

**МИНИСТЕРСТВО ОБРАЗОВАНИЯ И НАУКИ РОССИЙСКОЙ
ФЕДЕРАЦИИ**

ФЕДЕРАЛЬНОЕ АГЕНТСТВО ПО ОБРАЗОВАНИЮ

**САНКТ-ПЕТЕРБУРГСКИЙ ГОСУДАРСТВЕННЫЙ УНИВЕРСИТЕТ
ИНФОРМАЦИОННЫХ ТЕХНОЛОГИЙ, МЕХАНИКИ И ОПТИКИ**



ПОБЕДИТЕЛЬ КОНКУРСА ИННОВАЦИОННЫХ ОБРАЗОВАТЕЛЬНЫХ ПРОГРАММ ВУЗОВ

**НАУЧНО-ТЕХНИЧЕСКИЙ
ВЕСТНИК**

**САНКТ-ПЕТЕРБУРГСКОГО ГОСУДАРСТВЕННОГО
УНИВЕРСИТЕТА ИНФОРМАЦИОННЫХ ТЕХНОЛОГИЙ,
МЕХАНИКИ И ОПТИКИ**

март–апрель 2010

№ 2(66)



ГЛАВНЫЙ РЕДАКТОР

д.т.н., профессор В.О. Никифоров

РЕДАКЦИОННАЯ КОЛЛЕГИЯ

д.т.н., профессор А.А. Бобцов, д.т.н. А.В. Бухановский,
д.т.н., профессор В.А. Валетов, д.ф.-м.н., ст.н.с. Т.А. Вартанян,
д.т.н. М.А. Ган, д.т.н., профессор Ю.А. Гатчин, д.т.н., профессор А.В. Демин,
к.т.н., доцент Н.С. Кармановский (заместитель главного редактора),
д.ф.-м.н., профессор Ю.Л. Колесников, д.ф.-м.н., профессор С.А. Козлов,
д.т.н., профессор А.Г. Коробейников, д.т.н., профессор В.В. Курейчик,
д.т.н., доцент Л.С. Лисицына, к.т.н., доцент В.Г. Мельников,
д.т.н., профессор Ю.И. Нечаев, д.т.н., профессор Н.В. Никоноров,
д.т.н., профессор А.А. Ожиганов, д.т.н., профессор П.П. Парамонов,
д.ф.-м.н., ст.н.с. Е.Ю. Перлин, д.т.н., профессор И.Г. Сидоркина,
д.т.н. О.А. Степанов, д.т.н., профессор В.Л. Ткалич,
д.т.н., профессор А.А. Шалыто, д.т.н., профессор Ю.Г. Якушенков

Секретарь – Г.О. Артемова

Редактор – к.т.н., ст.н.с. Н.Ф. Гусарова

Перевод – Н.Г. Лещикова

Адрес: 197101, Санкт-Петербург, Кронверкский пр., 49, СПбГУ ИТМО,
комн. 330

Телефон / факс (812) 233 45 51

[http: //books.ifmo.ru/ntv](http://books.ifmo.ru/ntv)

E-mail: karmanov@mail.ifmo.ru

1. ОПТИЧЕСКИЕ И ОПТИКО-ЭЛЕКТРОННЫЕ СИСТЕМЫ.	
ОПТИЧЕСКИЕ ТЕХНОЛОГИИ.....	5
Воронин А.А., Смирнова Е.В., Смирнов А.П. К вопросу идентификации пород древесины с применением методов анализа спектров.....	5
Парфенов В.А., Геращенко А.Н., Геращенко М.Д., Григорьева И.Д. Лазерная очистка исторических памятников	11
Мешковский И.К., Олехнович А.О. Исследование процесса образования слоистости при выщелачивании натрийборосиликатных стекол.....	18
2. ФОТОНИКА И ОПТОИНФОРМАТИКА	23
Буяновская Е.М., Козлов С.А. Закономерности взаимодействия встречных световых волн из малого числа колебаний в нелинейных диэлектрических средах	23
3. АНАЛИЗ И СИНТЕЗ СЛОЖНЫХ СИСТЕМ. ЭЛЕМЕНТЫ И УСТРОЙСТВА ТЕХНИЧЕСКИХ СИСТЕМ	30
Дударенко Н.А., Полякова М.В., Ушаков А.В. Алгебраическая организация условий обобщенной синхронизируемости многоагрегатных динамических объектов.....	30
Ревин М.С., Савёлов Н.С. Алгоритм ускоренного повторного решения систем линейных алгебраических уравнений и его использование при математическом моделировании электронных устройств.....	37
4. КОМПЬЮТЕРНЫЕ СИСТЕМЫ И ИНФОРМАЦИОННЫЕ ТЕХНОЛОГИИ.....	43
Бессмертный И.А. Теоретико-множественный подход к логическому выводу в базах знаний.....	43
Сидоркина И.Г., Егошин А.В., Шумков Д.С., Кудрин П.А. Прогнозирование сложных сигналов на основе выделения границы реализаций динамических систем	49
Зубакин И.А., Фахми Ш.С. Классификация нестационарных изображений и разработка методики оценки алгоритмов кодирования источника.....	54
Коротков К.Г., Орлов Д.В., Величко Е.Н., Гатчина Ю.Ю. Процедура проведения измерений объектов природной среды методом газоразрядной визуализации.....	59
Хританков А.С. Оценка эффективности распределенных систем при решении задач переменного размера	66
Жукова А.Р., Мазнин М.А. Акторное расширение языка JAVA в среде MPS	72
5. МАТЕРИАЛОВЕДЕНИЕ И НАНОТЕХНОЛОГИИ	78
Чивилихин С.А., Голубок А.О., Мухин И.С. Рост нановискера под воздействием электронного пучка: математическая модель.....	78
6. ТЕПЛОФИЗИКА И ТЕОРЕТИЧЕСКАЯ ТЕПЛОТЕХНИКА.....	84
Шибут И.И. Применение суперкомпьютерных технологий для проведения теплового анализа радиоэлектронных устройств.....	84
7. СИСТЕМЫ АВТОМАТИЗИРОВАННОГО ПРОЕКТИРОВАНИЯ	90
Зубакин И.А., Фахми Ш.С., Шагаров С.С. Аппаратно-программное проектирование сложных функциональных блоков с использованием систем на кристалле	90
Москаленко С.В., Гатчин Ю.А. Оптимизация алгоритмов идентификации графового изоморфизма	98
8. МЕТОДЫ И СИСТЕМЫ ЗАЩИТЫ ИНФОРМАЦИИ	104
Хусаинова Э.Р. Модель оценки угроз информационной безопасности на основе материалов хакерских конференций.....	104
Спивак А.И. Оценка эффективности атак злоумышленника в процессе построения его модели	108
9. ЭКОНОМИКА И ФИНАНСЫ. МЕНЕДЖМЕНТ	113
Васюхин О.В., Павлова Е.А. Стратегия формирования и развития инновационного потенциала промышленного предприятия	113
КРАТКИЕ СООБЩЕНИЯ	120
SUMMARY	

1. OPTICAL AND OPTICAL ELECTRONIC SYSTEMS. OPTICAL TECHNOLOGIES	5
A. Voronin, E. Smirnova, A. Smirnov. On the issue of wood species identification with the use of spectrum analysis methods	5
V. Parfenov, A. Gerashchenko, M. Gerashchenko, I. Grigoryeva. Laser cleaning of historical monuments	11
I. Meshkovskiy, A. Olekhovich. Process of layering formation research at acid leaching of alkali borosilicate glass	18
2. PHOTONICS AND OPTICAL INFORMATICS	23
E. Buyanovskaya, S. Kozlov. Interaction laws of counter-propagating light waves consisting of few oscillations in the nonlinear dielectric media	23
3. ANALYSIS AND SYNTHESIS OF COMPLEX SYSTEMS. UNITS AND APPLIANCES OF ENGINEERING SYSTEMS	30
N. Dudarenko, M. Polyakova, A. Ushakov. Algebraic conditions of total synchronism for multivariate dynamic plants	30
M. Revin, N. Savyolov. Algorithm of accelerated resetting solution of linear algebraic equation systems and its use in mathematical modeling of electronic devices	37
4. COMPUTER SYSTEMS AND INFORMATION TECHNOLOGIES	43
I. Bessmertnyy. A set-theoretic approach to the logical inference in knowledge bases	43
I. Sidorkina, A. Egoshin, D. Shumkov, P. Kudrin. Complex signal prediction based on dynamic system allocation limits	49
I. Zubakin, Sh. Fahmi. Non-stationary images classification and development of source coding algorithms estimation method	54
K. Korotkov, D. Orlov, E. Velichko, Yu. Gatchina. Standard procedure of natural environment objects measuring by the GDV/EPC method	59
A. Khritankov. Performance evaluation model for distributed systems and tasks with variable size	66
A. Zhukova, M. Mazin. Actor extension of Java language in MPS environment	72
5. MATERIALS TECHNOLOGY AND NANOTECHNOLOGIES	78
S. Chivilikhin, A. Golubok, I. Mukhin. Nanowhisker growth under electron beam impact: mathematical model	78
6. THERMAL PHYSICS AND THEORETICAL HEAT ENGINEERING	84
I. Shibut. Application of supercomputer technologies for the thermal analysis of radio-electronic devices	84
7. COMPUTER-AIDED DESIGN SYSTEMS	90
I. Zubakin, Sh. Fahmi, S. Shagarov. Hardware-software design of complex functional units on "system on chip"	90
S. Moskalenko, Yu. Gatchin. Optimization of graph isomorphism identification algorithms	98
8. INFORMATION SECURITY METHODS AND SYSTEMS	104
E. Khusainova. Threat evaluation model of information safety on the basis of hacker conference materials	104
A. Spivak. The efficiency estimation of the intruder attacks in the process of his model creation	108
9. ECONOMICS AND FINANCES. MANAGEMENT	113
O. Vasyukhin, E. Pavlova. Strategy of formation and development of the industrial enterprise innovative potential	113
BRIEF REPORTS	120
A. Gus'kov, A. Dunaev, F. Podtelkin, A. Bukhanovsky. Intelligent technologies of users' support in the field of numerical quantum-chemical computing of nanostructure and nanomaterials modeling	120
E. Kaliteevskaya, V. Krutyakova, T. Razumova, A. Starovoytov. Study of mechanisms responsible for the molecular layers composition formation	121
A. Korobeynikov, K. Dukel'skiy, E. Ter-Nersesyants, Yu. Gatchin. The big core photonic crystal optical fiber development	122
Z. Bekyasheva, V. Vasil'ev, A. Vostrikov, A. Pavlov. Statistical properties of hypotheses caused by two-layer neural network with bidirectional interconnections, implemented by Fourier-holography technique	123
V. Gilina, D. Sachkov, V. Khramov. About multiwave generation of Er:YLF laser with diode pumping and passive Fe:ZnSe Q-switch	124
SUMMARY	125

