитель, нанопорошок серебра, дисперсант и полимер, которые выбираются из вышеприведенных положений. При этом процентное весовое содержание серебра может составить 20–40 %, а размер частиц серебра $40 \cdot 10^{-9} – 100 \cdot 10^{-9}$ м (40–100 нм).

Работа выполнена в соответствии с проектом 2.1.2/6494 АВЦП "Развитие научного потенциала высшей школы".

СПИСОК ЛИТЕРАТУРЫ

1. **Нагорный. В.С., Нагорный И.В.** Требования к параметрам рабочих жидкостей электрокаплеструйных технологий // Системный анализ в проектировании и управлении: Тр. XIII Междунар. науч.-практ. конф. СПб., 24–26 июня 2009. СПб.: Изд-во СПбГПУ, 2009. Ч. 2. С. 235–238.

2. **Нагорный. В.С.** Рабочие жидкости электрокаплеструйных технологий на базе наночастиц серебра: Области применения, требования // Вакуумная наука и техника: Матер. XVI науч.-техн. конф. с участием зарубежных специалистов / МИЭМ. М., 2009.

3. **Нагорный В.С.** Электрокаплеструйные регистрирующие устройства. Л.: Машиностроение, 1988. 269 с.

4. **Безруков В.И.** Научно-технические основы и аппаратное обеспечение автоматизированной электрокаплеструйной маркировки изделий: Дис. ... докт. техн. наук / Науч. конс. В.С. Нагорный; СПбГПУ. СПб., 2003. 537 с.

5. **Марков О.** Сольвентные чернила: Ваш выбор... / Компания САН. Новосибирск, 2006.

УДК 378.4

Ю.А. Ветринский

ОПЫТ ИСПОЛЬЗОВАНИЯ ТЕХНОЛОГИИ СМЕШАННОГО ОБУЧЕНИЯ В ПРАКТИКЕ ПРЕПОДАВАНИЯ ТЕХНИЧЕСКИХ ДИСЦИПЛИН

Особенность внедрения форм и методов электронного обучения в вузовскую практику состоит в придании учебному процессу смешанного характера, сочетающего аудиторные занятия и самостоятельную работу студентов с сетевыми источниками информации. При этом эффективность обучения во многом зависит от того, насколько рационально распределена учебная нагрузка между аудиторной и внеаудиторной составляющими. Единства мнений по вопросам распределения нагрузки к настоящему времени не достигнуто – диапазон существующих подходов простирается от полного отрицания аудиторно-лекционной формы обучения и отведения преподавателям роли разработчиков электронных онлайн-курсов до сведения всей "электронной" части учебного процесса к визуальным эффектам на лекциях и лабораторным занятиям с привлечением компьютерных моделей.

Попытки научного анализа эффективности технологий смешанного обучения (blended learning) с выработкой соответствующих рекомендаций в отечественной практике были предприняты сравнительно недавно. В известных работах обобщены два основных подхода к организации смешанного учебного процесса в высшем учебном заведении [1, 2]:

основной материал изучается в рамках дистанционных курсов, а его отработка и закрепление проводится на очных занятиях;

весь материал изучается на очных занятиях с привлечением сетевых источников информации в синхронном или асинхронном режимах.

Реализуемые в соответствии с этими подходами модели смешанного обучения достаточно универсальны и уже внедрены в некоторых отечественных вузах [3], однако утверждать, что данные модели стали эталонами для всей системы отечественного высшего образования, очевидно, преждевременно.

В Центре технологий электронного обучения СПбГПУ разрабатывается подход к организации смешанного обучения, основанный на результатах экспериментальных исследований, проводимых в рамках Федеральной целевой программы "Развитие научного потенциала высшей школы" в 2007–2009 годах.

Постановка эксперимента

Задача проводимых исследований включала экспериментальную отработку технологий доставки электронного образовательного контента. При этом под технологиями доставки понимались не только программно-аппаратные технологии дистанционного доступа к учебным материалам, но и способы доведения обучающей информации, позволяющие обеспечить усвоение студентами необходимого объема знаний и навыков.

Содержание эксперимента предусматривало преподавание дисциплины "Спутниковые системы связи" на радиофизическом факультете СПбГПУ в течение семестра по технологии смешанного обучения. По итогам семестра методика доставки электронного образовательного контента уточнялась и корректировалась, после чего проверялась на эффективность в следующем учебном году.

В эксперименте участвовали две академические группы кафедры "Радиотехника и телекоммуникации" СПбГПУ, одна из которых привлекалась как контрольная группа, в которой данная дисциплина читалась по "традиционным" методикам.

Техническая поддержка эксперимента осуществлялась с использованием программной системы управления обучением (Learning Management Systems – LMS) Moodle, выбранной из представленных на рынке образования систем подобного класса по критерию цена-качество [4] и установленной на сервере Центра технологий электронного обучения.

Электронный образовательный контент дисциплины был представлен:

материалами электронных лекций и справочными материалами в форматах редакторов MHTML, MS Word и Adobe Acrobat (рис. 1);

электронными лабораторными работами на базе лицензионной программной системы компьютерного моделирования "Альбатрос-Бюджет", размещенной на сервере Центра и доступной дистанционно в режиме онлайн через диалоговую форму LMS Moodle (рис. 2);

электронной библиотекой в формате DJV, тестами для самопроверки и контрольными тестами по изученному материалу дисциплины (рис. 3).

Традиционный образовательный контент был представлен учебными пособиями по изучаемой дисциплине, которыми основная и контрольная группы обеспечивались в полном объеме через библиотеку вуза.

Внеаудиторное интерактивное взаимодействие студентов с преподавателем и между собой осуществлялось посредством предоставляемых LMS Moodle услуг электронной почты, рассылок и системы тематических форумов.

Изучаемый подход к организации смешанного обучения предусматривал изложение теоретической части курса в форме очных лекций, проводимых в учебной аудитории без привлечения в ходе занятий сетевых источников информации. Студентам не предъявлялось требования конспектировать лекции – в противовес теории "моторного" запоминания основной упор делался на рациональное осмысление материала, достигаемое в ходе живого обмена мнениями, инициируемого лектором непосредственно в процессе занятия.

Проработка и закрепление учебного материала проводились студентами во внеаудиторное время с использованием печатных учебных пособий и электронных материалов, доступ к

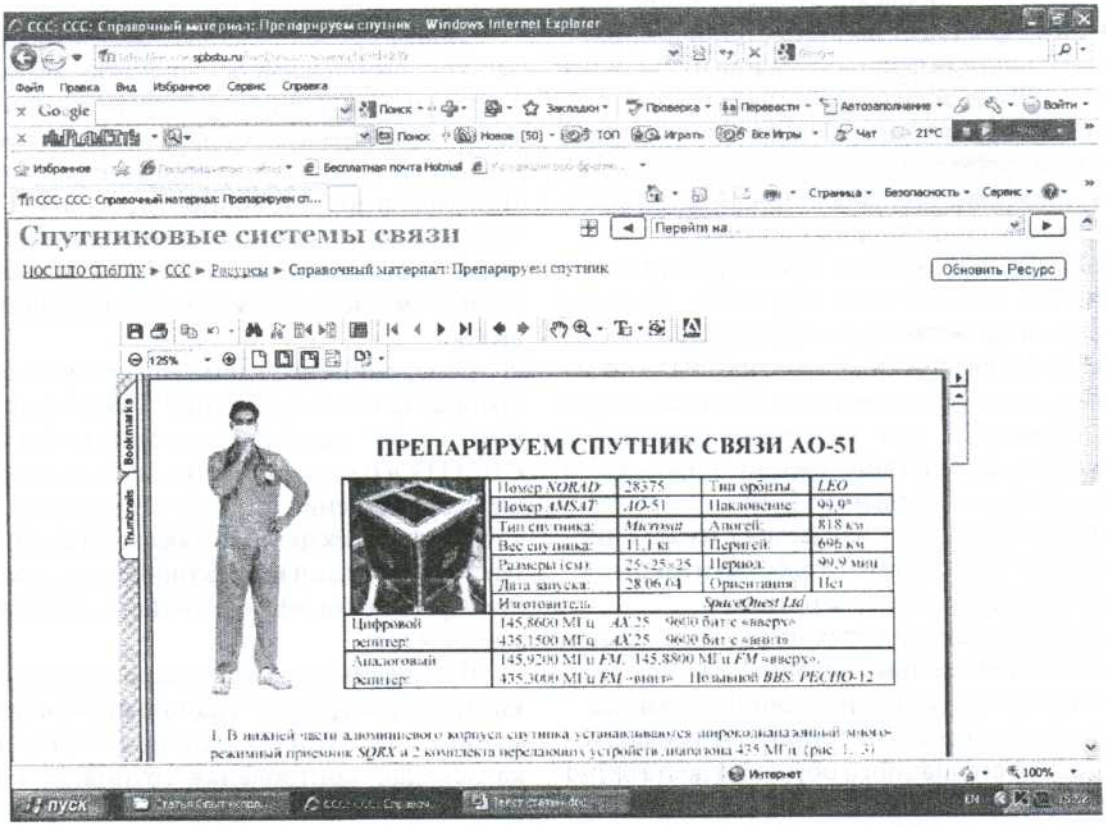


Рис. 1. Доставка учебных материалов через интерфейс LMS

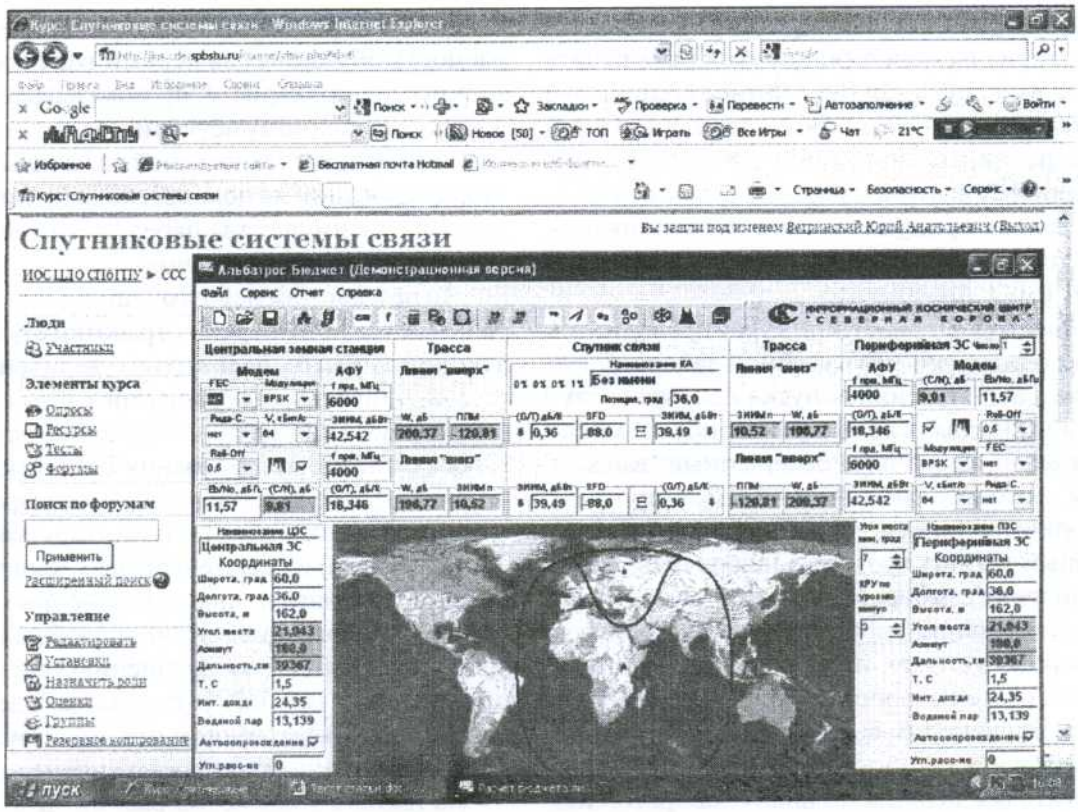


Рис. 2. Доступ к удаленным программам через интерфейс LMS

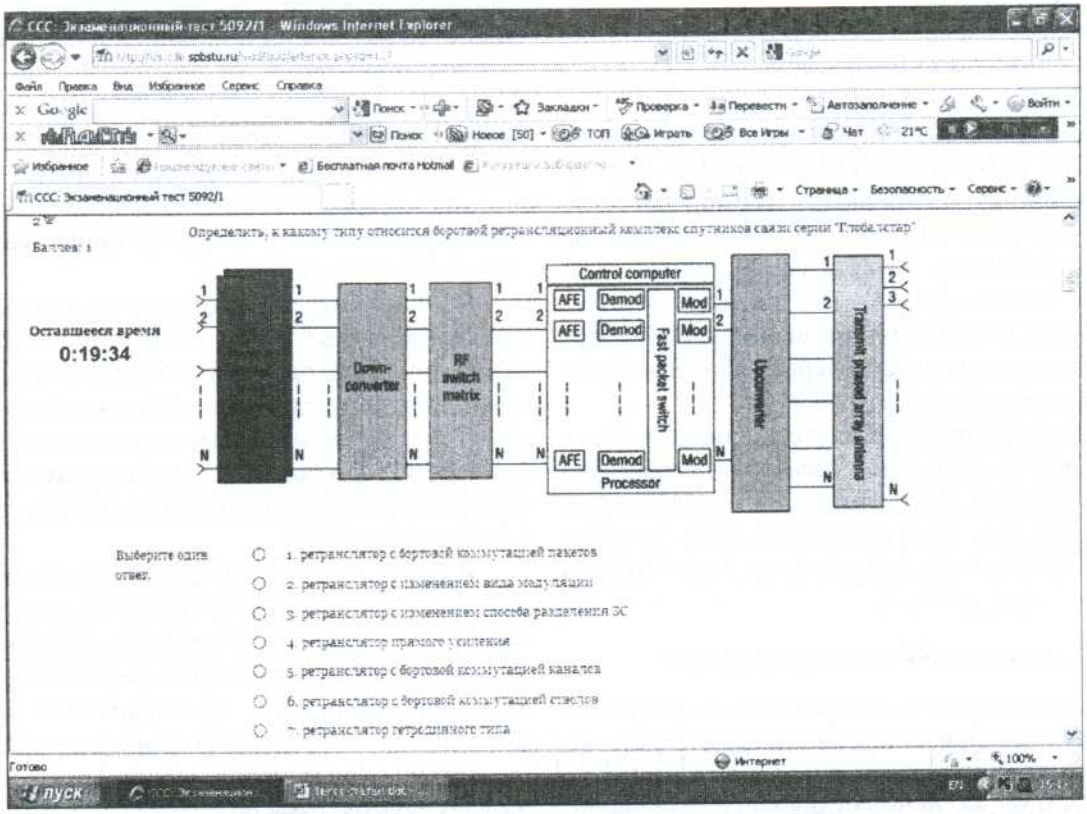


Рис. 3. Доступ к контрольным тестам через интерфейс LMS

которым осуществлялся через LMS (электронные материалы дополняли и уточняли печатное пособие, отражая последние тенденции в предметной области и предоставляя богатый иллюстративный материал). Внеаудиторное обращение студентов к материалу лекций инициировалось принудительно с использованием системы контрольного тестирования LMS Moodle [5].