УДК 681.785.554

**К ВОПРОСУ ИДЕНТИФИКАЦИИ ПОРОД ДРЕВЕСИНЫ
С ПРИМЕНЕНИЕМ МЕТОДОВ АНАЛИЗА СПЕКТРОВ**

А.А. Воронин, Е.В. Смирнова, А.П. Смирнов

Описан прибор для идентификации пород древесины и концепция построения автоматизированной системы распознавания древесины по породам на основе методов анализа спектральной информации.

Ключевые слова: идентификация, спектр, спектрофотометр, дифракционная решетка, фотометрический шар, порода древесины.

Введение

В настоящее время существует проблема идентификации грузов, провозимых через таможенные переходы. Фактически, сотрудник таможенных органов (таможенник) на посту не может произвести анализ груза на соответствие декларации. При появлении сомнений ему необходимо обратиться к экспертам для подтверждения или опровержения информации о соответствии перевозимого груза накладной на этот груз. Эта проблема касается всех видов грузов. Это и сыпучие материалы, и животные, и различного рода жидкие материалы (топливо, химикаты), и древесина. Особенно актуальной эта задача является при экспорте древесины. Измерительных приборов, определяющих или проверяющих наименование породы, на оснащении таможенных органов нет.

Традиционные классические методы спектрального анализа не дают точного однозначного ответа о принадлежности исследуемой древесины определенной породе, так как по своему химическому составу древесина практически одинакова, т.е. универсального метода распознавания породы древесины на базе химического состава не существует. В этом случае требуется комплексный подход, учитывающий индивидуальные особенности одних пород перед другими не только в сфере их химического состава, но и относительно текстуры среза или спектра диффузного отражения в определенном диапазоне длин волн.

Статья посвящена решению задачи идентификации пород древесины путем анализа спектральных данных, полученных при помощи портативного прибора, включающего в себя спектрофотометр и анализатор спектров.

Описание и принцип работы прибора идентификации

В рамках поставленной задачи на первый план выходит разработка и создание портативного прибора для получения спектральной информации об исследуемом объекте. Принцип действия разработанного прибора для идентификации лиственных и хвойных пород древесины «Кедр» основан на измерении и последующей математической обработке спектров видимого [1] и инфракрасного светового потока, отраженного от поверхности объекта (древесины). На рис. 1 приведена упрощенная структурная схема работы портативного прибора.

В основе конструкции прибора лежит конструкция спектрофотометра [2], упрощенная оптическая схема которого представлена на рис. 2.

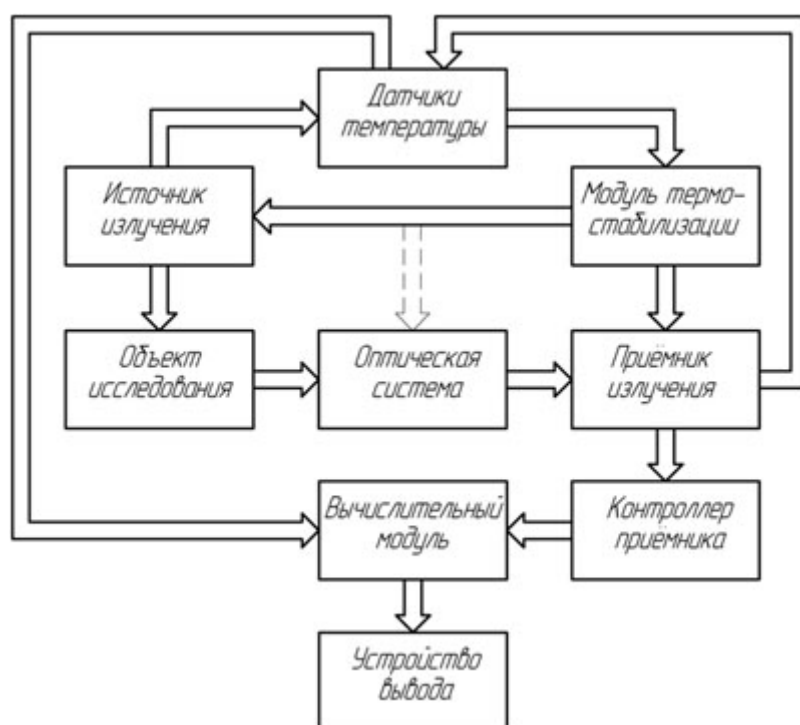


Рис. 1. Структурная схема портативного прибора идентификации древесины

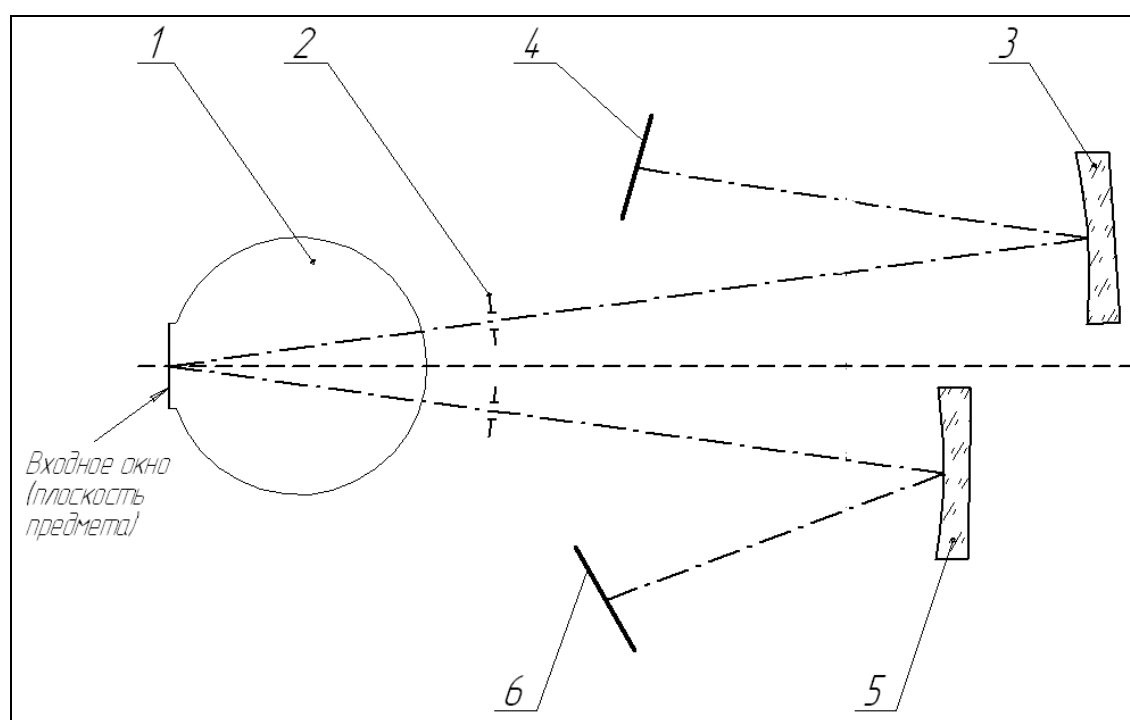


Рис. 2. Оптическая схема прибора для идентификации пород древесины:

- 1 – фотометрический шар со встроенными источниками излучения;
- 2 – входная щель с электромеханической шторкой; 3 – вогнутая стигматическая дифракционная решетка ($N=300 \text{ мм}^{-1}$, $R=125 \text{ мм}$); 4 – ПЗС-линейка производства фирмы Toshiba (3048 пикселей); 5 – вогнутая стигматическая дифракционная решетка ($N=100 \text{ мм}^{-1}$, $R=100 \text{ мм}$); 6 – фотогальванический приемник ИК-излучения типа ФД793-0,3 (96 элементов)

Оптическая схема содержит всего один оптический элемент – вогнутую дифракционную решетку. Конструктивно прибор выполнен в виде двух спектрометров, один из которых работает в видимой области спектра (0,4–0,9 мкм), а второй – в инфракрасной (0,9–4,0 мкм). Инфракрасный канал добавлен для повышения надежности контроля (идентификации) естественных объектов, а также для дальнейшей возможности определения влажности исследуемого естественного объекта (древесины). Световой поток от источников излучения видимого и инфракрасного диапазонов, размещенных в одном фотометрическом шаре, падает на торцевую поверхность исследуемого объекта, отражается от нее и через входную щель попадает на дифракционные решетки. Диспергирующие элементы разлагают сигнал, полученный от исследуемого образца, в линейный спектр излучения и проецируют его не на отдельные фотоприемники, а на ПЗС-линейку приемников. С приемников сигнал поступает на встроенный микроконтроллер, где происходит его первичная обработка и преобразование в цифровой вид для дальнейшей передачи в персональный компьютер, работающий в операционной среде Windows XP. Измеренный спектр отражения отображается на экране прибора.