Лабораторные работы по курсу проводились в форме очных занятий в учебной лаборатории. Причем для получения допуска к занятию студенты должны были во внеаудиторное время самостоятельно выполнить "облегченный" вариант работы, используя дистанционный доступ к лабораторной установке через интерфейс LMS. Подобный подход позволил исключить этапы ознакомления с лабораторной установкой и изучения содержания работы непосредственно в ходе занятия (что так или иначе происходило, если подготовка к работе осуществлялась студентами без использования LMS). В результате у студентов появлялась возможность не только выполнить все задания по плану лабораторной работы, но и оформить по ней отчет и защитить его в ходе занятия.

В качестве формальных результатов обучения использовались итоговые оценки студентов экспериментальной группы по дисциплине. Оценки формировались как среднее арифметическое из результатов контрольных тестов по материалам лекций, проводимых в процессе обучения, и результата экзаменационного теста по всему изученному материалу. Статистический анализ оценок студентов группы по всем контрольным мероприятиям проводился LMS автоматически с визуализацией результатов на каждом этапе обучения.

За период проведения исследований были проанализированы и обобщены результаты трех экспериментов, по которым могут быть сделаны первые выводы об эффективности предложенного подхода к организации смешанного обучения в техническом вузе.

Анализ и обобщение результатов

Распределение итоговых оценок среди студентов основной и контрольной групп по итогам трех семестров представлено диаграммами на рис. 4 (основная группа помечена штриховкой, контрольная – точками).

Анализ диаграммы, представляющей успеваемость обучаемых за первый год рассматриваемого периода (рис. 4, *а*), показывает, что переход к смешанному обучению в соответствии с описанной выше технологией на начальном этапе исследований не показал сколько-нибудь заметного преимущества перед "традиционной" аудиторной формой обучения.

По результатам первого этапа было принято решение увеличить время отработки студентами лекционных материалов путем увеличения числа принудительных обращений к ним с одного раза до двух. Побуждение к изучению осуществлялось системой тестирования LMS, открывающей на 24-часовой период доступ к контрольным тестам сразу после прочтения лекции (закрепление материала) и перед очередной лекцией (повторение материала). Анализ полученных результатов (рис. 4, *б*) показывает, что подобный подход приводит к заметному улучшению успеваемости экспериментальной группы, однако делать по ним вывод о качественном повышении эффективности обучения пока преждевременно.

На третьем этапе исследований было принято решение о последовательном наращивании объема контрольных тестов в процессе изучения дисциплины за счет введения дополнительных вопросов по ранее пройденному материалу. Именно этот подход показал лучшие на текущий момент результаты (рис. 4, *в*), позволяющие говорить о качественном повышении степени усвоения материала за счет внедрения технологии смешанного обучения в практику преподавания технических дисциплин.

Обобщая полученные результаты, можно сделать однозначный вывод об эффективности предложенного подхода к организации смешанного обучения, главная отличительная черта которого – изучение основного материала на очных занятиях под руководством преподавателя с последующей его внеаудиторной проработкой под управлением LMS.

В качестве наиболее эффективного элемента принятой модели смешанного обучения по результатам исследований определена система двухэтапного дистанционного тестирования обучаемых по материалам прослушанных лекций. Идеология побуждения студентов к учебе путем проведения контрольных мероприятий (опросов, коллоквиумов) сама по себе не нова, но использование технологий дистанционного обучения



Рис. 4. Успеваемость учебных групп за 2007 (*а*), 2008 (*б*) и 2009 (*в*) годы

поднимает ее на совершенно новый уровень. Дистанционное тестирование, проводимое посредством взаимодействия студентов с LMS без привязки ко времени занятия, позволяет преодолеть главные ограничения, характерные для "традиционных" форм контрольных мероприятий – дефицит учебного времени и физических сил преподавателя. Это в свою очередь дает возможность проведения глубоких проверок знаний по всему пройденному материалу при стопроцентном охвате личного состава учебной группы. Систематическое двухэтапное дистанционное тестирование обучаемых с формированием итоговой оценки за курс на базе текущих оценок за тесты проявило себя действенным стимулом, побуждающим обучаемых к самостоятельному внеаудиторному изучению пройденного материала.

Представленная в работе модель смешанного обучения может быть рекомендована к использованию в учебном процессе технических вузов как проверенная экспериментально и хорошо себя зарекомендовавшая.

СПИСОК ЛИТЕРАТУРЫ

1. **Мохова М.Н.** Активные методы в смешанном обучении в системе дополнительного педагогического образования: Дис. ... канд. пед. наук. М., 2005. 155 с.

2. **Капустин Ю.И.** Педагогические и организационные условия эффективного сочетания очного обучения и применения технологий дистанционного образования: Дис. ... докт. пед. наук. М., 2007. 419 с.

3. **Ширшов Е.В., Ефимова Е.В.** Организация учебной деятельности в вузе на основе электронных информационно-образовательных технологий: Монография. Архангельск: Изд-во Арханг. гос. техн. ун-та, 2006. 208 с.

4. **Ветринский Ю.А., Никитин А.Б., Сороцкий В.А., Цикин И.А.** Программные средства поддержки учебного процесса в информационно-образовательной среде университета // Научно-технические ведомости СПбГПУ. 2008. № 3 (60). Информатика. Телекоммуникации. Управление. С. 256–264.

5. **Ветринский Ю.А., Никитин А.Б., Сороцкий В.А., Цикин И.А.** Работа с информационными ресурсами в типовой информационно-образовательной среде: Учеб.-метод. пособие / Под ред. И.А. Цикина. СПб.: Изд-во Политехн. ун-та, 2008. 44 с.

УДК 355.4.6(2).301

Н.Б. Кунтурова

МЕТОДИКА ФОРМИРОВАНИЯ СТРУКТУРНЫХ ГРУПП КАНДИДАТОВ ПРИ ПОСТУПЛЕНИИ В МАГИСТРАТУРУ С ПРОГНОЗИРУЕМОЙ УСПЕШНОСТЬЮ ИХ ОБУЧЕНИЯ

Проблема обеспечения требуемого качества государственного кадрового заказа обусловливает необходимость построения эффективной системы управления качеством функционирования системы непрерывной подготовки кадров силовых структур. Данная система позволит на основе системного подхода получить адекватный аппарат для решения задач кадрового отбора и получения магистерского образования с учетом фактора профессиональной компетентности кадров. Это поможет на более качественном уровне решать задачи прогнозирования тенденций развития системы непрерывной подготовки специалистов силовых структур при том или ином воздействии, а также планирования процесса управления системой. Только при решении проблемы кадрового отбора кандидатов с учетом фактора профессиональной компетентности возможна разработка современной модели управления качеством системы непрерывной подготовки специалистов силовых структур, обеспечивающей эффективное выполнение государственного кадрового заказа.

Сущность процесса кадрового отбора кандидатов в магистратуру в соответствии с уровнем профессиональной компетентности состоит в таком целенаправленном изменении состояний подсистем подготовки и профессионального использования, которое обеспечивает достижение системой заданных целей функционирования – выполнение государственного кадрового заказа с требуемым качеством.

Достижение заданных целей функционирования системой непрерывной подготовки кадров силовых структур при том или ином варианте управления возможно только на основе использования методов моделирования.

Целью проведенной автором работы была разработка методики формирования структурных групп кандидатов, поступающих в магистратуру, с прогнозируемой успешностью их обучения.

На основе интегративного структурно-функционального подхода разработана методика формирования набора кандидатов для получения магистерского образования с учетом фактора их профессиональной компетентности.

На первом этапе сбора информации о кандидатах их результаты, характеризующие уровень теоретической и практической подготовки,

такие, как экспертные оценки, заносятся в матрицу экспертных оценок.

На втором этапе при помощи кластерного анализа все кандидаты распределяются на три структурные группы. Они различаются: по личностным качествам; качествам, характеризующим специфические способности кандидатов (склонности) к профессиональной деятельности.

При этом целенаправленно выявлены уровни профессиональной компетентности кандидатов (высоко-, средне-, малоуспешные) с целью формирования рекомендаций руководству вуза по приему слушателей в магистратуру. Этап реализуется в следующей последовательности:

1. Определяются численные значения показателя качества подготовки кандидатов к решению профессиональных задач – коэффициента профессиональной компетентности.

2. Обрабатывается информация матрицы экспертных оценок методом кластерного анализа в качественные показатели, такие, как высоко-, средне- и малоуспешные.

3. Формулируются рекомендации руководству вуза по использованию полученных качественных показателей кандидатов для формирования набора с прогнозируемым уровнем успешности дальнейшего обучения.

Разработка методики определяется необходимостью практической реализации разработанной ранее модели (рис. 1) с целью возможно более полного учета информации, имеющейся о каждом поступающем в магистратуру, и осуществления прогноза успешности обучения каждого слушателя.

Для реализации поставленной задачи была составлена матрица Q экспертных оценок кандидатов. В нее вошли:

коэффициент профессиональной компетентности;

уровень профессиональной подготовленности;

средний балл диплома;

категория профессиональной пригодности;

коэффициент интеллекта (IQ), вычисленный с помощью теста Айзенка;

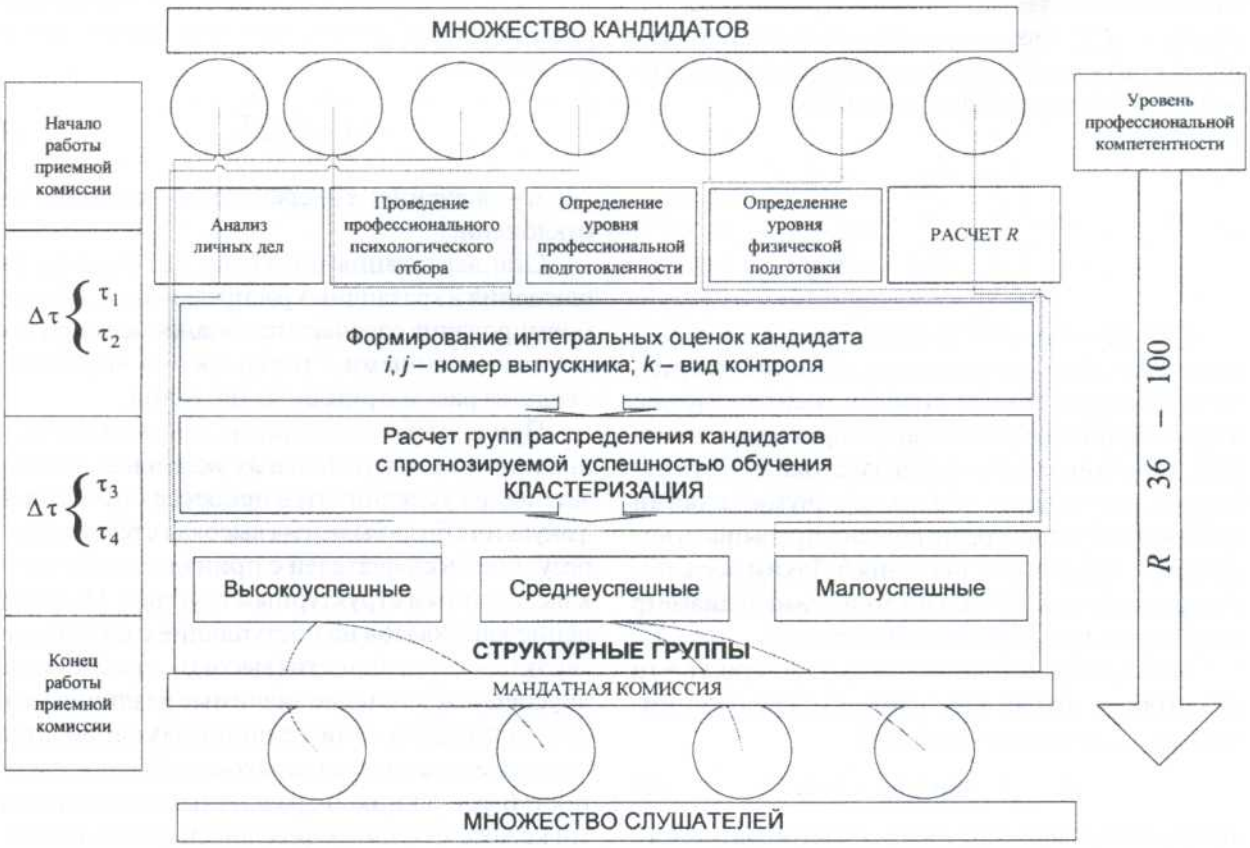


Рис. 1. Организация кадрового набора офицеров с прогнозируемой успешностью дальнейшего обучения

уровень физической подготовленности;

дисциплинарная практика (взыскания, поощрения).

При помощи метода кластерного анализа множество поступающих было разделено на три структурные группы (высоко-, средне- и малоуспешные). Кластерный анализ используется при построении разбиения, удовлетворяющего некоторому критерию оптимальности. Этот критерий представляет собой некоторый функционал, выражающий уровни желательности разбиений, который называют целевой функцией.

В качестве целевой функции была взята внутригрупповая сумма квадратов отклонения

$$W = \sum_{j=1}^{n}(x_j - \overline{x})^2 = \sum_{j=1}^{n} x_j^2 - \frac{1}{n}\left(\sum_{j=1}^{n} x_j\right)^2, \qquad (1)$$

где x_j представляет собой измерения j-го объекта.

Объекты i-й и j-й попадают в один кластер, когда расстояние (отдаленность) между точками x_i и x_j достаточно мало, и в разные кластеры, когда это расстояние достаточно велико. Значение $d(x_i, x_j)$ для x_i и x_j – расстояние между x_i и x_j и эквивалентно расстоянию между i-м и j-м объектами с соответственными характеристиками (F_1, F_2, F_3, ..., F_p). Расстояния между парами векторов $d(x_i, x_j)$ могут быть представлены в виде симметричной матрицы расстояний D:

$$D = \begin{pmatrix} 0 & d_{12} & \ldots d_{1n} \\ d_{21} & 0 & \ldots d_n \\ d_{n1} & d_{n2} & \ldots 0 \end{pmatrix}.$$

Суть данного метода в том, что два объекта – кандидата, принадлежащих одной и той же группе (кластеру), имеют коэффициент сходства, который меньше некоторого порогового значения. В терминах евклидова расстояния d это означает, что расстояние между двумя точками (объектами) кластера не должно превышать некоторого порогового значения h. Таким образом, h определяет максимально допустимый диаметр подмножества, образующего кластер.

Положив расстояние между кластером $i + j$ и некоторым другим кластером k, равным минимальному из этих расстояний

$$d_{i+j,k} = \min(d_{ik} + d_{jk}), \qquad (2)$$

получим первый шаг работы агломеративного иерархического алгоритма.

Последующие шаги аналогичны. Тогда формула перерасчета расстояний принимает вид

$$d_{i+j,k} = \frac{n_i}{n_i + n_j} d_{ik} + \frac{n_j}{n_i + n_j} d_{jk}. \qquad (3)$$

В данном случае расстояние между двумя кластерами на каждом шаге работы алгоритма оказывается равным среднему арифметическому из расстояний между всеми такими парами элементов, что один элемент пары принадлежит к одному кластеру, другой – к другому. Процесс объединения идет до тех пор, пока количество групп не станет равным трем (высоко-, средне-, малоуспешные). При этом мерой различия служит квадрат евклидова расстояния:

$$d(x_i, x_j) = \left[\sum_{k=1}^{p}\left(x_{ki} - x_{kj}\right)^2\right]^{\frac{1}{2}}. \qquad (4)$$

Чтобы устранить неоднородность измерения исходных данных, все их значения предварительно нормируются. Нормирование исходных данных для кластерного анализа исследуемого процесса проводится посредством деления исходных величин на среднеквадратичное отклонение соответствующих показателей:

$$Z_i = \frac{x_i - \overline{x}}{\sigma}, \qquad (5)$$

где x_i – значение; $\overline{x}$ – среднее; σ – стандартное отклонение.

Стандартизация позволяет сравнивать наблюдения из различных распределений. Методы нормирования означают признание всех признаков равноценными с точки зрения выяснения сходства рассматриваемых объектов.

При анализе результатов, полученных слушателями на вступительных экзаменах, и показателей их успешности в процессе обучения на факультете была выявлена высокая степень связи результатов слушателей с принадлежностью их к выделенным структурным группам. Исследование кандидатов на поступление с различным характером успешности (высоко-, средне- и малоуспешные) выявило значимые различия между слушателями из числа принятых кандидатов, что свидетельствует о широком разбросе возможностей выпускников с различными показателями коэффициента профессиональной компетентности.



Рис. 2. Реальная успешность в обучении
$K_{тм}$ – критериальная тестовая мера



Рис. 3. Методика кадрового отбора кандидатов с прогнозируемой успешностью обучения

Рассмотрим реальный процесс: при осуществлении приема на обучение в магистратуру претендентов 80 человек, набирают же всего 50 слушателей (рис. 2).

1. При наборе без методики формирования структурных групп кандидатов на поступление в магистратуру с прогнозируемой успешностью результата их обучения взяли бы всех – 80 человек. Из них 42 + 14 = 56 человек будут успешно обучаться, т. е. коэффициент успешности $\gamma = 56/80 = 0{,}7$.

2. При наборе с учетом методики формирования структурных групп кандидатов на поступление в магистратуру с прогнозируемой успешностью результата их обучения: 42 + 8 = 50; из них успешно обучающиеся составляют 38 человек, а коэффициент $\gamma = 42/50 = 0{,}84$.

Методика распределения кандидатов на поступление в магистратуру на структурные группы значительно увеличивает качество приема на факультет.

Принадлежность слушателей к структурным группам – важный фактор, определяющий успех обучения и воспитания в магистратуре вузов силовых структур.

Значительное место в методике (рис. 3) занимает выбор способа формализации информации, полученной о каждом из кандидатов. В качестве аппарата формализации основных отношений структуры используется матричный метод.

Исследование кандидатов на поступление с различным характером успешности (высоко-, средне- и малоуспешные) выявило значимые различия между слушателями из числа принятых кандидатов, что свидетельствует о различных возможностях выпускников с разными показателями коэффициента профессиональной компетентности.

Предлагаемая методика обеспечивает структурирование группы кандидатов, принимающих участие в отборе, на группы (высоко-, средне-, малоуспешные). Существенное отличие предлагаемого подхода к набору слушателей в магистратуру вузов силовых структур – использование современного математического аппарата обработки информации.

СПИСОК ЛИТЕРАТУРЫ

1. **Беспалько В.П.** Элементы теории управления процессом обучения. Ч. 1, 2. М.: Знание, 1980. 81 с.

2. **Суходольский Г.В.** Математико-психологические модели деятельности. СПб., 1994. 62 с.

УДК 355.4.6(2).301

Н.Б. Кунтурова

ОСОБЕННОСТИ ОЦЕНИВАНИЯ ПРОФЕССИОНАЛЬНОЙ КОМПЕТЕНТНОСТИ БУДУЩИХ СПЕЦИАЛИСТОВ В ПРОЦЕССЕ ИХ ОБУЧЕНИЯ В ВУЗЕ

Развитие профессиональной компетентности будущего специалиста происходит в процессе его обучения и саморазвития в вузе. Изучение процессов и результатов собственной деятельности обусловливает образование у обучаемых свойств, способствующих формированию их готовности к самостоятельному решению предстоящих задач, требующих компетентности и мастерства. Профессиональная компетентность как определенный процесс имеет стадии, или фазы: начало, развитие (экстенсивное, интенсивное, стагнация, деградация и др.), окончание.

В современной психолого-педагогической литературе вопрос исследования профессиональной компетентности связывают с уровнем подготовки специалиста. Причем и уровень подготовки определяют как уровень усвоения компетентности: максимально возможный, достаточный, необходимый. Но измерений уровня профессиональной компетентности, ос-

нованных на параметрических методах исследования, нет.

Целью авторского исследования была разработка эффективного научно-методического аппарата оценивания профессиональной компетентности будущих специалистов силовых ведомств в процессе их вузовского образования.

Рассматривая процесс обучения в вузе как некоторую операцию Q, состоящую из последовательных этапов – семестров, можно представить результаты семестровых экзаменов как контроль процесса обучения.

В качестве исходных данных могут выступать: количество обучаемых, время обучения и множество возможных результатов сдачи семестровых экзаменов. Характеристики этих случайных величин могут быть получены статистическим путем. Пусть факторы ξ, влияющие на результаты сдачи экзаменов, представляют собой случайные величины с вероятностными характеристиками, тогда показатель эффективности W, зависящий от них, также величина случайная. При таких условиях максимизировать случайную величину невозможно, она остается случайной и неконтролируемой. Если заменить ξ их средними значениями, задача становится детерминированной и может быть решена. Факторами, от которых зависит успешность операции, также являются: заданные, заранее известные факторы α (условия выполнения операции); зависящие от обучаемого решения x.

На успех операции Q можно влиять, выбирая решение x, при этом случайные факторы ξ и α могут быть заменены средними значениями, тогда показатель эффективности W зависит от этих случайных величин и случаен. В качестве показателя эффективности выберем среднее значение этой случайной величины $\overline{W} = M[W]$ и такое решение x, при которых этот усредненный по условиям показатель обращается в максимум:

$$\overline{W} = M\left[W\left(\alpha, x, \xi\right)\right] \Rightarrow \max. \qquad (1)$$

Эффективность каждой отдельной операции, проводимой при конкретных значениях случайных факторов α и ξ, может сильно отличаться от ожидаемой как в большую, так и меньшую сторону. При оптимизации операции "в среднем" из-за большого числа повторений происходит компенсация ее результата.

Параметры процесса Q можно описать. На каждом шаге выбирается решение, от которого зависит эффективность данного шага по отношению к эффективности операции в целом. Примем эти шаговые решения за $x_1, x_2, x_3, \ldots, x_m$:

$$X = (x_1, x_2, x_3, \ldots, x_m).$$

Каждое шаговое решение – суммарный результат, который зависит от переменной y_{ik} – экзаменационной оценки, принимающей возможные значения из множества {3; 4; 5}. Управление в данном случае – это выбор одного из решений. Если X^* – решение, при котором достигается максимальная эффективность

$$W = \sum_{i=1}^{m} w_i \Rightarrow \max, \qquad (2)$$

то это оптимальное решение. Оно состоит из совокупности оптимальных шаговых решений:

$$X^* = (x_1^*, x_2^*, x_3^*, \ldots, x_m^*). \qquad (3)$$

Максимальная эффективность в этом случае

$$W^* = \max_{x \in X} \{W(x)\}. \qquad (4)$$

Чтобы диагностировать качество результатов деятельности обучаемых, нужен оценивающий показатель – обобщенная интегральная оценка – мера качества, выражающая степень достижений поставленных целей с учетом совокупных затрат. В качестве такого результирующего показателя качества обучения может служить коэффициент профессиональной компетентности R. Это количественный дифференциальный показатель профессионального потенциала каждого выпускника. Для расчета коэффициента использовались результаты сдачи экзаменов, рассматриваемые в качестве экспертных оценок. Оценивались теоретические знания, умения и практические навыки, приобретенные выпускником в процессе обучения и стажировок. Если оценка по дисциплине y_{ik}, коэффициент нагрузки по дисциплине $r = t \cdot 100\,\%$, где t – время, отводимое на изучение дисциплины, то параметр качества усвоения учебной информации $g = r \cdot y_{ik}$. Тогда фактическая эффективность каждого обучаемого

$$W_{\text{факт}} = 12/5\left(\sum r_{i3} \cdot 3 + \sum r_{i4} \cdot 4\right) + \sum r_{i5} \cdot 5, \qquad (5)$$

где $\sum r_i = \sum_{j=1}^{n} g_j$ при $i = \overline{3,5}$. Максимально возможный результат эффективности обучения при этом составит

$$W_{\max} = \sum r \cdot 5. \qquad (6)$$

Тогда можно вычислить коэффициент оценивания профессиональной компетентности каждого выпускника:

$$R_i = \frac{W_{\text{факт}}}{W_{\max}} = \frac{12/5\left(\sum r_{i3}\cdot 3 + \sum r_{i4}\cdot 4\right) + \sum r_{i5}\cdot 5}{\sum r\cdot 5} 100\ \%. \quad (7)$$

Если принять общее число предметов, по которым выставляется экспертная оценка, за n, а число обучаемых в конкретной группе за p, то количество принятых решений будет X_m, где m – число шагов (число сессий). Тогда можно рассчитать величину суммарного коэффициента профессиональной компетентности для всей группы. В этом случае эффективность конкретной группы обучаемых составит

$$W_m = \sum_{j=1}^{m} x_j. \quad (8)$$

Величина коэффициента профессиональной компетентности группы обучаемых

$$R = \frac{W_m}{W_{opt}}, \quad (9)$$

где $W_{opt} = p\Sigma r\cdot 5$ – максимально возможная эффективность группы.

В системе с дискретными состояниями можно оценить динамический коэффициент компетентности группы в целом на каждом шаге j:

$$R_j = \frac{W_{\text{гр}}}{W^*_{\text{opt гр}}}. \quad (10)$$

По своей сути коэффициент профессиональной компетентности R есть оценка меры продуктивности профессиональной деятельности выпускника. Эта формула дает возможность вычислять коэффициент профессиональной компетентности для исследуемой выборки обучаемых. Алгоритм его вычисления представлен на рис. 1.

На основании проведенного экспертного и кластерного анализа отзывы на выпускников были разделены на пять уровней (I–V), соответствующих различным показателям качества профессиональной подготовки. В процессе исследования выявлена положительная связь между коэффициентом профессиональной компетентности каждого конкретного выпускника, вычисленного на основании результатов обучения в вузе, и реальной успешностью его профессиональной деятельности на первичной должности.

В процессе исследования при анализе выпуска (125 чел.) рассматривался закон распределения наблюдаемой случайной величины R. Были получены результаты, свидетельствующие о том, что коэффициент профессиональной компетентности R как случайная величина имеет нормальное распределение. Значения коэффициента профессиональной компетентности изменяются в границах от 100 до 36 единиц.

На основании проведенных расчетов можно приближенно определить гипотетический закон распределения случайной величины – коэффициента профессиональной компетентности R и проверить гипотезу о виде закона распределения.

Данные расчета коэффициента профессиональной компетентности R были представлены в сгруппированном виде, и осуществлено сравнение эмпирической и теоретической функции распределения $F(x)$, для этого был использован критерий χ^2.

Графики эмпирической и теоретической функции распределения изображены на рис. 2.

Критерий χ^2 позволил осуществлять проверку гипотезы о согласии: вычисленное значение $\chi^2_{\text{эмп}} = 122{,}56$, критическое значение для уровня значимости $a = 0{,}05$ $\chi^2_{\text{крит}} = 157{,}61$, следовательно, $\chi^2_{\text{эмп}} < \chi^2_{\text{крит}}$.

Можно сделать вывод о том, что $F(x) = F_N(X;\overline{Q})$, т. е. полученное эмпирическое распределение не отличается от нормального. Для увеличения совпадения с нормальным распределением данного эмпирического распределения случайной величины – коэффициента профессиональной компетентности R – необходимо увеличить выборку выпускников.

Таким образом, можно утверждать, что коэффициент профессиональной компетентности R как случайная величина имеет нормальное распределение.

При сравнении распределения, полученного опытным путем, выпускников с рассчитанным коэффициентом профессиональной компетентности $R(2)$ по уровням профессиональной компетентности со значениями нормального распределения для коэффициента профессиональной компетентности $R(1)$ (рис. 3) обнаружено их соответствие друг другу.

В выборке выпускников среднее арифметическое $M = 68$, а стандартное отклонение $\sigma = 10{,}67$.