Блок регистрации спектральной информации (спектрофотометр и вычислительный модуль) для портативного прибора идентификации лесо- и пиломатериалов лиственных и хвойных пород древесины предназначен для сбора и формирования базы (банка) данных спектральных характеристик различных пород деревьев. Прибор обеспечивает оперативное измерение оптических (спектральных) характеристик различных пород (образцов) древесины в виде лесо- и пиломатериалов в диапазоне длин волн 0,4–4,0 мкм для последующей обработки полученных данных с использованием комплекта программных средств (специального программного обеспечения) и выдачей результатов идентификации в буквенно-цифровом виде на устройство вывода ППИ «Кедр» в любое время года при различных погодных условиях.

Первый модуль в персональном компьютере – модуль первичной обработки спектра. В нем производится коррекция полученного спектра на спектр темнового сигнала, усреднение, нормирование и аппроксимация полученного спектра, затем передача его в модуль, в котором реализуется алгоритм анализатора спектральной информации.

Краткие сведения о математическом алгоритме обработки спектральной информации

На рис. 3 изображен примерный вид спектров древесины породы Береза повислая, полученный при помощи прибора. Как видно, существует общая тенденция, которая позволяет подобрать возможную тренд-линию. В то же время имеются различия и в абсолютной величине каждого из спектров, что говорит об отличиях свойств одного образца исследования от другого.

Однако интерес представляют не столько спектры отдельных пород, сколько сравнение их между собой. Так, на рис. 4 представлены усредненные спектры различных пород древесины. Как видно из рисунка, спектры различных пород различаются как по абсолютному значению, так и по уровню наклона кривых (первой производной). В связи с этим был разработан алгоритм распознавания, который учитывает статистическую спектральную информацию об отдельных породах и строит так называемые «коридоры» максимальных и минимальных значений параметров (абсолютных значений спектров и первых производных). Все спектры образцов одной породы формируют коридоры, представляющие собой нижнюю и верхнюю границы спектров. Далее алгоритм оценивает в процентах соответствие полученного спектра древесины породам, находящимся в базе данных. Коридоры спектров для некоторых пород древесины представлены на рис. 5.

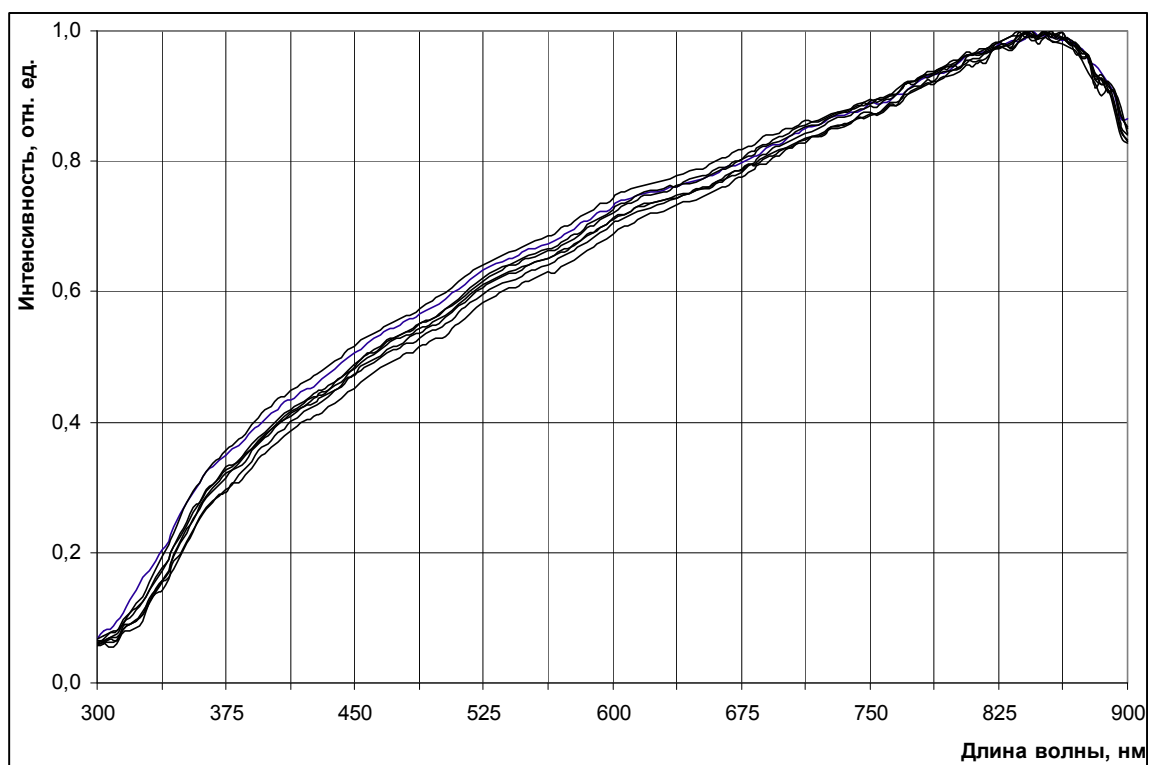


Рис. 3. Спектры диффузного отражения древесины породы Береза повислая (8 образцов)