Нормальное распределение характеризует такие случайные величины, на которые воздей-

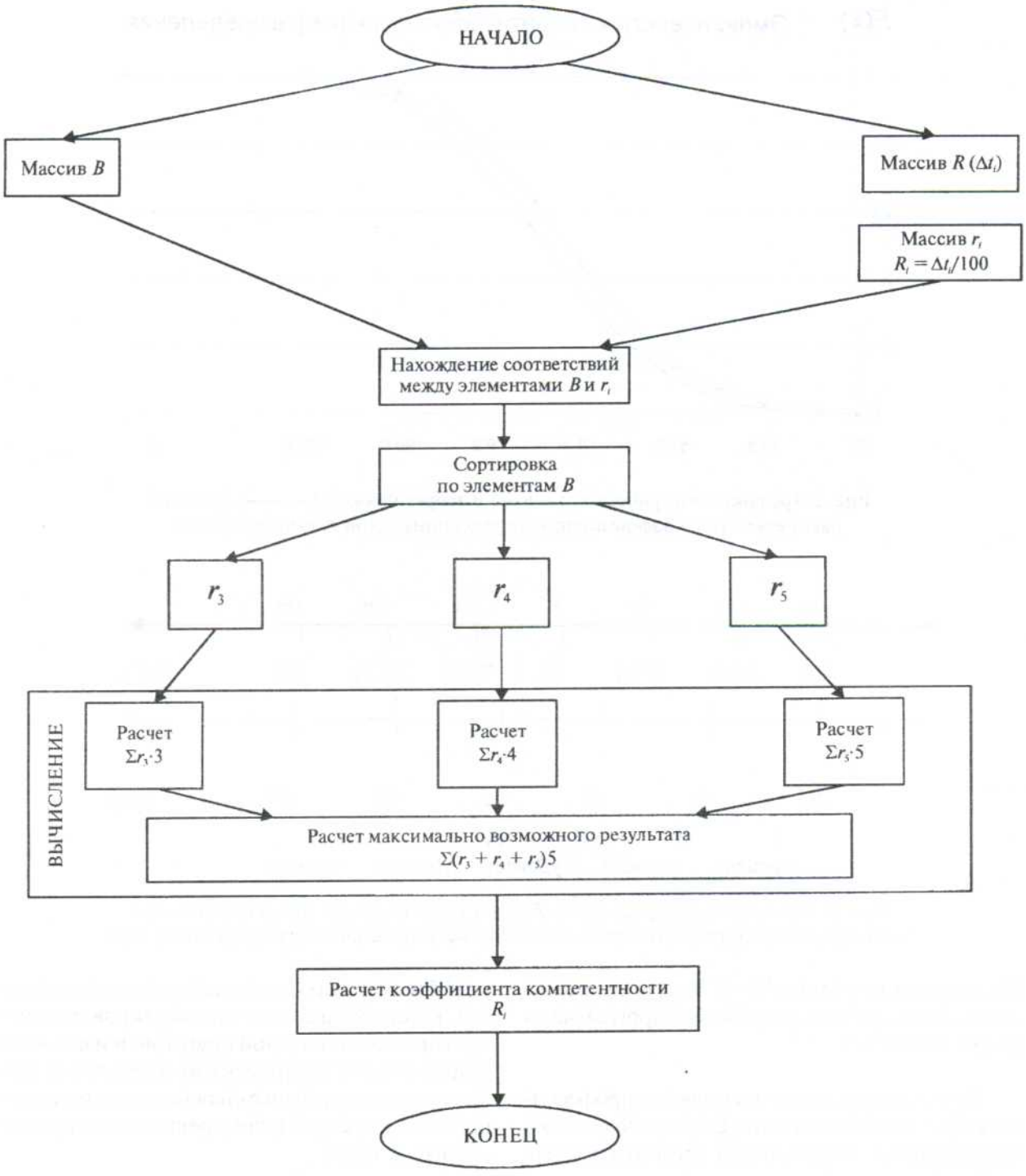


Рис. 1. Алгоритм расчета коэффициента профессиональной компетентности

ствует большое количество разнообразных факторов, причем сила воздействия одного отдельно взятого фактора значительно меньше суммы воздействий остальных факторов. В результате получается, что чаще наблюдаются некоторые средние значения измеряемого параметра, реже крайние, и чем сильнее отличается какое-то значение от среднего, тем реже оно встречается.

Используя свойства нормального распределения, можно определить прогнозируемый процент выпускников, имеющих коэффициент профессиональной компетентности более 75 единиц.

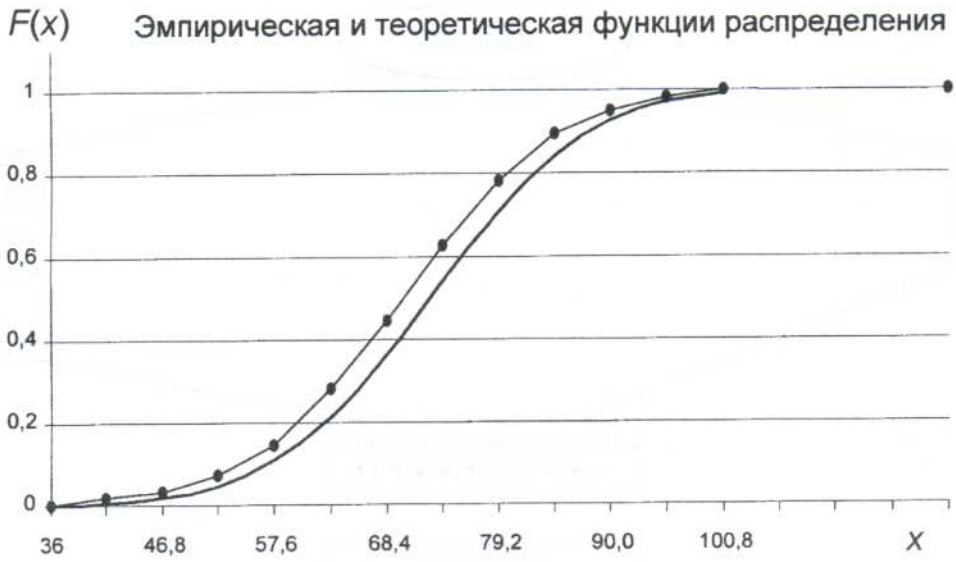


Рис. 2. Графики эмпирической (—●—) и теоретической (———) функций распределения коэффициента профессиональной компетентности

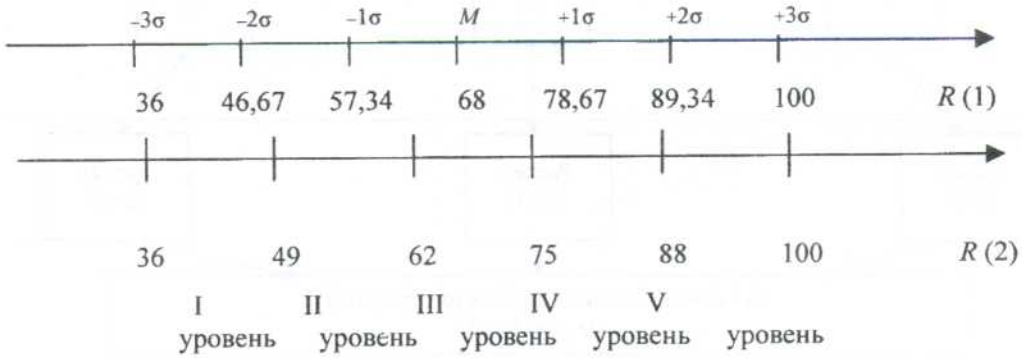


Рис. 3. Сравнение распределения $R(2)$ по уровням (I–V) профессиональной компетентности с рассчитанными значениями нормального распределения $R(1)$

Он составит примерно 15–17 %. Для автоматизации этого процесса разработана программа на языке Matlab 6.5.

Предложен аппарат оценивания профессиональной компетентности в системе подготовки специалистов, включающий алгоритм расчета коэффициента профессиональной компетентности, который позволяет диагностировать качество профессиональной подготовки и является эффективным инструментом (средством), используемым при оценке степени готовности специалиста к продуктивному решению профессиональных задач.

СПИСОК ЛИТЕРАТУРЫ

1. **Александров В.В., Горский Н.Л.** Алгоритмы и программы структурного метода обработки данных. Л.: Наука, 1983. 367 с.

2. **Вентцель Е.С.** Исследование операций. М., 1972.

3. Военная педагогика / Под ред. А.В. Барабанщикова; ВПА. М., 1986. 268 с.

4. **Татур Ю.Г.** Компетентность в структуре модели качества подготовки специалиста // Высшее образование сегодня. 2004. № 3. С. 20–26.

СВЕДЕНИЯ ОБ АВТОРАХ

КОНТАКТНЫЕ ДАННЫЕ

АКСЕНОВ Константин Александрович – *доцент кафедры автоматизированных систем управления Уральского государственного технического университета имени первого Президента России Б.Н. Ельцина, кандидат технических наук, доцент.*

620078, г. Екатеринбург, ул. Мира, д. 32. Тел. (343) 375-41-45.

АНИКИН Игорь Вячеславович – *заведующий кафедрой систем информационной безопасности Казанского государственного технического университета имени А.Н. Туполева (КАИ), кандидат технических наук, доцент.*

420111, Казань, ул. К. Маркса, д. 10. Тел. раб. (843) 231-00-69, e-mail anikinigor@evm.kstu-kai.ru

АРЗУМАНЯН Максим Юрьевич – *аспирант кафедры информационных технологий в экономике Санкт-Петербургского государственного университета телекоммуникаций имени проф. М.А. Бонч-Бруевича.*

191186, Санкт-Петербург, наб. р. Мойки, д. 61. Тел. (812) 315-25-38, e-mail arz83@mail.ru

БЕКУРИН Дмитрий Владимирович – *ведущий инженер ОАО "Авиадвигатель", кандидат физико-математических наук.*

614090, Пермь, Букирева, д. 15. Тел. (342)2396-450.

БЕЛЯКОВА Татьяна Павловна – *инженер-программист ООО "Дигитон", магистр Санкт-Петербургского государственного университета телекоммуникаций имени проф. М.А. Бонч-Бруевича.*

197110, Санкт-Петербург, Крестовский пр., д. 15, пом. 12-Н. Тел. (812) 601-68-12, e-mail tatyana@digiton.ru

БОЛДЫРЕВ Юрий Яковлевич – *заведующий кафедрой математического и программного обеспечения высокопроизводительных вычислений Санкт-Петербургского государственного политехнического университета, доктор технических наук, профессор.*

195251, Санкт-Петербург, ул. Политехническая, д. 29, корп. 4. Тел. (812) 596-28-61, e-mail boldyrev@phmf.spbstu.ru

БРАТИН Даниил Юрьевич – *аспирант кафедры информатики Санкт-Петербургского государственного университета.*

198504, Санкт-Петербург, Старый Петергоф, Университетский пр., д. 28. Тел. (812) 445-34-18.

БРЫКСИН Виталий Михайлович – *научный сотрудник лаборатории сетевых информационных технологий Югорского НИИ информационных систем.*

656049, г. Барнаул, пр. Ленина, д. 61. Тел. (3852) 36-70-18.

ВАН КАЙ – *аспирант кафедры автоматизированных систем управления Уральского государственного технического университета имени первого Президента России Б.Н. Ельцина.*

620078, г. Екатеринбург, ул. Мира, д. 32. Тел. (343) 375-41-45.

ВАРГАУЗИН Виктор Анатольевич – *доцент кафедры радиотехники и телекоммуникаций Санкт-Петербургского государственного политехнического университета, кандидат технических наук, доцент.*

195251, Санкт-Петербург, ул. Политехническая, д. 29, корп. 2. Тел. (812) 552-61-89, e-mail var@mail.spbstu.ru

ВЕТРИНСКИЙ Юрий Анатольевич – *доцент кафедры радиотехники и телекоммуникаций Санкт-Петербургского государственного политехнического университета, кандидат технических наук, доцент.*

195251, Санкт-Петербург, ул. Политехническая, д. 29. Тел. (812) 552-76-39, e-mail vetrinsky@mail.spbstu.ru

ГАВРИЛОВСКАЯ Надежда Владимировна – *аспирантка кафедры теоретической кибернетики и прикладной математики Алтайского государственного университета.*

656049, г. Барнаул, пр. Ленина, д. 61. Тел. (3852) 36-70-18, e-mail nadia@math.asu.ru

ГИЛЬМУЛЛИН Тимур Мансурович – *ассистент кафедры систем информационной безопасности Казанского государственного технического университета имени А.Н. Туполева (КАИ), аспирант.*

420111, Казань, ул. К. Маркса, д. 10. Тел. раб. (843) 231-00-69, e-mail LTim@Land.ru

ГОРБУНОВ Владимир Павлович – *аспирант кафедры радиофизики Санкт-Петербургского государственного политехнического университета.*

195251, Санкт-Петербург, ул. Политехническая, д. 29, корп. 2. Тел. (812) 552-97-86, e-mail Gorbunov_V@bk.ru

ГРУДИНИН Владимир Алексеевич – *генеральный директор ООО "Дигитон", аспирант Санкт-Петербургского государственного университета телекоммуникаций имени проф. М.А. Бонч-Бруевича.*

197110, Санкт-Петербург, Крестовский пр., д. 15, пом. 12-Н. Тел. (812) 601-68-12, e-mail vlad@digiton.ru

ДОРОСИНСКИЙ Леонид Григорьевич – *декан радиотехнического факультета Уральского государственного технического университета – УПИ имени первого президента России Б.Н. Ельцина, заведующий кафедрой автоматизированных систем управления, доктор технических наук, профессор.*

620002, г. Екатеринбург, ул. Мира, д. 32. Тел. (343) 375-41-45, e-mail asu@rtf.ustu.ru

ЗРАЕНКО Алексей Сергеевич – *аспирант кафедры автоматизированных систем управления Уральского государственного технического университета имени первого Президента России Б.Н. Ельцина.*

620078, г. Екатеринбург, ул. Мира, д. 32. Тел. (343) 375-41-45, e-mail zraenko@yandex.ru

ИВАНОВСКИЙ Сергей Алексеевич – *доцент кафедры математического обеспечения и применения ЭВМ Санкт-Петербургского государственного электротехнического университета "ЛЭТИ", кандидат технических наук, доцент.*

620078, г. Екатеринбург, ул. Мира, д. 32. Тел. (343) 375-41-45.

ИКРЯНОВ Игнат Иванович – *аспирант Московского физико-технического института (Государственного университета).*

141700, Московская обл., г. Долгопрудный, Институтский пер., д. 9. Тел. (495) 408-45-54, факс (495) 408-42-54.

КАЗАКОВ Борис Борисович – *аспирант кафедры математического обеспечения и применения ЭВМ Санкт-Петербургского государственного электротехнического университета "ЛЭТИ".*

620078, г. Екатеринбург, ул. Мира, д. 32. Тел. (343) 375-41-45, e-mail Boris.kazakov@gmail.com

КАРЕЛИН Игорь Владимирович – *студент кафедры компьютерных интеллектуальных систем в проектировании Санкт-Петербургского государственного политехнического университета, бакалавр.*

196135, Санкт-Петербург, ул. Алтайская, д. 20. Тел. (812) 373-42-33, e-mail ikarus47@mail.ru

КАРЕЛИН Владимир Александрович – *ведущий научный сотрудник Санкт-Петербургского российского института радионавигации и времени, кандидат технических наук.*

196135, Санкт-Петербург, ул. Алтайская, д. 20. Тел. (812) 373-42-33, e-mail renyxa46@mail.ru

КОНОВАЛОВ Владимир Николаевич – *старший преподаватель кафедры систем автоматического управления и электротехники Калужского филиала ГОУ ВПО "Московский государственный технический университет имени Н.Э. Баумана".*

248600, г. Калуга, ул. Баженова, д. 2. Тел. (4842) 745095, e-mail v.konovalov@list.ru

КОНОВАЛОВ Иван Владимирович – *инженер ЗАО научно-производственной фирмы "Сигма".*

248600, г. Калуга, ул. Баженова, д. 2. Тел. (4842) 745095, e-mail ivank-85@mail.ru

КРУПЕНКО Денис Александрович – *аспирант кафедры систем автоматизированного проектирования Санкт-Петербургского государственного электротехнического университета "ЛЭТИ" имени В.И. Ульянова (Ленина).*

197376, Санкт-Петербург, ул. Проф. Попова, д. 5. Тел. (343) 277-57-76, e-mail dakrupenko@rambler.ru

КУДРИН Максим Юрьевич – *аспирант кафедры информатики Московского физико-технического института.*

141700, Московская обл., г. Долгопрудный, Институтский пер., д. 9. Тел. (495) 408-66-95, e-mail mku@yandex.ru

КУНТУРОВА Надежда Борисовна – *профессор кафедры математики Военно-космической академии имени А.Ф. Можайского, кандидат педагогических наук, доцент.*

197082, Санкт-Петербург, П-82, ул. Ждановская, д. 13. E-mail giz26@rambler.ru

МАКАРОВА Елена Анатольевна – *доцент кафедры технической кибернетики Уфимского государственного авиационного технического университета, кандидат технических наук, доцент, докторант.*

450000, г. Уфа, ул. К. Маркса, д. 12. Тел. (347) 273-78-35, e-mail ea-makarova@mail.ru

МАРЦЕНЮК Михаил Андреевич – *заведующий кафедрой компьютерных систем и телекоммуникаций Пермского государственного университета, доктор физико-математических наук, профессор.*

614090, Пермь, ул. Букирева, д. 15. Тел. (342) 2396-450.

НАГОРНЫЙ Владимир Степанович – *профессор кафедры транспортных и технологических систем Санкт-Петербургского государственного политехнического университета, доктор технических наук, заслуженный деятель науки Российской Федерации.*

195251, Санкт-Петербург, ул. Политехническая, д. 29. Тел (812) 552-60-88, e-mail nagorny@jet.hop.stu.neva.ru

НЕНАШЕВ Александр Васильевич – *доцент кафедры радиоэлектроники Сургутского государственного университета Ханты-Мансийского автономного округа (ХМАО) – Югры, кандидат технических наук, доцент.*

628412, г. Сургут, пр. Ленина, д. 1. Тел. (346-2) 76-31-24, e-mail navas1@mail.ru

НОВИКОВ Юрий Николаевич – *доцент кафедры радиофизики Санкт-Петербургского государственного политехнического университета, кандидат физико-математических наук.*

195251, Санкт-Петербург, ул. Политехническая, д. 29, корп. 2. Тел. (812) 552-97-86, e-mail jn@mail.nnz.ru

ОРЛОВ Сергей Владимирович – *аспирант кафедры электронных вычислительных машин Тульского государственного университета.*

300600, г. Тула, пр. Ленина, д. 92. Тел. (4872) 33-24-45, e-mail s.orlov@list.ru

ОХОТНИКОВ Вячеслав Анатольевич – *аспирант кафедры радиоэлектроники Сургутского государственного университета Ханты-Мансийского автономного округа (ХМАО).*

628412, г. Сургут, пр. Ленина, д. 1. Тел. (346-2) 76-31-24, e-mail satellite@inbox.ru

ПЕТУХОВ Евгений Павлович – *аспирант кафедры математического и программного обеспечения высокопроизводительных вычислений Санкт-Петербургского государственного политехнического университета.*

195251, Санкт-Петербург, ул. Политехническая, д. 29, корп. 4. Тел. (812) 596-28-61, e-mail eugene@lamm.spbstu.ru

ПОБЕДАШ Павел Николаевич – *докторант кафедры системного анализа и исследования операций Сибирского государственного аэрокосмического университета.*

660014, г. Красноярск, пр. им. газеты "Красноярский рабочий", д. 31. Тел. (3912) 91-91-41, e-mail pobed_pnp@mail.ru

ПОПЦОВ Николай Александрович – *аспирант кафедры компьютерных систем и телекоммуникаций Пермского государственного университета.*

614090, Пермь, ул. Букирева, д. 15. Тел. (342)2396-450, e-mail Nikolai_work@mail.ru

СЕРЕДИНСКИЙ Юрий Владимирович – *студент кафедры радиофизики Санкт-Петербургского государственного политехнического университета.*

195251, Санкт-Петербург, ул. Политехническая, д. 29, корп. 2. Тел. (812) 552-97-86, e-mail seredinskyy@mail.ru

СИДОРОВ Юрий Евгеньевич – *профессор кафедры радиотехники и телекоммуникаций Санкт-Петербургского государственного политехнического университета, доктор технических наук, профессор.*

195251, Санкт-Петербург, ул. Политехническая, д. 29, 2-й учеб. корп., ауд. 440. Тел./факс (812) 552-95-16, e-mail jusidorov@mail.ru

СОКОЛОВСКИЙ Артем Константинович – *аспирант кафедры системы качества Санкт-Петербургского государственного политехнического университета.*

195251, Санкт-Петербург, Политехническая ул., д. 29. Тел. (812) 297-20-95, e-mail artem.sokolovsky@gmail.com

СОКОЛОВ Евгений Владимирович – *аспирант кафедры информатики Московского физико-технического института.*

1414700, Московская обл., пос. Долгопрудный, Институтский пер., д. 9. Тел. (495)-4086695, e-mail kason@yandex.ru

СОЛНУШКИН Константин Сергеевич – *ведущий программист главного информационно-вычислительного комплекса Санкт-Петербургского государственного политехнического университета.*

195251, Санкт-Петербург, ул. Политехническая, 29. Тел. (812) 596-28-61, e-mail ks@spbstu.ru

СОРОЦКИЙ Владимир Александрович – *доцент кафедры радиотехники и телекоммуникаций Санкт-Петербургского государственного политехнического университета, кандидат технических наук, доцент.*

195251, Санкт-Петербург, ул. Политехническая, д. 29. Тел. (812) 552-76-79, моб. тел. 8-921-3465518, e-mail sorotsky@mail.spbstu.ru

СТЕНИН Алексей Алексеевич – *аспирант кафедры компьютерных систем и технологий Московского инженерно-физического института (Национальный исследовательский ядерный центр), инженер НПП "Фактор-ТС".*

123290, Москва, 1-й Магистральный пр-д, д. 11, стр. 1, бизнес-центр "Интерьер". Тел. (495) 644-3130, e-mail staa4891@list.ru

СТУКОВ Михаил Сергеевич – *аспирант кафедры компьютерных систем и технологий Московского инженерно-физического института (Национальный исследовательский ядерный центр), инженер НПП "Фактор-ТС".*

123290, Москва, 1-й Магистральный пр-д, д. 11, стр. 1, бизнес-центр "Интерьер". Тел. (495) 644-3130, e-mail m.stukov@gmail.com

СУХОВ Антон Александрович – *аспирант кафедры персональных компьютеров и сетей Московского государственного университета приборостроения и информатики.*

113093, Москва, 1-й Щипковский пер., д. 23. E-mail suhov.anton@rambler.ru

ТАЗЕТДИНОВ Андрей Дамирович – *директор Санкт-Петербургского центра информационных технологий АНО ВПО "Международный банковский институт", кандидат технических наук.*

191023, Санкт-Петербург, Невский пр., д. 60. Тел. (812) 570-55-52, e-mail 191_nkp@bk.ru

ТЕЙХРИБ Антон Павлович – *аспирант кафедры автоматизированных систем управления Уральского государственного технического университета имени первого Президента России Б.Н. Ельцина.*

620002, г. Екатеринбург, ул. Мира, д. 32. Тел. (343) 375-41-45, e-mail a_p_t@mail.ru

ТОКАРЕВ Вячеслав Леонидович – *профессор кафедры электронных вычислительных машин Тульского государственного университета, доктор технических наук, профессор.*

300600, г. Тула, пр. Ленина, д. 92. Тел. (4872) 33-24-45.

ТОМИЛИН Александр Вячеславович – *аспирант кафедры компьютерных систем и технологий Московского инженерно-физического института (Национальный исследовательский ядерный центр), инженер НПП "Фактор-ТС".*

123290, Москва, 1-й Магистральный пр-д, д. 11, стр. 1, бизнес-центр "Интерьер". Тел. (495) 644-3130, e-mail plantera_@mail.ru

ТОРМАСОВ Александр Геннадьевич – *заместитель заведующего кафедрой информатики Московского физико-технического института, доктор физико-математических наук.*

1414700, Московская обл., г. Долгопрудный, Институтский пер., д. 9. Тел. (495)-4086695, e-mail tor@crec.mipt.ru

ТУТЫГИН Владимир Семенович – *доцент кафедры информационных и управляющих систем Санкт-Петербургского государственного политехнического университета, кандидат технических наук, доцент.*

195251, Санкт-Петербург, ул. Политехническая, д. 21. Тел. (812) 552-76-66.

УТКИН Марк Аркадьевич – *доцент кафедры радиотехники и телекоммуникаций Санкт-Петербургского государственного политехнического университета, кандидат технических наук, доцент.*

195251, Санкт-Петербург, ул. Политехническая, д. 29. Тел. (812) 552-76-79, моб. тел. 8-921-9499142, e-mail utkinmark@gmail.com

ФИРСОВ Антон Николаевич – *ассистент кафедры математического обеспечения вычислительных систем Пермского государственного университета.*

614990 Пермь, ул. Букирева, д. 15. Тел. (342) 239-66-27, e-mail a_firsov@mail.ru

ХВОРОВА Любовь Анатольевна – *доцент кафедры теоретической кибернетики и прикладной математики Алтайского государственного университета, кандидат технических наук, доцент.*

656049, г. Барнаул, пр. Ленина, д. 61. Тел. (3852) 36-70-18.

ЦИКИН Игорь Анатольевич – *заведующий кафедрой радиотехники и телекоммуникаций Санкт-Петербургского государственного политехнического университета, доктор технических наук, профессор.*

195251, Санкт-Петербург, ул. Политехническая, д. 29, корп. 2. Тел. (812) 552-61-89, e-mail tsikin@mail.spbstu.ru

ШЕДОВ Сергей Владимирович – *аспирант кафедры информационных и управляющих систем Санкт-Петербургского государственного политехнического университета.*

195251, Санкт-Петербург, Политехническая ул., д. 21. Тел. (812) 552-76-66.

ШУМИЛОВ Анатолий Владимирович – *студент кафедры радиотехники и телекоммуникаций Санкт-Петербургского государственного политехнического университета.*

195251, Санкт-Петербург, ул. Политехническая, д. 29, 2-й учеб. корп., ауд. 440. Тел./факс (812) 552-95-16.

АННОТАЦИИ

КЛЮЧЕВЫЕ СЛОВА

Тазетдинов А. Д. ОБ ИСПОЛЬЗОВАНИИ ГРАФОВ ПОНЯТИЙ ДЛЯ СТРУКТУРИРОВАНИЯ И УПРАВЛЕНИЯ ИНФОРМАЦИЕЙ В АВТОМАТИЗИРОВАННЫХ ОБУЧАЮЩИХ СИСТЕМАХ.

Вместо традиционного способа структурирования обучающей информации на разделы, темы, параграфы предложено использовать графы понятий. В этом случае для обучающей информации введены дополнительные характеристики и предложен способ их измерения, что позволяет более точно определять параметры обучающей информации и тем самым более эффективно управлять процессом обучения.

УПРАВЛЕНИЕ. СТРУКТУРИРОВАНИЕ ИНФОРМАЦИИ. АВТОМАТИЗИРОВАННЫЕ ОБУЧАЮЩИЕ СИСТЕМЫ. МОДЕЛИРОВАНИЕ. ОПТИМИЗАЦИЯ.

Соколовский А. К. ОЦЕНКА ПРИВЛЕКАТЕЛЬНОСТИ ПРОЕКТОВ В СФЕРЕ ИНФОРМАЦИОННЫХ ТЕХНОЛОГИЙ.

Рассмотрены проблемы, связанные с оценкой привлекательности проектов в сфере информационных технологий (ИТ). Предложена методика по оценке привлекательности ИТ проектов, основанная на требованиях заинтересованных сторон и направленная на повышение достоверности оценки привлекательности и ускорение процесса принятия управленческого решения при помощи ранжирования результатов оценивания.

ИТ ПРОЕКТ. ПРИВЛЕКАТЕЛЬНОСТЬ. ОЦЕНКА. МОДЕЛИРОВАНИЕ. ОПТИМИЗАЦИЯ.

Сухов А. А. ПРОБЛЕМА ЧЕЛОВЕКО-МАШИННОГО ИНТЕРФЕЙСА В ИНФОРМАЦИОННЫХ СИСТЕМАХ.

Рассмотрен метод создания человеко-машинного интерфейса при помощи нейронных сетей. Метод позволит использовать плюсы экспертных оценок, одновременно автоматизировав и стандартизировав процесс оценки.

ИНТЕРФЕЙС. ОЦЕНКА ИНТЕРФЕЙСА. НЕЙРОННЫЕ СЕТИ.

Хворова Л. А., Брыксин В. М., Гавриловская Н. В. ИСПОЛЬЗОВАНИЕ ИНФОРМАЦИОННЫХ ТЕХНОЛОГИЙ ПРИ ПРОГНОЗИРОВАНИИ УРОЖАЙНОСТИ ЗЕРНОВЫХ КУЛЬТУР.

Рассмотрены вопросы использования информационных технологий при оценке урожайности зерновых культур на примере яровой пшеницы на территории Западной Сибири. Приведены основные результаты научно-исследовательской работы по прогнозу урожайности сельскохозяйственных культур в условиях Тюменской, Новосибирской областей и Алтайского края, а именно поэтапная оценка урожайности с использованием разработанных авторами технологии определения года-аналога и эмпирической модели, технологии имитационного комплекса MIDC и данных дистанционного зондирования Земли.

ГОД-АНАЛОГ. УРОЖАЙНОСТЬ. ПРОДУКЦИОННЫЙ ПРОЦЕСС. ВЕГЕТАЦИОННЫЙ ПЕРИОД. МОДЕЛЬ. ПРОГНОЗ. ДИСТАНЦИОННОЕ ЗОНДИРОВАНИЕ.

Болдырев Ю. Я., Варгаузин В. А., Петухов Е. П., Цикин И. А. ТЕХНОЛОГИЯ СЕТЕВОГО ДОСТУПА К ПРОГРАММНОЙ СРЕДЕ MATLAB НА ОСНОВЕ КЛАСТЕРНОЙ ВЫЧИСЛИТЕЛЬНОЙ СИСТЕМЫ.

Рассмотрены принципы сетевого доступа к программному комплексу MATLAB. Изложены особенности работы комплекса на вычислительной кластерной системе.

УДАЛЕННЫЙ ДОСТУП. ПРОГРАММНЫЙ КОМПЛЕКС MATLAB. МНОГОПОЛЬЗОВАТЕЛЬСКАЯ РАБОТА. ВЫЧИСЛИТЕЛЬНЫЙ КЛАСТЕР.

Фирсов А. Н. ОЦЕНКА НАДЕЖНОСТИ АСИНХРОННЫХ РАСПРЕДЕЛЕННЫХ СИСТЕМ, УСТОЙЧИВЫХ К ПРОИЗВОЛЬНЫМ ОТКАЗАМ.

Предложены формула и алгоритм для оценки надежности устойчивых к произвольным отказам систем, работающих в гетерогенной по надежности среде. Демонстрируются ситуации, когда существующие решения неэффективны, и предлагаются способы их усовершенствования.