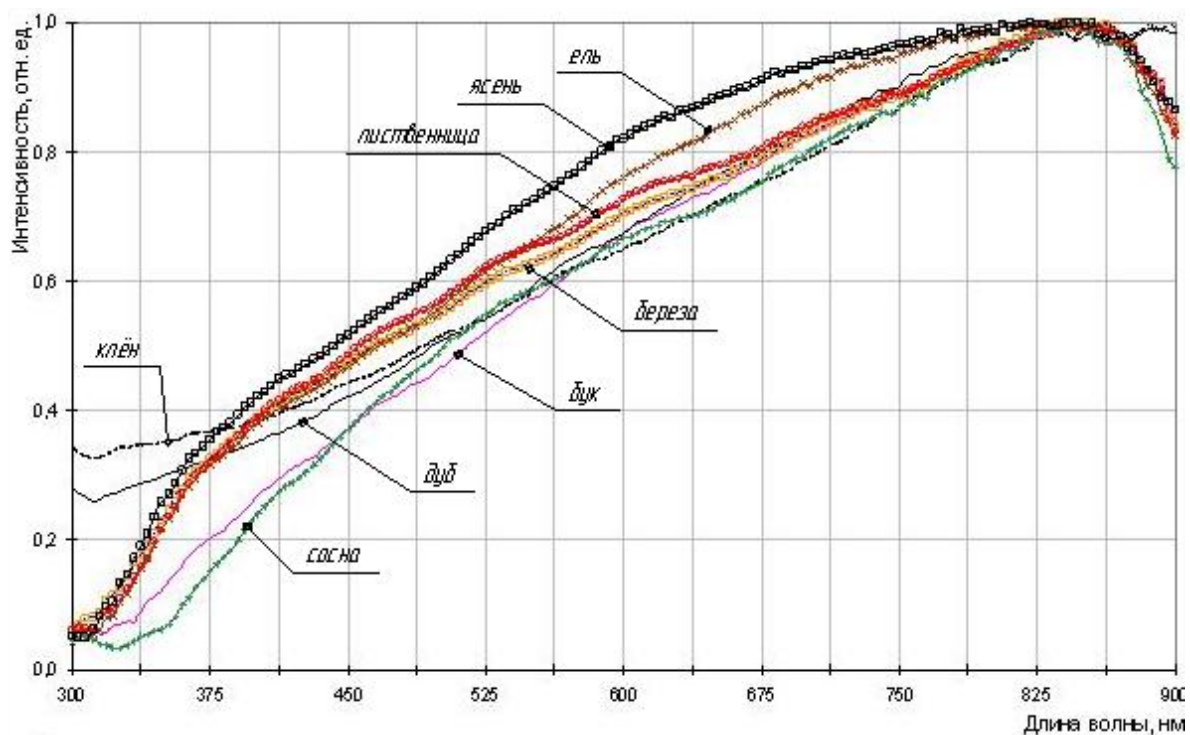


Рис. 4. Спектры диффузного отражения различных пород древесины

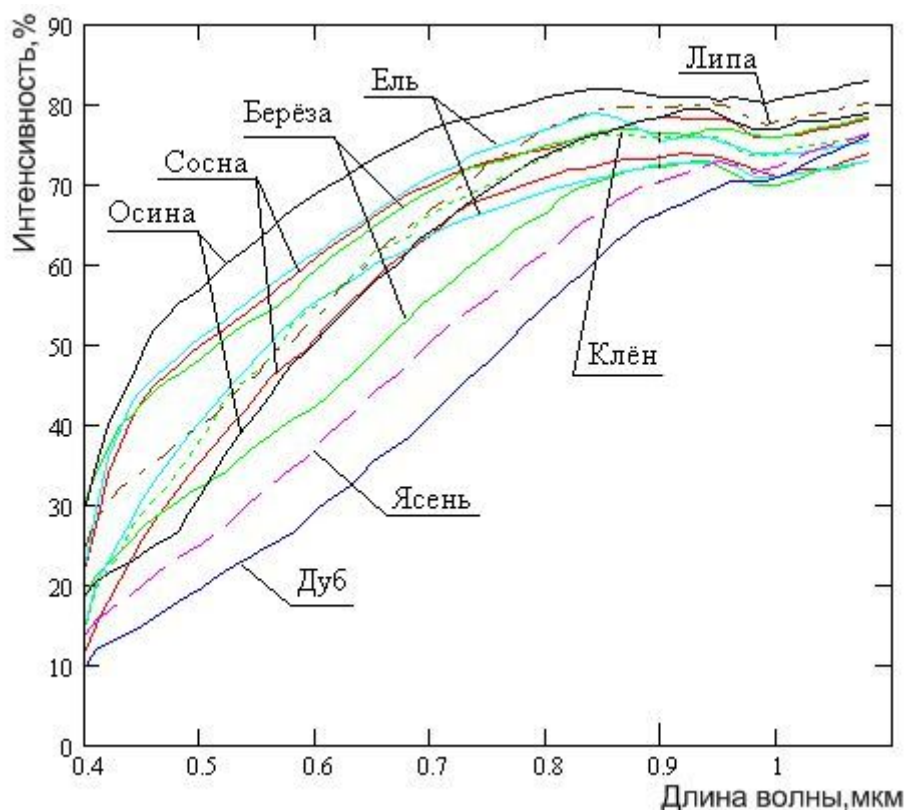


Рис. 5. Коридоры некоторых пород древесины

Алгоритм сравнивает полученный спектр со всеми спектрами пород древесины, хранящимися в базе данных по породам, и дает вероятностную оценку принадлежности образца к той или иной породе древесины. Принцип работы алгоритма базируется на последовательном сравнении спектра от исследуемого образца с так называемыми «эталонными» спектрами как по абсолютной величине, так и по первой производной (наклон кривой спектра). Каждая отличительная особенность спектров одной породы от всех других называется «признаком» породы. Количество признаков каждой из пород в базе данных определяется при составлении базы данных спектров алгоритмом распознавания, а в процессе распознавания породы алгоритм проверяет спектр исследуемого образца на соответствие признакам каждой из пород в базе данных. В результате алгоритм вычисляет соответствие спектра исследуемого образца каждой из пород, находящихся в базе данных, по формуле:

$$P_n = \frac{N_+}{N_{\text{общ}}} \cdot 100\%,$$

где N_+ — количество признаков, которым удовлетворяет спектр исследуемого объекта, $N_{\text{общ}}$ — общее количество признаков для породы, P_n — степень соответствия (уверенность), выраженная в процентах для каждой из пород. Экспериментально установлено, что в случае, если образец не соответствует породе, уверенность всегда меньше 25%. Если образец абсолютно точно соответствует породе из базы данных, то уверенность всегда выше 33%. Если рассчитанная уверенность находится в рамках между 25% и 33%, прибор просит повторить измерение или попробовать другие образцы для исследования.

После отработки алгоритма результат передается в модуль взаимодействия с человеком, в котором принимается решение о типе выводимой информации по следующему алгоритму.

1. В случае удачного распознавания прибор обращается к базе голосовых сообщений и воспроизводит название распознанной породы.
2. В случае неудачного распознавания прибор обращается к базе голосовых сообщений и предлагает на выбор различные варианты действий:
 - а) произвести повторное измерение;
 - б) завершить измерения с отрицательным результатом – порода не соответствует эталонным образцам либо вообще не является древесиной;
 - в) произвести диагностические операции с прибором (замена лампы, чистка оптики).

Следует отметить, что для полноценной работы алгоритма ему необходима обширная база данных по породам, т.е. исчерпывающее описание характеристик различных видов и модификаций одной и той же породы. Если же база данных недостаточна, возможны погрешности и даже ошибки в определении породы, поэтому немаловажным фактором становится сбор статистической информации.

Заключение

Разработан портативный прибор идентификации, позволяющий проводить идентификацию породы древесины. При участии авторов составлена база данных, изучена достоверность идентификации объектов исследования. Достоверность идентификации зависит от множества факторов. Показано, что для повышения уверенности распознавания необходимо не только иметь обширную базу данных объекта исследования, но и предусмотреть температурную коррекцию результатов измерения.

По результатам многократных исследований была выявлена также необходимость доработки как программных, так и аппаратных средств, однако уже на данном уровне принципиально вопрос об автоматической идентификации породы древесины решен практически и реализован в приборах.

Литература

- 1 Колгин Е.А., Ухов А.А., Воронин А.А., Кострюков А.А., Черноглазов В.С., Савушкин А.В. Спектрометрическое устройство для идентификации пород древесины // Петербургский журнал электроники. – 2008. – № 2 (55) – 3 (56). – С. 116–120.
- 2 Колгин Е.А., Ухов А.А., Савушкин А.В. Спектрометры на основе полихроматора и одномерной ПЗС-матрицы: опыт разработки и применения // Петербургский журнал электроники. – 2008. – № 2 (55) – 3 (56). – С. 120–127.

<i>Воронин Андрей Анатольевич</i>	– Санкт-Петербургский государственный университет информационных технологий, механики и оптики, аспирант, V-electronics@mail.ru
<i>Смирнова Елена Викторовна</i>	– Санкт-Петербургский государственный университет информационных технологий, механики и оптики, аспирант, smirnova-elen@yandex.ru
<i>Смирнов Александр Павлович</i>	– Санкт-Петербургский государственный университет информационных технологий, механики и оптики, доктор технических наук, профессор

УДК 681.7.069.223:621.791.72:903.32

ЛАЗЕРНАЯ ОЧИСТКА ИСТОРИЧЕСКИХ ПАМЯТНИКОВ

В.А. Парфенов, А.Н. Геращенко, М.Д. Геращенко, И.Д. Григорьева

Представлены результаты экспериментальных работ по лазерной очистке исторических памятников из меди. Очистка поверхности производилась при помощи импульсного Nd:YAG лазера (длина волны 1064 нм, длительность импульса 30–110 мкс, плотность энергии 10–25 Дж/см²). Эффективность лазерной обработки контролировалась с помощью рентген-флуоресцентного спектроанализатора. Проведенные исследования продемонстрировали высокую эффективность лазерной очистки и возможность применения лазера с указанными параметрами при реставрации выколотной медной скульптуры.

Ключевые слова: лазерная очистка, импульсный Nd:YAG лазер, памятники истории и культуры, реставрация, рентген-флуоресцентный спектральный анализ.

Введение

В последние годы лазерная техника находит все более широкое применение в реставрации произведений искусства [1, 2]. Одно из основных применений лазеров в данной области – это очистка поверхности исторических памятников от загрязнений и природных наслоений [3]. В статье представлены результаты экспериментальных работ по лазерной очистке фрагментов экстерьерных памятников Санкт-Петербурга, которые проводились в 2007–2009 гг. в рамках сотрудничества Санкт-Петербургского государственного электротехнического университета с рядом музеев и реставрационных компаний. Основная цель работы состояла в отработке технологии лазерной очистки и выработке практических рекомендаций для специалистов-реставраторов по ее применению при работе с конкретными видами памятников.

В качестве комментария следует отметить, что хотя впервые лазеры были использованы в реставрации еще в начале 1970-х годов [4], в настоящее время метод лазерной очистки хорошо отработан только применительно к каменным памятникам, а при работе с историческими объектами из металлов он нуждается в дополнительном изучении и отработке технологии. Однако даже при очистке камня до сих пор остается много открытых вопросов, в том числе проблема удаления биогенных загрязнений. Заметим также, что, несмотря на значительный опыт применения лазеров в реставрации за рубежом, в России применение лазерных технологий в этой области началось сравнительно недавно. Таким образом, поставленные в работе задачи исследования являются актуальными как с научной, так и с практической точек зрения.