ОТКАЗОУСТОЙЧИВОСТЬ. ПРОИЗВОЛЬНЫЕ (ВИЗАНТИЙСКИЕ) СБОИ. АСИНХРОННЫЕ РАСПРЕДЕЛЕННЫЕ СИСТЕМЫ.

Сидоров Ю. Е., Шумилов А. В. ИССЛЕДОВАНИЕ ПОМЕХОУСТОЙЧИВОСТИ РАНГОВОГО ОБНАРУЖИТЕЛЯ СИГНАЛОВ ПРИ АПРИОРНОЙ НЕОПРЕДЕЛЕННОСТИ.

Рассмотрен непараметрический ранговый обнаружитель радиосигналов в шумах с неизвестным распределением. Его помехоустойчивость оценена с помощью метода статистического имитационного моделирования (метод Монте-Карло).

ОБНАРУЖЕНИЕ СИГНАЛОВ. АПРИОРНАЯ НЕОПРЕДЕЛЕННОСТЬ. СТАТИСТИЧЕСКОЕ ИМИТАЦИОННОЕ МОДЕЛИРОВАНИЕ. РАНГОВЫЕ ПРАВИЛА.

Грудинин В. А., Белякова Т. П. СИСТЕМА ОПЕРАТИВНОГО КОНТРОЛЯ СОСТОЯНИЯ ПОКРЫШЕК ГРУЗОВЫХ АВТОМОБИЛЕЙ.

Приведены основные результаты исследований, проведенных в рамках разработки системы оперативного оповещения водителя о возникновении дефекта (отслоение части корда) колеса грузового автомобиля во время его движения.

АКУСТИЧЕСКИЕ СИГНАЛЫ. СПЕКТРАЛЬНЫЙ АНАЛИЗ. АМПЛИТУДНАЯ МОДУЛЯЦИЯ. ПИК-ФАКТОР. АНАЛИЗ ОГИБАЮЩЕЙ.

Казаков Б. Б., Ивановский С. А. ЭФФЕКТИВНЫЕ МЕТОДЫ ВЫДЕЛЕНИЯ ДВИЖЕНИЯ ОБЪЕКТОВ В ПОСЛЕДОВАТЕЛЬНОСТИ ИЗОБРАЖЕНИЙ.

Разработан подход к выделению движения в последовательности изображений на основе оптического потока. Применение регуляризированного локального метода вычисления оптического потока позволяет отслеживать движение объектов в реальном времени на широком спектре оборудования (в том числе во встраиваемых устройствах). Для проведения вычислительных экспериментов разработан ПК визуального исследования алгоритмов.

СЛЕЖЕНИЕ. ОПТИЧЕСКИЙ ПОТОК. СЕГМЕНТАЦИЯ. CUDA. GPU. ВИЗУАЛЬНОЕ ПРОГРАММИРОВАНИЕ.

Томилин А. В., Стуков М. С., Стенин А. А. ПРИМЕНЕНИЕ ДИВЕРСИФИКАЦИИ В ЗАДАЧАХ ДИНАМИЧЕСКОЙ ЗАЩИТЫ ПАМЯТИ.

Рассмотрены вопросы практического применения диверсификации для задач динамической защиты памяти C/C++ программ от атак, использующих уязвимости типа "переполнение буфера". Дан анализ методов противодействия попыткам НСД, использующим переполнение буфера, и приведена классификация атак. Область применения – задачи защиты информации компьютерных систем ответственного назначения.

ПЕРЕПОЛНЕНИЕ БУФЕРА. ДИВЕРСИФИКАЦИЯ. ЗАЩИТА ПАМЯТИ.

Тутыгин В. С., Шедов С. В. НОВЫЕ АЛГОРИТМЫ ОБНАРУЖЕНИЯ И ОПРЕДЕЛЕНИЯ ПАРАМЕТРОВ ЗАШУМЛЕННЫХ СИГНАЛОВ.

Рассмотрены новые подходы к решению задачи обнаружения и определения параметров зашумленных эхо-сигналов спектрометров ЯМР и ЛЧМ-сигналов, основанные на дополнительной обработке спектров, полученных с помощью БПФ и использующие нетрадиционное сочетание нормированной кросс-корреляции, сплайн-интерполяции, передискретизации и аппроксимации специальной функцией.

ЯДЕРНЫЙ МАГНИТНЫЙ РЕЗОНАНС. ЭХО-СИГНАЛЫ. ЛЧМ-СИГНАЛЫ. БЫСТРОЕ ПРЕОБРАЗОВАНИЕ ФУРЬЕ. ДОПОЛНИТЕЛЬНАЯ ОБРАБОТКА. НОРМИРОВАННАЯ КРОСС-КОРРЕЛЯЦИЯ. СПЛАЙН-ИНТЕРПОЛЯЦИЯ. ПЕРЕДИСКРЕТИЗАЦИЯ. АППРОКСИМАЦИЯ СПЕЦИАЛЬНОЙ ФУНКЦИЕЙ.

Ненашев А. В., Охотников В. А. РАЗРАБОТКА МЕТОДА АНАЛИЗА И МОДЕЛИРОВАНИЯ НЕЛИНЕЙНЫХ РАДИОТЕХНИЧЕСКИХ УСТРОЙСТВ.

Предложен метод анализа и моделирования нелинейных радиотехнических устройств, при котором нелинейные элементы заменяются источниками тока или ЭДС согласно теореме компенсации. Решение находится методом последовательных приближений. Рассмотрение ведётся на примере усилительного каскада на ПТШ.

НЕЛИНЕЙНЫЕ УСТРОЙСТВА. МОДЕЛИРОВАНИЕ. ТЕОРЕМА КОМПЕНСАЦИИ. ПОСЛЕДОВАТЕЛЬНЫЕ ПРИБЛИЖЕНИЯ. УСИЛИТЕЛЬ.

Карелин И. В., Карелин В. А. ПОЛУЧЕНИЕ МАКСИМАЛЬНО ПРАВДОПОДОБНЫХ ОЦЕНОК МАЛЫХ ОТКЛОНЕНИЙ ЧАСТОТ ВЫСОКОСТАБИЛЬНЫХ ГАРМОНИЧЕСКИХ СИГНАЛОВ.

Предложен алгоритм синтеза весовой функции фильтра для оценки среднего отклонения частоты сигналов прецизионных генераторов, который в отличие от методов классической теории нелинейной и линейной фильтрации не требует упрощающих ограничений относительно модели спектральной плотности мощности помехи. Показано, что точность получаемых оценок совпадает с потенциальной точностью оптимальных нелинейных фильтров метода максимального правдоподобия (ММП) при значительно меньших аппаратных и программных затратах для практической реализации синтезируемых фильтров.

ВЕСОВАЯ ФУНКЦИЯ. УСРЕДНЯЮЩИЙ ФИЛЬТР. ПРЕЦЕЗИОННЫЙ ГЕНЕРАТОР. РАДИОТЕХНИЧЕСКАЯ СИСТЕМА. МЕТОД МАКСИМАЛЬНОГО ПРАВДОПОДОБИЯ.

Коновалов В. Н., Коновалов И. В. ПОДХОД К МНОГОУРОВНЕВОЙ ОПТИМИЗАЦИИ КОМБИНАЦИОННЫХ ЛОГИЧЕСКИХ СХЕМ.

Рассмотрен алгоритм минимизации логических схем, полученных в результате многоуровневой декомпозиции. Алгоритм основан на проведении анализа схемы, введении логической недоопределённости во входные числовые последовательности элементов схемы и поиске элементов с непротиворечивыми последовательностями, которые могут быть в дальнейшем соединены между собой, что приводит к исключению из схемы соответствующих фрагментов.

КОМБИНАЦИОННАЯ ЛОГИЧЕСКАЯ СХЕМА. МИНИМИЗАЦИЯ БУЛЕВЫХ ФУНКЦИЙ. ЛОГИЧЕСКИЙ СИНТЕЗ. ДЕКОМПОЗИЦИЯ. ОПТИМИЗАЦИЯ.

Серединский Ю. В., Горбунов В. П., Новиков Ю. Н. ПРИМЕНЕНИЕ ФЕРРИТОВ ДЛЯ ОСЛАБЛЕНИЯ ВЛИЯНИЯ КАБЕЛЕЙ НА ХАРАКТЕРИСТИКИ ИЗЛУЧЕНИЯ АНТЕННЫХ СИСТЕМ.

Исследована проблема паразитного влияния кабелей на характеристики антенных систем многоярусных антенных комплексов. Для ее решения предложено использовать ферритовые материалы (кольца, стержни). Показано, что правильное применение ферритов позволяет ослабить влияние кабелей, проходящих вблизи антенн.

АНТЕННАЯ СИСТЕМА. ЭЛЕКТРОМАГНИТНАЯ СОВМЕСТИМОСТЬ. ИЗЛУЧЕНИЕ КАБЕЛЕЙ. ФЕРРИТОВЫЕ СТЕРЖНИ. ФЕРРИТОВЫЕ КОЛЬЦА.

Сороцкий В. А., Уткин М. А. ВЛИЯНИЕ ЭЛЕКТРОМАГНИТНОЙ СВЯЗИ АНТЕНН РАДИОПЕРЕДАЮЩИХ УСТРОЙСТВ НА ПРОЦЕССЫ В КЛЮЧЕВЫХ ГЕНЕРАТОРАХ.

Рассмотрены особенности функционирования двух РПдУ, обусловленные их взаимовлиянием. Показано, что электромагнитная связь антенн РПдУ может привести к ухудшению характеристик и возникновению аварийных режимов работы ключевых генераторов.

РАДИОПЕРЕДАЮЩЕЕ УСТРОЙСТВО. КЛЮЧЕВОЙ ГЕНЕРАТОР. АНТЕННА. ВЗАИМНОЕ ВЛИЯНИЕ.

Сороцкий В. А., Уткин М. А. ИССЛЕДОВАНИЕ ХАРАКТЕРИСТИК МНОГОЯЧЕЙКОВЫХ ТИРИСТОРНЫХ ГЕНЕРАТОРОВ НА ЦИФРОВОЙ МОДЕЛИ.

Рассмотрены вопросы разработки цифровой модели многоячейковых тиристорных генераторов. Приведены результаты моделирования энергетических характеристик и проведен их анализ.

ТИРИСТОР. МАКРОМОДЕЛЬ. ГЕНЕРАТОР. ЦИФРОВАЯ МОДЕЛЬ.

Макарова Е. А. УПРАВЛЕНИЕ ПОВЕДЕНИЕМ МАКРОЭКОНОМИЧЕСКИХ АГЕНТОВ НА ОСНОВЕ МЕХАНИЗМА САМОВЫРАВНИВАНИЯ ФИНАНСОВЫХ ПОТОКОВ.

Предложена динамическая модель поведения макроэкономического агента. Представлен алгоритм корректировки планового темпа расхода финансовых ресурсов агента на основе информации об объеме накопленных запасов. Сформированы три варианта интерпретации функциональной схемы динамической модели с позиций экономической теории, теории управления и теории систем; проведен их анализ.

МАКРОЭКОНОМИЧЕСКИЙ АГЕНТ. ДИНАМИЧЕСКАЯ МОДЕЛЬ. НЕРАВНОВЕСНЫЙ РЕЖИМ. ПОТОК. ЗАПАС.

Токарев В. Л., Орлов С. В. ЗАДАЧА АВТОМАТИЗАЦИИ ПРОЦЕССА ФОРМИРОВАНИЯ ЛИНГВИСТИЧЕСКИХ МОДЕЛЕЙ ДЛЯ ПОДДЕРЖКИ ПРИНЯТИЯ РЕШЕНИЙ.

Рассмотрен метод формирования лингвистических моделей системы для поддержки принятия решений. Предложена методика автоматического определения наилучших параметров модели.

ЛИНГВИСТИЧЕСКАЯ МОДЕЛЬ. ПРИБЛИЖЕННОЕ МОДЕЛИРОВАНИЕ. ИНТЕГРИРОВАННАЯ СИСТЕМА ПОДДЕРЖКИ ПРИНЯТИЯ РЕШЕНИЙ.

Победаш П. Н. АНАЛИЗ ДВУХКРИТЕРИАЛЬНОЙ МОДЕЛИ ОПТИМАЛЬНОГО УПРАВЛЕНИЯ РЕАЛЬНЫМИ ИНВЕСТИЦИЯМИ С НЕОПРЕДЕЛЕННЫМ СПРОСОМ.

Приведены оценки оптимального значения свертки в модели оптимального управления реальными инвестициями с неопределенным спросом с двумя экономическими агентами – производителем и налоговым центром, полученные на основе дискретного принципа максимума. Это позволяет оценивать эффективность соответствующего инвестиционного проекта на этапе его предварительного анализа.

ИНВЕСТИЦИОННЫЙ ПРОЕКТ. МНОГОШАГОВАЯ ЗАДАЧА ЛИНЕЙНОГО ПРОГРАММИРОВАНИЯ.

Попцов Н. А., Бекурин Д. Б., Марценюк М. А. НЕЙРОСЕТЕВОЕ УПРАВЛЕНИЕ ЛОКОМОЦИЯМИ САМООБУЧАЮЩЕЙСЯ МОДЕЛИ БИООБЪЕКТА.

Рассмотрена задача самообучения биообъектов локомоциям на примере модели амебы, управляемой самообучающейся нейросетью. Модель амебы состоит из двух эллипсоидов, связанных между собой тонким каналом, и может в определенных пределах изменять длину канала, а также объем и форму эллипсоидов. Гидродинамика движения объекта рассматривается в приближении Стокса. Объект стремится достичь некоторой цели – точки в пространстве, находящейся на заданном расстоянии от объекта. Рассмотрены два варианта управления локомоциями с самообучением и проведено сравнение их эффективности в зависимости от удаленности цели и её первоначальной ориентации относительно объекта.

САМООБУЧЕНИЕ НЕЙРОСЕТИ. МОДЕЛЬ ЛОКОМОЦИЙ. МОДЕЛЬ ПЛАВАНИЯ АМЕБЫ.

Кудрин М. Ю., Соколов Е. В., Тормасов А. Г. ВЫЯВЛЕНИЕ СОСТОЯНИЙ ГОНКИ С ПОМОЩЬЮ ГРАФА СОВМЕСТНОГО ИСПОЛНЕНИЯ ПОТОКОВ.

Предложен способ проверки наличия состояний гонки в потоках, использующих разделяемую память. Он применим, когда некорректную работу программы можно определить по итоговому состоянию исполнения потоков. В таких случаях он позволяет гарантированно обнаруживать все состояния гонки, не требуя при этом полного перебора вариантов совместного исполнения потоков программы.

МУЛЬТИПРОЦЕССОРНОЕ ПРОГРАММИРОВАНИЕ. СОСТОЯНИЕ ГОНКИ. СИНХРОНИЗАЦИЯ ПОТОКОВ.

Крупенко Д. А. АДАПТАЦИЯ СИСТЕМЫ ПРОЕКТИРОВАНИЯ AUTOCAD ДЛЯ РАЗРАБОТКИ КОНСТРУКЦИЙ МИКРОСБОРОК СВЧ.

Рассмотрены методы адаптации системы автоматизированного проектирования AutoCAD для обеспечения возможности ее работы по проектированию СВЧ-устройств с использованием динамических блоков.

КОНСТРУКТОРСКАЯ ДОКУМЕНТАЦИЯ. AUTOCAD. VISUAL LISP. ДИНАМИЧЕСКИЙ БЛОК. МИКРОСБОРКА. МОДУЛЬ СВЧ.

Солнушкин К. С. СИСТЕМА УЧЕТА ДОСТУПА К ЛИЦЕНЗИОННОМУ ПРОГРАММНОМУ ОБЕСПЕЧЕНИЮ В ЦЕНТРАХ КОЛЛЕКТИВНОГО ПОЛЬЗОВАНИЯ.

Представлена открытая модульная архитектура программной системы учета использования лицензионного ПО. На основе накопленных учетных данных имеется возможность проводить ретроспективный анализ и выявлять статистические закономерности. Система также имеет интерфейс с планировщиками задач, уменьшая вероятность неуспешного запуска задач из-за отсутствия необходимых лицензий.

ЛИЦЕНЗИОННОЕ ПРОГРАММНОЕ ОБЕСПЕЧЕНИЕ. МОНИТОРИНГ. УЧЁТ. ПЛАНИРОВЩИК ЗАДАЧ.

Сухов А. А. ОЦЕНКА ЭМОЦИОНАЛЬНОЙ СОСТАВЛЯЮЩЕЙ ИНТЕРФЕЙСОВ ТРЕНАЖЕРОВ.

Рассмотрен метод оценки человеко-машинного интерфейса при помощи нескольких нейронных сетей, который позволяет более объективно оценить качество интерфейса.

ИНТЕРФЕЙС. ОЦЕНКА ИНТЕРФЕЙСА. НЕЙРОННЫЕ СЕТИ. КОМПЬЮТЕРНЫЕ ТРЕНАЖЕРЫ.

Братин Д. Ю. ОСОБЕННОСТИ ЯЗЫКА FDL V. 1.0 ДЛЯ ОПИСАНИЯ ФРОНТЕНДА ВЕБ-ПРИЛОЖЕНИЙ.

Описан микроязык на базе XML, созданный для декларативного описания фронтенда веб-ресурса. Язык даёт возможность описывать структуру ресурса, функциональные возможности программной части и шаблоны страниц. Показано, что его отличием является удобная модель абстракции от реализации приложения и наглядные средства обращения к её компонентам.

ДЕКЛАРАТИВНЫЙ ЯЗЫК. ЯЗЫК ОПИСАНИЯ ПОЛЬЗОВАТЕЛЬСКИХ ИНТЕРФЕЙСОВ. РАЗРАБОТКА ВЕБ-ПРИЛОЖЕНИЙ.

Гильмуллин Т. М., Аникин И. В. МОДЕЛИРОВАНИЕ ОБЪЕКТОВ ИНФОРМАЦИОННОЙ БЕЗОПАСНОСТИ ДЛЯ ЗАДАЧИ ОЦЕНКИ РИСКОВ.

Предложены формальные модели объектов информационной безопасности для задачи оценки рисков ИБ. Разработана концептуальная модель предметной области “Риск информационной безопасности”. В рамках разработанных моделей сформулированы и доказаны фундаментальные теоремы для задачи управления рисками ИБ.

МОДЕЛИРОВАНИЕ. ИНФОРМАЦИОННАЯ БЕЗОПАСНОСТЬ. ОЦЕНКА РИСКОВ. НЕЧЕТКИЕ МНОЖЕСТВА.

Зраенко А. С., Аксенов К. А., Ван Кай. КОАЛИЦИОННАЯ МОДЕЛЬ МУЛЬТИАГЕНТНОГО ПРОЦЕССА ПРЕОБРАЗОВАНИЯ РЕСУРСОВ.

Рассмотрена задача поддержки принятия эффективных управленческих решений средствами мультиагентного моделирования. Для ее решения предложено использовать разработанную коалиционную модель на основе механизмов процессов преобразования ресурсов.

ПОДДЕРЖКА ПРИНЯТИЯ РЕШЕНИЙ. МУЛЬТИАГЕНТНЫЕ СИСТЕМЫ. МОДЕЛИРОВАНИЕ. ОПТИМИЗАЦИЯ. ПЛАНИРОВАНИЕ. ПРОЦЕССЫ ПРЕОБРАЗОВАНИЯ РЕСУРСОВ.

Арзуманян М. Ю. МОДЕЛИ ОПИСАНИЯ ПРЕДПРИЯТИЙ И ИХ МОНИТОРИРУЕМЫЕ СВОЙСТВА.

Рассмотрены основные существующие модели описания предприятий. Предложена новая модель, основанная на сервисном представлении, разобраны некоторые параметры сервисной модели. Предпринята попытка определения мониторируемых свойств для каждой модели.

ИНФОРМАЦИОННАЯ ЭКОНОМИКА. МОНИТОРИНГ. ПРЕДПРИЯТИЕ. МОДЕЛИ ПРЕДПРИЯТИЯ. СЕРВИСНОЕ ПРЕДСТАВЛЕНИЕ.

Тейхриб А. П., Доросинский Л. Г. СРАВНИТЕЛЬНЫЙ ОБЗОР МОДЕЛЕЙ ЛОГИЧЕСКОГО ВРЕМЕНИ.

Рассмотрено несколько моделей логического времени. Выполнено приведение моделей к единому формализму, а также определены свойства каждой модели.

ЛОГИЧЕСКОЕ ВРЕМЯ. РАСПРЕДЕЛЕННЫЕ СИСТЕМЫ. ПРАВДОПОДОБНОСТЬ. ИЗОМОРФНОСТЬ. МОДЕЛИРОВАНИЕ.

Икрянов И. И. МОДЕЛИ ДВИЖЕНИЯ ИНДИКАТОРОВ КУРСА И ГЛИССАДЫ С МАЯТНИКОВОЙ СИСТЕМОЙ СТАБИЛИЗАЦИИ ОПТИЧЕСКИХ ОСЕЙ.

Объектами исследования выбраны индикаторы курса и глиссады (ИК/ИГ), предназначенные для визуальной посадки вертолетов на палубы морских судов днем и ночью, в простых и сложных метеоусловиях. Предложены две модели ИК/ИГ для случаев качки корабля на регулярном и нерегулярном волнении. Приведено сравнение полученных с помощью различных моделей результатов, которое показывает их хорошее совпадение.

МОДЕЛЬ ДВИЖЕНИЯ. ОПТИЧЕСКИЕ ИНДИКАТОРЫ. ПОГРЕШНОСТЬ СТАБИЛИЗАЦИИ. СЛУЧАЙНЫЙ ПРОЦЕСС. АППРОКСИМАЦИЯ.

Нагорный В. С. БАЗОВЫЕ РЕЦЕПТУРЫ НАНОСТРУКТУРИРОВАННЫХ СЕРЕБРОМ РАБОЧИХ ЖИДКОСТЕЙ ДЛЯ ЭЛЕКТРОКАПЛЕСТРУЙНЫХ ТЕХНОЛОГИЙ.

Рассмотрен электрокаплеструйный метод высокого давления нанесения проводящих топологий на диэлектрические подложки. На основании проведенного патентного поиска и библиографического анализа предложен способ получения базовых рецептур рабочих жидкостей с наночастицами серебра.

ЭЛЕКТРОКАПЛЕСТРУЙНЫЙ МЕТОД. ПРОВОДЯЩИЕ ТОПОЛОГИИ. РАБОЧАЯ ЖИДКОСТЬ. НАНОЧАСТИЦЫ СЕРЕБРА.

Ветринский Ю. А. ОПЫТ ИСПОЛЬЗОВАНИЯ ТЕХНОЛОГИИ СМЕШАННОГО ОБУЧЕНИЯ В ПРАКТИКЕ ПРЕПОДАВАНИЯ ТЕХНИЧЕСКИХ ДИСЦИПЛИН.

Рассмотрен один из подходов к организации смешанного обучения в техническом университете. Представлены результаты экспериментальных исследований, подтверждающие эффективность предложенного подхода.

ДИСТАНЦИОННЫЙ ДОСТУП. СИСТЕМА УПРАВЛЕНИЯ ОБУЧЕНИЕМ. СМЕШАННОЕ ОБУЧЕНИЕ. УЧЕБНЫЙ КУРС.

Кунтурова Н. Б. МЕТОДИКА ФОРМИРОВАНИЯ СТРУКТУРНЫХ ГРУПП КАНДИДАТОВ ПРИ ПОСТУПЛЕНИИ В МАГИСТРАТУРУ С ПРОГНОЗИРУЕМОЙ УСПЕШНОСТЬЮ ИХ ОБУЧЕНИЯ.

Предложена методика формирования групп кандидатов при поступлении в магистратуру в вузах. Использование данной методики обеспечивает структурирование группы кандидатов и прогнозирует успешность их обучения и воспитания в магистратуре вузов.

ПРОФЕССИОНАЛЬНАЯ КОМПЕТЕНТНОСТЬ. КВАЛИМЕТРИЯ. ОБРАЗОВАТЕЛЬНАЯ ТРАЕКТОРИЯ. КАЧЕСТВО ПРОФЕССИОНАЛЬНОЙ ПОДГОТОВКИ. ПРОГНОЗ. УСПЕШНОСТЬ ПРОФЕССИОНАЛЬНОЙ ДЕЯТЕЛЬНОСТИ.

Кунтурова Н. Б. ОСОБЕННОСТИ ОЦЕНИВАНИЯ ПРОФЕССИОНАЛЬНОЙ КОМПЕТЕНТНОСТИ СПЕЦИАЛИСТОВ В ПРОЦЕССЕ ИХ ОБУЧЕНИЯ В ВУЗЕ.

Предложен оригинальный аппарат оценивания профессиональной компетентности в системе подготовки специалистов. Аппарат включает алгоритм расчета коэффициента профессиональной компетентности, который позволяет диагностировать качество профессиональной подготовки и является эффективным инструментом (средством), используемым при оценке степени готовности специалиста к продуктивному решению профессиональных задач. Использование алгоритма в практике управленческой деятельности вуза способствует выявлению перспектив развития обучаемых и позволяет решить задачу повышения эффективности использования выпускников.

ПРОФЕССИОНАЛЬНАЯ КОМПЕТЕНТНОСТЬ. КВАЛИМЕТРИЯ. ОБРАЗОВАТЕЛЬНАЯ ТРАЕКТОРИЯ. КАЧЕСТВО ПРОФЕССИОНАЛЬНОЙ ПОДГОТОВКИ. ПРОГНОЗ. УСПЕШНОСТЬ ПРОФЕССИОНАЛЬНОЙ ДЕЯТЕЛЬНОСТИ.

ABSTRACTS

KEYWORDS

Tazetdinov A. D. ABOUT USING GRAPHS OF CONCEPTS FOR STRUCTURIZATION AND MANAGEMENT INFORMATION IN THE AUTOMATED TRAINING SYSTEMS.

It is offered to use graphs of concepts instead of a traditional way of structuring training information on sections, themes and paragraphs. In this case for the training information additional characteristics are entered, and offered the way of their measurement that allows determining more precisely parameters of the training information and by that, more effectively to operate process of training.

MANAGEMENT. STRUCTURING INFORMATION. AUTOMATED TRAINING SYSTEMS. MODELLING. OPTIMIZATION.

Sokolovsky A. K. EVALUATION OF PROJECT ATTRACTIVENESS IN INFORMATION TECHNOLOGY FIELD.

The issues that deal with the evaluation of the information technology (IT) projects attractiveness are considered. The article proposes the methodology of the evaluation of IT projects attractiveness, based on requirements of stakeholders. The methodology is intended to increase the trustworthiness of attractiveness evaluation and accelerate the managerial decision making process with the help of evaluation results ranking.

IT PROJECT. ATTRACTIVENESS. EVALUATION. MODELLING. OPTIMIZATION.

Suhov A. A. PROBLEMS OF HUMAN-MACHINE INTERFACES IN INFORMATION SYSTEMS.

The method of creating human-machine interface by means of neural networks is described. This method allows the use of pluses peer reviews, both automated and standardized process of evaluation.

INTERFACE. EVALUATION OF THE INTERFACE. NEURAL NETWORKS.

Khvorova L. A., Bryksin V. M., Gavrilovskaya N. V. USE OF INFORMATION TECHNOLOGIES IN FORECASTING GRAIN YIELDS.

This article is devoted to information technology in assessing the yield of grain crops on the example of spring wheat in Western Siberia. This article presents the main results of research on forecasting crop production in the Tyumen, Novosibirsk and Altai regions, namely: a phased evaluation of yield, using technology developed by the authors definition of analog and empirical models, the complex technology of simulation and data MIDC Remote Sensing of the Earth.

YEAR-ANALOG. YIELD. PRODUCTION PROCESSES. THE GROWING SEASON. MODEL. THE FORECAST. REMOTE SENSING.

Boldyrev Yu. Ya., Vargauzin V. A., Petukhov E. P., Tsikin I. A. TECHNOLOGY FOR REMOTE NETWORK ACCESS TO MATLAB SOFTWARE ENVIRONMENT HOSTED ON COMPUTER CLUSTER SYSTEM.

Principles of remote network access to MATLAB software environment are considered. Specificity of particular software operation on computer cluster system is stated.

REMOTE ACCESS. SOFTWARE ENVIRONMENT MATLAB. MULTIUSER WORK. COMPUTER CLUSTER.

Firsov A. N. RELIABLITY ESTIMATION OF ASYNCRONOUS BYSINATINE FAULT-TOLERANT DISTRIBUTED SYSTEMS.

A formula and an effective polynomial algorithm for estimation of reliability of asynchronous byzantine fault-tolerant distributed system consisting of components of different reliability are presented. Situations are demonstrated where existing solutions are ineffective, and approaches to their improvement are suggested.

FAULT-TOLERANCE. BYSANTINE FAULTS. ASYNCRONOUS DISTRIBUTED SYSTEMS.

Sidorov Yu. E., Shumilov A. V. NONPARAMETRIC RADIO SIGNAL DETECTOR.

The article describes a nonparametric rank detector of radio signal in the presence of noise of unknown distribution. The noise resistance of the detector is estimated using statistical imitation modeling (Monte Carlo simulation).

SIGNAL DETECTION. A PRIORI UNCERTAINTY. STATISTICAL IMITATION MODELING. RANK RULES.

Grudinin V. A., Belyakova T. P. ACTIVE CONTROL SYSTEM OF TRUCK TIRES INTACTNESS STATUS.

The article presents the main results of research carried out to develop truck tires active intactness status control system. System is intended for alerting driver about cord separation of truck's tire then driving.