Традиционные методы очистки памятников в реставрации

Очистка поверхности произведений искусства и иных объектов культурно-исторического наследия от загрязнений и природных наслоений является одним из важнейших этапов реставрации. Традиционные методы очистки памятников основаны, главным образом, на использовании механического и химического воздействия, в том числе разнообразных ручных инструментов (скребков, кисточек, щеток и т.д.), растворителей и моющих средств [5]. К сожалению, такие методы не всегда позволяют добиться желаемых результатов. Основные их недостатки – возможное изменение, повреждение и даже удаление части исторического материала, что является неприемлемым в практике реставрации. Другая проблема заключается в том, что при очистке механическим способом реставраторы зачастую сталкиваются с невозможностью удаления загрязнений с труднодоступных участков объекта (складки одежды у скульптур, орнаментальный рельеф на фасадах исторических зданий и т.д.). Кроме того, применение моющих средств и растворителей и пролонгирующий эффект их воздействия может вызвать дальнейшее изменение физико-химических свойств и ухудшение внешнего

облика исходного материала. Как будет видно из последующего изложения, технология лазерной очистки представляет собой альтернативу традиционным методам реставрации.

Физические принципы лазерной очистки памятников

В основе технологии лазерной очистки лежит локальный нагрев слоя загрязнений под воздействием лазерного излучения с последующим его удалением с поверхности памятника в результате фотоабляции. Фотоабляция (греч. *φος* – свет, позднелат. *ablatio* – отнятие) представляет собой процесс удаления частиц загрязняющего вещества с поверхности обрабатываемого объекта в результате поглощения им высокоинтенсивного света лазера. Физическая природа этого явления довольно сложна и сочетает в себе как испарение инородных частиц, так и их удаление в твердой фазе [3, 6]. В упрощенном виде процесс удаления загрязнений в результате быстрого термического поглощения энергии лазерного излучения частицами загрязняющего вещества схематически показан на рис. 1.

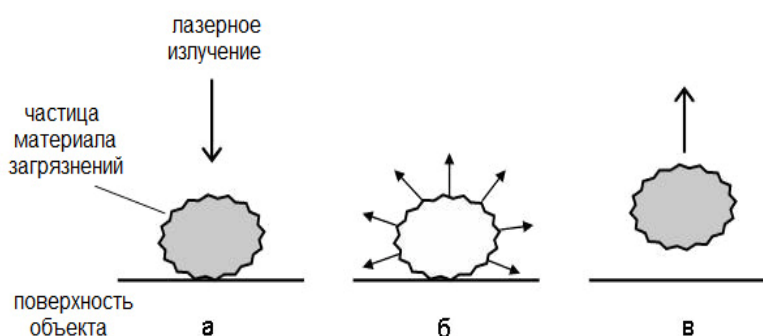


Рис. 1. Принцип лазерной очистки

Различают два вида фотоабляции: фототермическую и фотохимическую. При очистке камня и металлов удаление загрязнений происходит, в основном, в результате фототермической абляции, а при очистке произведений живописи, напротив, преобладают процессы фотохимической абляции. В обоих случаях для того, чтобы при лазерной очистке происходило селективное удаление одних лишь поверхностных загрязнений без повреждения субстрата, следует принимать в расчет его физико-химические и оптические свойства, что на практике приводит к необходимости применения вполне определенных типов лазеров и индивидуального подбора их выходных параметров в каждом конкретном случае.

Для удаления загрязняющего вещества необходимо, чтобы плотность энергии лазерного излучения превышала некоторый минимальный пороговый уровень, при котором начинается процесс абляции. Но при этом работать можно только при таких уровнях плотности энергии излучения, которые лежат ниже порогового уровня абляции подложки (т.е. материала самого памятника). Это наглядно показано на рис. 2, где приведены характерные кривые, показывающие качественный характер зависимости степени фотоабляции слоя загрязнений и подложки при воздействии на них излучением импульсного лазера.

Помимо плотности энергии, другим важным параметром является длина волны излучения лазера. Ее нужно выбирать таким образом, чтобы слой загрязнений имел более высокий (по сравнению с подложкой) коэффициент оптического поглощения. В этом случае по мере удаления слоев загрязнений (интенсивно поглощающих свет лазера) лазерный пучок будет постепенно приближаться к материалу подложки (т.е. к поверхности памятника), который при правильном выборе длины волны излучения лазера

почти не поглощает его излучение. После полного удаления загрязнений процесс фотоабляции самопроизвольно прекращается, так как практически весь падающий свет будет отражаться от поверхности памятника.

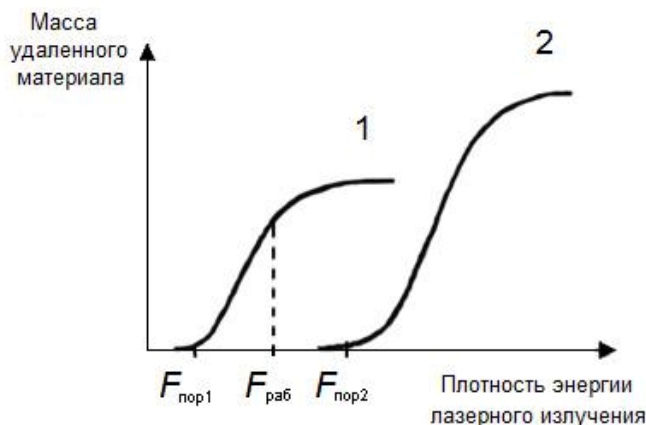


Рис. 2. Характерные зависимости степени абляции материала загрязнения (1) и материала памятника (2) от энергии лазера: $F_{\text{пор1}}$ – пороговый уровень фотоабляции вещества загрязнения, $F_{\text{пор2}}$ – пороговый уровень фотоабляции материала памятника, $F_{\text{раб}}$ – рабочий уровень плотности энергии лазера, который нужно поддерживать для селективного удаления загрязнений в процессе лазерной очистки памятника

Таким образом, при правильном выборе параметров лазера лазерная обработка является селективным и полностью управляемым процессом, что позволяет исключить риск повреждения поверхности памятника. Это выгодно отличает лазерную очистку от, например, химической обработки, при которой неконтролируемому воздействию химического реагента подвергается не только область обработки, но и вся прилегающая часть поверхности объекта (которая чаще всего не нуждается в очистке). В этом заключается принципиальное отличие лазерной очистки от традиционных методов реставрации – обработка поверхности памятника с помощью лазера является наиболее щадящей технологией, которая позволяет эффективно удалять даже самые стойкие загрязнения и при этом не только не повреждать поверхность объекта, но и сохранять оригинальную (историческую) патину.

К сказанному выше следует добавить, что для обеспечения лучшей сохранности объекта реставрации целесообразно применение так называемой влажной лазерной очистки [3]. В этом случае на поверхность объекта перед обработкой наносят тонкий слой воды (или спиртосодержащей жидкости), которая проникает внутрь слоя загрязнений. При быстром термическом расширении и абляции частиц этого слоя степень термического воздействия лазерного излучения на поверхность памятника оказывается меньше, чем при сухой очистке, поскольку в процессе адгезии, помимо сил Ван-дер-Ваальса, участвуют также капиллярные силы [7]. Кроме того, свой вклад вносит еще и энергия «термовзрыва», возникающего при вскипании молекул воды в порах загрязняющего вещества в результате поглощения им световой энергии лазерного излучения.

Экспериментальные результаты

Объекты исследования

Типичными материалами петербургских памятников из металлов являются медь, бронза и чугун, а для каменной скульптуры характерны мрамор, известняк и гранит. В качестве основного объекта исследования в данной работе использовались фрагменты так называемой выколотной медной скульптуры с центральной башни здания Главного Адмиралтейства и крыши Зимнего дворца в Санкт-Петербурге. Проводились также и

эксперименты по лазерной очистке памятников из мрамора, включая удаление с их поверхности микроскопических грибов и других биологических поражений, что составляло отдельную и очень важную часть наших исследований. Но эти вопросы заслуживают отдельного серьезного обсуждения и будут рассмотрены в более поздних публикациях авторов.

Подвергавшиеся лазерной обработке фрагменты медных памятников имели характерные для подобных объектов загрязнения и поражения поверхностного слоя. Под воздействием влажной атмосферы, солевых и серноокислых соединений в воздухе медь разрушается с образованием хлоридов, сульфатов, сульфидов, зеленых карбонатов, черной окиси и красной закиси меди. При этом цвет, состав и строение коррозионного слоя зависит от присутствия в воздухе различных газов, твердых частиц разных веществ, а также от состава самого медного сплава.

Экспериментальное оборудование

Эксперименты по лазерной очистке фрагментов медных памятников были выполнены при помощи специализированного реставрационного лазера Smart Clean II (производство компании ElEn Spa., Италия). Это импульсный Nd:YAG лазер, предназначенный для очистки поверхности исторических объектов из камня и металла. Основные характеристики его излучения таковы: длина волны $\lambda=1064$ нм, длительность импульса – 30–110 мкс, энергия – 0,2–2 Дж, частота повторения импульсов – 1–30 Гц. В качестве комментария заметим, что в блоке питания лазера (в цепи разряда лампы накачки) используется специальная электронная схема, которая позволяет обеспечить нетипичную для режима свободной генерации (очень малую) длительность импульса. Именно эта особенность позволяет использовать данный лазер как для очистки мрамора (где указанная длительность импульса является оптимальной [8]), так и исторических объектов из металлов [9].

С целью уменьшения риска возможных повреждений медной поверхности использовался метод влажной лазерной очистки. Работы по очистке всех экспериментальных образцов начинались с пробных расчисток отдельных небольших (размером около 2×2 см²) участков, причем при самых низких уровнях плотности энергии излучения. Затем плотность энергии постепенно увеличивалась для достижения требуемой степени очистки (решение о приемлемости результатов лазерной обработки принималось специалистами-реставраторами).