ACOUSTIC SIGNAL. SPECTRAL ANALYSIS. AMPLITUDE MODULATION. CREST FACTOR. ENVELOPE ANALYSIS.

Kazakov B. B., Ivanockij S. A. EFFECTIVE METHODS OF OPTICAL FLOW BASED OBJECT TRACKING IN VIDEO SEQUENCES.

A novel approach to motion tracking based on optical flow calculation has been developed. Using of regularized local methods in optical flow calculations allows to track objects in real time on the wide variety of hardware (including embed systems). A development environment for visual algorithm analysis has been developed in order to perform large number of numerical experiments.

TRACKING. OPTICAL FLOW. SEGMENTATION. CUDA. GPU. VISUAL PROGRAMMING.

Tomilin A. V., Stukov M. S., Stenin A. A. DYNAMIC MEMORY PROTECTION BY THE PROGRAM DIVERSITY.

The way to introduce diversity into C/C++ application's runtime buffer overflow attacks protection was examined. To thwart such attacks, several approches of the defence were developed, moreover, attacks classifications were suggested. The applicability area is the information security of the responsible computer systems.

BUFFER OVERFLOW. PROGRAM DIVERSITY. DYNAMIC MEMORY PROTECTION.

Tutygin V. S., Shedov S. V. NEW ALGORITHMS OF DETECTION AND PARAMETERS DEFINITION OF NOISED SIGNALS.

A new approach to the decision of a problem of detection and parameters definition noised echo-signals of a nuclear magnetic resonance spectrometers and the LFM-signals based on additional processing of the spectra received by means of fast Fourier Transform procedure and using nonconventional combination of normalized cross-correlation, spline-interpolation, rediscretization and approximations by special function are considered.

NUCLEAR MAGNETIC RESONANCE. NOISED ECHO-SIGNALS. LFM-SIGNALS. FAST FOURIER TRANSFORM. ADDITIONAL PROCESSING. NORMALIZED CROSS-CORRELATION. SPLINE-INTERPOLATION. REDISCRETIZATION. APPROXIMATION BY SPECIAL FUNCTION.

Nenashev A. V., Okhotnikov V. A. WORKING OUT THE METHOD OF THE ANALYSIS AND MODELING OF NONLINEAR RADIO ENGINEERING DEVICES.

The method of analysis and modeling of the radio engineering devices has offered. According to indemnification theorem nonlinear elements are replaced with the sources of a current or an EMP. By using a method of consecutive approximates the decision of a scheme has found. For example was taken the scheme of the amplifier cascade on SFET.

NONLINEAR DEVICES. MODELING. INDEMNIFICATION THEOREM. CONSECUTIVE APPROXIMATES. THE AMPLIFIER.

Karelin I. V., Karelin V. A. OPTIMAL ESTIMATION OF HIGH STABILITY HARMONIC WAVE SMALL FREQUENCY DEFLECTION.

The article presents the filter weight function synthesis algorithm for the average precision oscillator signals frequency error estimation. This algorithm does not require simplifying limitations relative to a noise spectral density model in contrast to the classic theory of liniar and nonlinear filtering approaches. It is shown that the estimation accuracy is equal to the maximum-likelihood method but the suggested filters realization requires less hardware and software expenses.

WEIGHT FUNCTION. AVERAGING FILTER. PRECISION GENERATOR. RADIO SYSTEM. MAXIMUM-LIKELIHOOD METHOD.

Konovalov V. N., Konovalov I. V. APPROACH TO MULTILEVEL OPTIMIZATION OF COMBINATIONAL CIRCUIT.

An algorithm of minimizing of logical circuits obtained as a result of multilevel decomposition is considered. Algorithm based on circuit analysis, deployment of underdetermined elements in logical orders generated on logical element's inputs and finding elements with consistent logical orders, which can be connected. This brings removal of appropriate parts of circuit.

COMBINATIONAL LOGICAL CIRCUIT. MINIMIZATION OF BOOLEAN FUNCTIONS. LOGICAL SYNTHESYS. DECOMPOSITION. OPTIMIZATION.

Seredinskiy Y. V., Gorbunov V. P., Novikov Y. N. THE USE OF FERRITE MATERIALS TO DECREASE CABLES INFLUENCE ON ANTENNA SYSTEMS RADIATION CHARACTERISTICS.

The research is dedicated to the problem of nearby passing cables parasitic influence on antenna systems in multilevel antenna complexes. To solve the problem the use of ferrite materials (ferrite beams and rods) was offered. Represented experimental research shows how the right use of ferrite materials helps to decrease the cables influence to characteristics of antenna systems.

ANTENNA SYSTEM. ELECTROMAGNETIC COMPATIBILITY. RADIATION CABLE. FERRITE ROD. FERRITE RING.

Sorotsky V. A., Utkin M. A. AN INFLUENCE OF RADIO TRANSMITTERS ANTENNAS ELECTROMAGNETIC INTERCONNECTION ON THE PROCESSES IN SWITCH GENERATORS.

The processes in radio transmitters caused by their mutual influence due to electromagnetic interconnection of antennas are covered. It has been shown that such interconnection can lead to poor performance of switch generators and even cause emergency modes.

RADIO TRANSMITTER. SWITCH GENERATOR. ANTENNA. MUTUAL INFLUENCE.

Sorotsky V. A., Utkin M. A. RESEARCH OF MULTICELL THYRISTOR GENERATORS CHARACTERISTICS USING DIGITAL MODEL.

The problems of multicell thyristor generators digital model development are covered. It presents the results of power characteristics modeling and their analysis.

THYRISTOR. MACROMODEL. GENERATOR. DIGITAL MODEL.

Makarova E. A. CONTROL OF MACROECONOMIC AGENT BEHAVIOR ON BASIS OF SELF-EQUALIZING MECHANISM OF FINANCIAL FLOWS.

Dynamic model of behavior of macroeconomic agent is suggested. Presented algorithm is available for correcting planned rate of financial resources expenditure on basis of information of accumulated stock volume. Three interpretation versions of functional diagram of dynamic model have been formed in terms of economic theory, control theory and system theory; analysis of these diagrams is carried out.

MACROECONOMIC AGENT. DYNAMIC MODEL. NON-EQUILIBRIUM REGIME. FLOW. STOCK.

Tokarev V. L., Orlov S. V. THE PROBLEM OF AUTOMATIONS OF PROCESS OF FORMATION OF LINGUISTIC MODEL FOR INTEGRATED DECISION-MAKING SUPPORT.

The method of linguistic model of system building for support decision making is considered. The technique of automatic determination of the best parameters of model is offered.

LINGUISTIC MODEL. ROUGH MODELLING. INTEGRATED DECISION SUPPORT SYSTEM.

Pobedash P. N. AN ANALYSIS OF TWO-CRITERIA OPTIMAL CONTROL PROBLEM OF REAL INVESTMENT UNDER UNCERTAINTY BY DISCRETE PRINCIPLE OF MAXIMUM.

The estimations of criteria value in the model of real investment with two economic agents under uncertainty are suggested. This estimation is founded by discrete principle of maximum. This values estimate the effectiveness of investment project.

INVESTMENT PROJECT. MULTISTAGE LINEAR OPTIMAL CONTROL PROBLEM.

Poptsov N. A., Bekurin D. B., Martsenyuk M. A. NEURONET OPERATING BY LOCOMOTION SYSTEM OF SELF-TRAINING BIO-OBJECT MODEL.

Self-training of bio-object (amoeba) locomotion system model which is operating by neuronet is considered. Amoeba model was prototyped by two ellipsoids connected with each other by thin channel. The length of the channel and the forms of ellipsoids are controlled by operating system. The motions of the model in the fluid are considered in Stocks approximation. Two scenarios of self-training are considered and their effectiveness are compared for different distances from target and initial object orientation angles.

SELF-TRAINING NEURONET. MODEL OF LOCOMOTION SYSTEM. AMOEBA SWIMMING MODELLING.

Kudrin M. U., Sokolov E. V., Tormasov A. G. THE RACE CONDITION DETECTION USING THE GRAPH of JOINT THREAD EXECUTION.

The method to detect race conditions between threads, which use a shared memory is presented. This method is applicable when incorrect behaviour of a program could be determined by a final state of threads' execution. For these instances it allows finding guaranteed all the race conditions without full search of cases of joint thread execution.

MULTIPROCESSOR PROGRAMMING, RACE CONDITION, THREAD SYNCHRONIZATION.

Krupenko D. A. METHODS OF ADAPTATION OF SYSTEM OF DESIGNING AUTOCAD FOR WORKING OUT OF THE DESIGN OF MICROASSEMBLAGES RF.

In article methods of adaptation of system of automated designing AutoCAD for a possibility of its work on designing RF of devices with use of dynamic blocks are considered.

THE DESIGN DOCUMENTATION. AUTOCAD. VISUAL LISP. THE DYNAMIC BLOCK. MICROASSEMBLAGE. MODULE RF.

Solnushkin K. S. THE SYSTEM OF ACCOUNTING FOR ACCESS TO THE LICENSED SOFTWARE IN CENTER FOR COLLECTIVE USE.

An open, modular architecture for a license usage accounting software system is presented. Accumulation of accounting data allows for a retrospective analysis as well as detecting statistical regularities. The proposed system also interfaces with job schedulers so that to minimize the risk of unsuccessful task launching due to the lack of required licenses.

LICENSED SOFTWARE. MONITORING. ACCOUNTING. JOB SCHEDULER.

Suhov A. A. THE EVALUATION OF THE EMOTIONAL COMPONENT OF THE COMPUTER SIMULATORS.

The method for estimating human-machine interface using multiple neural networks, which allows you to more objectively assess the quality of the interface is described.

INTERFACE. EVALUATION OF THE INTERFACE. NEURAL NETWORKS. COMPUTER SIMULATORS.

Bratin D. Y. SPECIALITIES OF THE FDL V. 1.0 THE FRONTEND DESCRIPTION LANGUAGE.

The article describes the dialect based on XML, designed for declarative description frontend web resource. Language makes it possible to describe the structure of the resource capabilities of the software and page templates. It is shown that his distinction is a convenient model of abstraction from the implementation of applications and visual aids to access its components.

DECLARATIVE LANGUAGE. USER-INTERFACE DESCRIPTION LANGUAGE. WEB-APPLICATION DEVELOPMNET.

Gilmullin T. M., Anikin I. V. INFORMATION SECURITY OBJECTS MODELING FOR RISK ASSESSMENT.

A formalized models of information security objects suggested for the task of information security risks assessment. We suggested conceptual model for the subject «Information security risk». Some fundamental theorems are formulated and proved.

SIMULATION. INFORMATION SECURITY. RISK ASSESSMENT. FUZZY SETS.

Zraenko A. S., Aksyonov K. A., Van Kai. COALITION MODEL OF MULTI-AGENT PROCESS OF RESOURCES TRANSFORMATION.

The problem of effective administrative decisions acceptance support by means multi-agent modeling is considered. For its decision use of the developed coalition model on the basis of the resources transformation processes mechanisms is offered.

DECISION-MAKING SUPPORT. MULTI-AGENT SYSTEMS. MODELING. OPTIMIZATION. PLANNING. RESOURCES TRANSFORMATION PROCESSES.

Arzumanyan M. Y. ENTERPRISE DESCREPTIONAL MODELS AND THEIR MONITORED VALUES.

The core enterprise description models is examined. Made an attempt to define monitored values for each model. Proposed the new model based on the service description, developed some measures of the model.

INFORMATION ECONOMY. MONITORING. ENTERPRISE. ENTERPRISE MODELS. SERVICE CONCEPTION.

Teihrib A. P., Dorosinsky L. G. COMPARATIVE REVIEW OF MODELS OF LOGICAL TIME.

The several models of logical time are described. Models are represented in unified formalism. Properties of every model are defined.

LOGICAL TIME. DISTRIBUTED SYSTEMS. PLAUSIBILITY. ISOMORPHISM. MODELLING.

Ikryanov I. I. MODELS of MOVEMENT of INDICATORS of the COURSE And GLIDEPATH With PENDULAR SYSTEM of STABILIZATION of OPTICAL AXES.

The objects of research in this article are the course displays and glidepath indicators (CD/GI) assigned for visual landing of helicopters on naval vessels boards during the day and night, in simple and adverse weather conditions. The mathematical models of CD/GI for regular and irregular ship motion are considered in this article. The comparison of these models shows the good coincidence between the results obtained using the each model.

MOVEMENT MODEL. OPTICAL INDICATORS. THE STABILIZATION ERROR. CASUAL PROCESS. APPROXIMATION.

Nagorniy V. S. BASE COMPOUNDINGS NANOSTRUCTURED SILVER OF WORKING LIQUIDS FOR ELECTRICAL DROP-JET TECHNOLOGIES.

High pressure electrical drop-jet layout plating method on dielectric substrate examined. On the grounds of carried out patent search and bibliometric analysis base compounding of working liquids with silver nanoparticles generation method suggested.

ELECTRICAL DROP-JET METHOD. CONDUCTIVE TOPOLOGY. WORKING FLUID. NANOPARTICLES OF SILVER.

Vetrinsky Yu. A. THE EXPERIENCE OF BLENDED LEARNING TECHNOLOGY USE IN PRACTICE OF TECHNICAL SUBJECTS TEACHING.

One approach to blended learning organization in technical university is considered. The results of experimental investigations, confirmed the proposed approach effectiveness, are presented.

BLENDED LEARNING. EDUCATIONAL COURSE. LEARNING MANAGEMENT SYSTEM. REMOTE ACCESS.

Kunturova N. B. METHOD OF SHAPING OF THE STRUCTURED GROUPS OF CANDIDATES AT THE ENTERING IN MAGISTRACY WITH PREDICTABLE SUCCESS OF THEIR EDUCATION.

The method of the groups candidate shaping at the entering in the high schools MA course is offered. Use of this method provides structuring of the group of candidates and forecasts success of their education and training in the high schools MA course.

PROFESSIONAL COMPETENCE. QUALIMETRY. EDUCATIONAL TRAJECTORY. QUALITY TRAINING. FORECAST. SUCCESSFUL PROFESSIONAL ACTIVITY.

Kunturova N. B. PECULIARITIES OF PROFESSIONAL COMPETENCE OF SPECIALISTS IN PROCESS OF EDUCATION IN HIGH SCHOOL.

An original estimation apparatus of professional competence in the specialists training system is offered. This apparatus includes the calculation algorithm of the professional competence factor, which allows to diagnose the quality of the professional training and is an efficient instrument (the method), used for estimation of specialist's readiness degree for productive decision of the professional tasks. Use of the algorithm in practice of the high school management activity promotes revealing of the development prospects of the trained specialists and allows solving problem of effectiveness increasing of the graduates use.

PROFESSIONAL COMPETENCE. QUALIMETRY. EDUCATIONAL TRAJECTORY. QUALITY TRAINING. FORECAST. SUCCESSFUL PROFESSIONAL ACTIVITY.