Обработка поверхности медных пластин производилась лазерным пучком диаметром около 2 мм. Доставка лазерного излучения в зону обработки и его сканирование по поверхности очищаемого предмета осуществлялись посредством оптоволоконного кабеля с ручным «фокусатором», входящего в состав лазерной установки (рис. 3). Рабочие уровни плотности энергии, при которых выполнялась очистка, составляли 10–25 Дж/см² (разброс значений связан с характером загрязнений на различных участках поверхности), а частота повторения импульсов – 5–10 Гц. При указанных выходных параметрах лазера с поверхности меди были успешно удалены не только очаги коррозии, но и самые стойкие загрязнения.

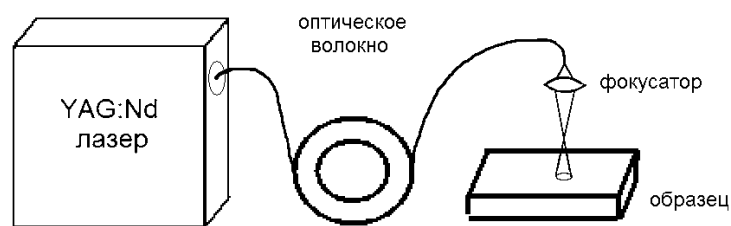


Рис. 3. Схема экспериментальной установки

Состояние поверхности образцов контролировалось при помощи оптического (модель МБС-10, Россия) и электронного сканирующего (модель АВТ-55, изготовитель – фирма Acashi, Япония) микроскопов. Анализ этих изображений (особенно показательны изображения, полученные на электронном микроскопе) показал, что на поверхности меди имеются следы микроплавления (см. рис. 4)). Однако такие микроповреждения были признаны реставраторами вполне приемлемыми. Дело в том, что традиционная механическая очистка экстерьерной выколотной скульптуры из меди осуществляется металлическими ручными и электромеханическими щетками, что зачастую приводит к появлению глубоких царапин. Кроме того, после очистки выколотная скульптура обычно грунтуется, а затем окрашивается. Поэтому отмеченные при лазерной обработке микродефекты поверхности никак не могут сказаться на экспозиционном виде и сохранности таких памятников.

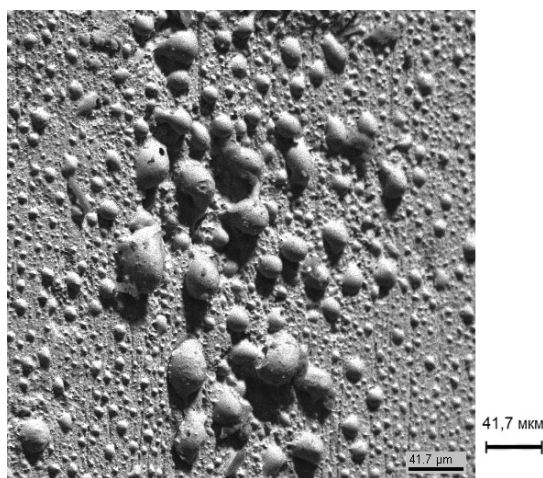
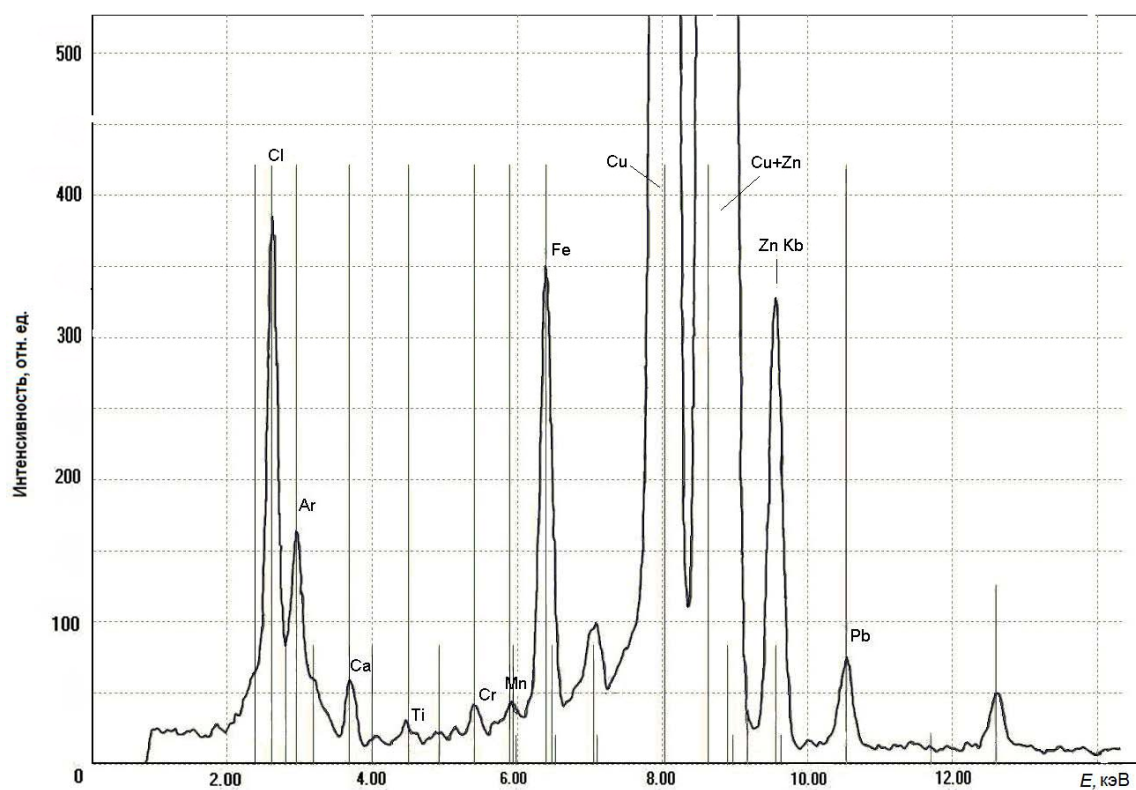
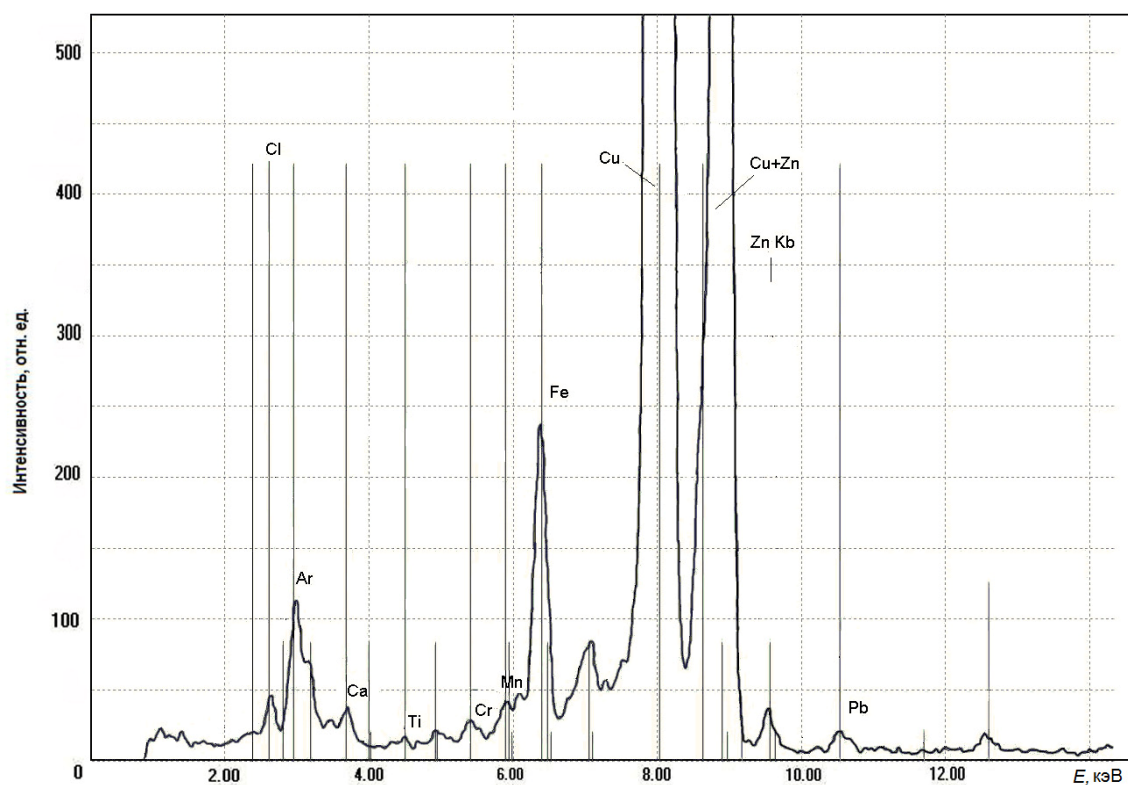


Рис. 4. Изображение поверхности медного памятника после лазерной обработки, полученное на электронном сканирующем микроскопе

Совсем другое дело – химическая чистота лазерной очистки. Рассмотрение этого вопроса было особенно актуально из-за того, что в отечественной реставрационной практике качество очистки памятников обычно оценивается весьма субъективно («на глаз»). С целью количественной оценки результатов работы с лазером были проведены специальные измерения с использованием рентген-флуоресцентного спектрометра (модель «Х-Арт М», разработчик – ЗАО «Комита», Россия). На рис. 5 приведены характерные рентгеновские спектры, полученные до и после завершения процесса очистки. Как видно из рисунка, после обработки лазером спектральные линии, характеризующие наличие хлора, кальция, титана, хрома, железа, цинка и свинца, либо исчезли полностью, либо стали значительно слабее, чем в исходном состоянии медной поверхности. Это свидетельствует о том, что в результате лазерной очистки с поверхности меди были удалены основные виды загрязнений антропогенного характера, которые могли быть результатом воздействия на медь химических примесей ливневых вод и выхлопных газов автомобильного транспорта.



а



б

Рис. 5. Спектры, полученные на рентген-флюоресцентном спектрометре для фрагмента медной скульптуры до (а) и после (б) лазерной очистки

Заключение

На основании проведенных исследований можно сделать вывод о том, что лазерная обработка, основанная на использовании эффекта фотоабляции, является эффективным способом очистки медной выколотной скульптуры от характерных природных наслоений и загрязнений, а специализированный реставрационный лазер Smart Clean II может быть рекомендован к использованию при решении подобных задач.

Авторы выражают благодарность В.И. Кудряшову за помощь в проведении измерений на спектроанализаторе «Х-Арт М». Благодарим также А.А. Борхварта (ЗАО «Акме-дек»), предоставившего для исследования фрагменты медных памятников, и К.А. Бурнякову, принимавшую участие в экспериментах.

Литература

1. Salimbeni R. Laser techniques in Conservation in Europe // SPIE Proceedings. – 2005. – V. 5857. – P. 8–18.
2. Парфенов В.А. Лазерные технологии реставрации и исследования произведений искусства // Исследования в консервации культурного наследия. Вып. 2. Материалы международной научно-методической конференции, посвященной 50-летию юбилею ГосНИИР. – М.: Индрик, 2008. – 320 с. – С. 217–226.
3. Cooper M. Laser Cleaning in Conservation: An Introduction, Butterworth-Heinemann. – Oxford, 1998.
4. Lazzarini L., Asmus J.F., Marchesini M.L. Lasers for the cleaning of statuary, initial results and potentialities // 1st Int. Symposium on the Deterioration of Building Stone. – La Rochelle, 1972. – P. 89–94.
5. Никитин М.К., Мельникова Е.П. Химия в реставрации. – Л.: Химия, ЛО, 1990. – 293 с.
6. Luk'yanchuk B.S. Laser Cleaning. – World Scientific Singapore, 2002.
7. Lu Y.F., Ren Z.M. Laser microprocessing and the applications in microelectronics industry // SPIE Proceedings. – 2000. – V. 4157. – P. 191–199.
8. Siano S., Giamello M., Bartoli L., Mencaglia A., Parfenov V., Salimbeni R. Phenomenological characterisation of stone cleaning by different laser pulse duration and wavelength, // J. Nimmrichter, W. Kautek, M. Schreiner (eds.). Laser in the conservation of artworks. – LACONA VI Proceedings, Vienna, Austria, Sept. 21–25, 2005. – Springer-Verlag, Berlin Heidelberg, 2007. – P. 87–96.
9. Siano S., Grazie F., Парфенов В.А. Оптимизированная лазерная очистка позолоченных бронзовых поверхностей // Оптический журнал. – 2008. – Т. 75. – № 7. – С. 18–29.

<i>Парфенов Вадим Александрович</i>	– Санкт-Петербургский государственный электротехнический университет, кандидат технических наук, доцент vadim_parfenov@mail.ru
<i>Геращенко Анастасия Николаевна</i>	– Санкт-Петербургский государственный электротехнический университет, аспирант, opinio@inbox.ru
<i>Геращенко Михаил Дмитриевич</i>	– Санкт-Петербургский государственный электротехнический университет, аспирант, sir_kofaioh@inbox.ru
<i>Григорьева Ирина Дмитриевна</i>	– Санкт-Петербургский государственный электротехнический университет, аспирант, grigoryevai@rambler.ru

УДК 544.77.032.1

**ИССЛЕДОВАНИЕ ПРОЦЕССА ОБРАЗОВАНИЯ СЛОИСТОСТИ
ПРИ ВЫЩЕЛАЧИВАНИИ НАТРИЙБОРСИЛИКАТНЫХ СТЕКОЛ****И.К. Мешковский, А.О. Олехнович**

Представлена модель образования слоистых неоднородностей при выщелачивании натрийборсиликатного стекла состава Na 7/23 [1]. Полученные выщелачиванием элементы могут использоваться в качестве оптических элементов.

Ключевые слова: выщелачивание, моделирование, слоистость, страты, пористость, натрийборсиликатное стекло, пористая матрица, двухфазное стекло.

Введение

Изучению слоистости пористого стекла, полученного выщелачиванием натрийборсиликатных двухфазных стекол, посвящено несколько работ [1–5] (рис. 1). Особенно актуальны эти работы в связи с применением пористого стекла для изготовления оптических элементов [6–9]. Причины образования упомянутой слоистости недостаточно изучены. Настоящая работа посвящена моделированию процессов, развивающихся при выщелачивании, которые могут быть ответственны за образование слоистых осадков в пористом стекле.

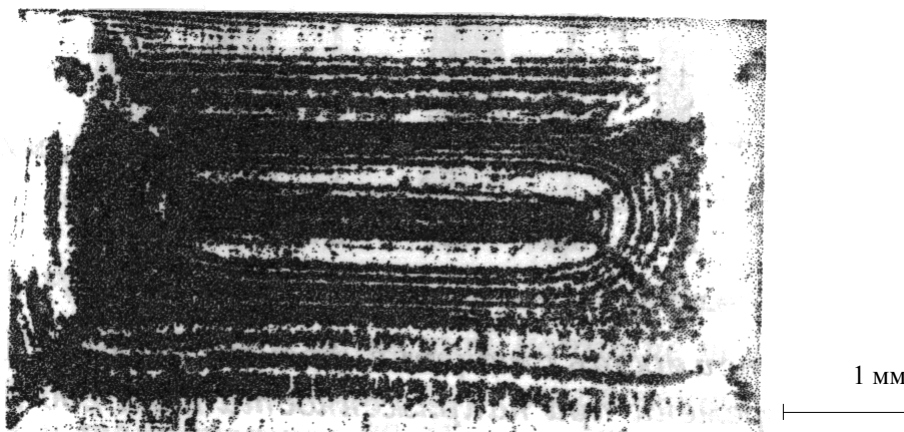


Рис. 1. Картина слоистых образований в пористом стекле. Изображение получено при фотографировании на просвет выщелоченного натрийборсиликатного образца из стекла ДВ-1М

Как известно, в натрийборсиликатном стекле, например состава Na 7/23 [1], прошедшем соответствующую термическую обработку, сосуществуют две фазы, пронизывающие одна другую: кремнеземистый каркас и растворимые в кислотах натрийборатные области, в которые входит некоторая доля кремнезема.

В процессе выщелачивания растворимая часть стекла постепенно переходит в раствор. Этот процесс характеризуется диффузионной кинетикой [10, 11]. При этом происходит постепенное углубление процесса растворения в поры образца, подвергнувшегося выщелачиванию. В работе впервые смоделирован механизм образования периодических неоднородностей преломления в пористых стеклах, полученных при выщелачивании натрийборсиликатных стекол ДВ-1М [12]. Актуальность работы связана с применением пористых стекол для изготовления оптических элементов, например активных элементов лазеров [13].

Описание модели

Механизм процесса может быть представлен следующим образом. По нашему предположению, процесс выщелачивания стекла начинается с адсорбции на поверхности ионов Cl^- и диффузии вглубь стекла протонов. Катионы водорода появляются в результате распада ионов гидроксония H_3O^+ , которые, как известно, присутствуют в водном растворе соляной кислоты. В контексте работы речь идет об ионах гидроксония, столкнувшихся с поверхностью образца и адсорбированных на ней. Протоны, оказавшиеся под поверхностью стекла, способны вносить существенные изменения в структуру стекла. Для обеспечения квазинейтральности системы на поверхность стекла притягиваются ионы Cl^- . Таким образом, в начальный момент на поверхности стекла формируется двойной электрический слой, который в дальнейшем сдвигается вместе с зоной раздела фаз. По нашим оценкам, напряженность поля в нем составляет от $5 \cdot 10^5$ до $1 \cdot 10^6$ В/см². Процесс формирования двойного электрического слоя связан с диффузией и накоплением протонов в приповерхностном слое стекла и соответствующим выстраиванием сольватированных ионов хлора непосредственно на поверхности стекла. Интересным экспериментальным фактом является задержка процесса выщелачивания от момента контакта раствора кислоты с поверхностью стеклянного образца. Эта задержка равна примерно 120 с. Именно в это время, по нашему мнению, формируется двойной электрический слой. За это время не наблюдается никаких изменений размеров образца стекла и не происходит изменение веса. После 120 с образцы начинают изменяться в размерах и весе.

По мере смещения процесса в толщу образца диффузионное сопротивление образующихся пор возрастает, и в зоне реакции начинается рост концентрации диффундирующих растворимых компонентов и снижение концентрации растворяющих компонентов. Кроме натрийборатных компонентов, в порах накапливается растворенный коллоидальный кремнезем, что приводит к торможению процесса растворения. Когда концентрация щелочных компонентов возрастает до определенного предела, происходит коагуляция коллоидного образования и выпадение на поверхность пор. Это приводит к резкому снижению диффузионного сопротивления, так как коллоидальный кремнезем занимал основное пространство пор, через которые происходит диффузия. Коагуляция коллоидного образования освобождает пространство пор, через которое происходит отток в раствор накопленных щелочных компонентов и приток из раствора кислоты. Процесс выщелачивания вновь интенсифицируется, и цикл повторяется вновь.

Расчетная модель

Представленный механизм может быть смоделирован следующим образом. Будем исследовать процесс, протекающий при контакте двухкомпонентной твердой фазы (А) и раствора (Б) (рис. 2). Месторасположение границы твердой и жидкой фаз перемещается вправо по мере растворения (А), перемещение границ описывается функцией $x=l(t)$. Концентрацию компонент твердой фазы (А) в растворе (Б) мы будем обозначать $U(x,t)$ и $V(x,t)$.

Предположим, что после попадания в раствор происходит диффузия компонентов, при этом компонент $V(x,t)$ выпадает в осадок со скоростью, пропорциональной его концентрации. Исходя из этих предположений, уравнения описывающие процесс диффузии могут быть записаны так:

$$\frac{\partial U}{\partial t} = D_{11} \frac{\partial^2 U}{\partial x^2} + D_{12} \frac{\partial^2 V}{\partial x^2}, \quad (1)$$

$$\frac{\partial V}{\partial t} = D_{21} \frac{\partial^2 U}{\partial x^2} + D_{22} \frac{\partial^2 V}{\partial x^2} - kV. \quad (2)$$

Здесь члены, содержащие коэффициенты диффузии D_{ik} – стандартные члены, описывающие двухкомпонентную диффузию (см., например, [14]), а член kV определяет процесс выпадения осадка.

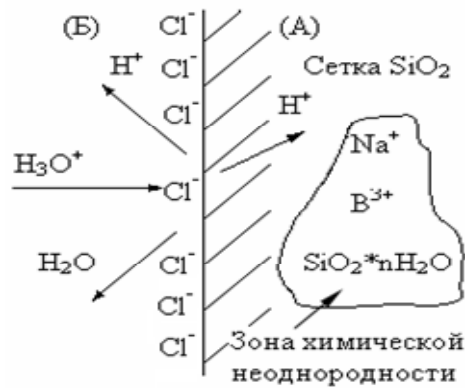


Рис. 2. Схема процесса распада ионов гидроксония при адсорбции на поверхности и образования потока протонов под поверхностью стекла

Граница фаз достигнет точки с координатой X в момент времени t , удовлетворяющий уравнению

$$x=l(t). \quad (3)$$

В этой точке выпадает осадок

$$N(x) = k \int_{\tau(x)}^{\infty} V(x,t) dt, \quad (4)$$

где через $\tau(x)$ мы обозначили функцию, обратную к (3), т.е. t – это момент времени, в который начнется выпадение осадка в токе с координатой x , $N(x)$ – приведенная концентрация осадка (если $n(x)$ – поверхностная концентрация осадка, то $SN(x)=n(x)$, где S – площадь поперечного сечения канала (А)).

Решение уравнения в общем виде позволило получить нам следующие результаты. Общее решение уравнения для случая

$$\Delta = D_{12}D_{21} - D_{11}D_{22} > 0 \quad (5)$$

имеет для пространственных координат вид

$$N = N_0 \cos(\omega x + \varphi_0) + N_1(x), \quad (6)$$

где N_0 и φ_0 – произвольные постоянные, $\omega = \sqrt{\frac{kD_{11}}{\Delta}}$, $N_1(x)$ – частное решение уравнения (4). Не вдаваясь в механизм возникновения колебаний (для этого было бы необходимо составлять краевые условия, исследовать вид $V[x, \tau(x)]$ и $\frac{\partial V[x, \tau(x)]}{\partial x}$ и т.д.), мы

можем утверждать, что в случае (5) происходит периодическое по пространственной переменной выпадение осадка с пространственным периодом

$$T = \frac{2\pi}{\omega} = 2\pi \sqrt{\frac{\Delta}{kD_{11}}}. \quad (7)$$

В важном для нас случае, когда $D_{12}D_{21} \gg D_{11}D_{22}$ (скажем, если $D_{22}=0$) и $D_{12} \sim D_{21} \sim D_0$,

$$T \sim 2\pi \frac{D_0}{\sqrt{kD_{11}}} . \quad (8)$$

Отметим, что формула (11) имеет смысл лишь при $N > 0$, что возможно, если

$$N_1 > |N_0|. \quad (9)$$

В случае $\tau = \frac{1}{c}x$ (граница движется с постоянной скоростью c) и $V[x, \tau(x)] = a_1$, $U[x, \tau(x)] = a_2$, $\frac{dV[x, \tau(x)]}{dx} = \epsilon$, где a и ϵ – константы (такое предположение вполне допустимо, так как эти величины изменяются достаточно медленно; впрочем, его можно было бы и не делать, это существенно усложнило бы дальнейшие выкладки, но привело бы к полностью аналогичным результатам),

$$f(x) = a_1 - \frac{D_{21}}{D_{11}} - \frac{\Delta a}{c} = const ,$$

и частное решение $N_1 = f$. При этом значительные по амплитуде колебания осадка будут иметь место лишь в случае

$$D_{11} \gg \frac{D_{21}a_2 + \Delta a c^{-1}}{a_1} . \quad (10)$$

Резюмируя рассуждения, относящиеся к формулам (5) и (10), мы можем сформулировать условия, при которых процесс выпадения осадка носит выраженный периодический характер:

- превалирует «перекрестная диффузия» (в (10) член $D_{12}D_{21}$ превышает $D_{11}D_{22}$);
- D_{11} – достаточно большая величина (в смысле оценки (10)).

Из этих двух условий следует, что для одновременного выполнения (5) и (10) коэффициент D_{22} должен быть мал.

При выщелачивании пористого стекла роль компонента U могут играть щелочные компоненты, а роль V – растворимая часть SiO_2 . При этом процесс диффузии SiO_2 практически целиком определяется сопутствующей диффузией щелочных компонентов, и коэффициент D_{22} пренебрежимо мал.

Отметим, что в широком смысле рассмотренный процесс описывает самоорганизацию открытой диффузной системы (открытость связана с наличием подвижной границы фаз).

Заключение

В статье представлена модель, дающая возможность объяснения появления периодического изменения плотности при выщелачивании двухфазного натрийборсиликатного стекла. Приведен математический аппарат, используемый для описания процесса выщелачивания, и определены условия, при которых выпадение осадка носит ярко выраженный периодический характер.

Литература

1. Молчанова О.С. Натриевоборсиликатные и пористые стекла. – М.: Оборонгиз, 1961, 162 с.
2. Жданов С.П. Структура пористых стекол по адсорбционным данным // Труды ГОИ. – 1956. – Т. 24. – Вып. 145. – С. 86–114.

3. Добычин Д.П. Проведение теплехимического диффузионного процесса с постоянной скоростью // ДАН СССР. – 1958. – Вып. 119. – С. 967–970.
4. Альтшулер Г.Б., Баханов В.А., Дульнева Е.Г., Мешковский И.К. Исследование оптических характеристик активных элементов из кварцевого микропористого стекла // Оптика и спектроскопия. – 1983. – Т. 55. – Вып. 2. – С. 369–374.
5. Сиренек В.А., Антропова Т.В., Кутчиев А.И., Чирков А.В., Анфимова И.Н. Расчет кинетики взаимодействия щелочно-боросиликатных стекол с водными растворами // Физика и химия стекла. – 2004. – Т. 30. – Вып. 1. – С. 24–35.
6. Дульнев Г.Н., Земский В.И., Крынецкий Б.Б., Мешковский И.К., Прохоров А.М., Стельмах О.М. Твердотельный перестраиваемый лазер на микрокомпозиционном матричном материале // Изв. АН СССР. – 1979. – Сер. физ. – Т. 43. – Вып. 2. – С. 237–238.
7. Альтшулер Г.Б., Дульнева Е.Г., Крылов К.И., Мешковский И.К., Урбанович С.С. Генерационные характеристики лазера на родамине 6Ж в микропористом стекле // ЖПС АН БССР. – 1982. – Т. 36. – Вып. 4. – С. 592–599.
8. Карпов Н.В., Мешковский И.К., Петров Ю.Н., Петров Р.С., Прохоров А.М. Лазерное управление проницаемостью молекулярного сита // Письма в ЖЭТФ. – 1979. – Т. 30. – вып. 1. – С. 48–52.
9. Алиев Ф.М., Баушев В.Н., Дульнев Г.Н., Мешковский И.К., Прохоров А.М. Опто-электронная ячейка на термически управляемом рассеянии света в гетерогенной системе «пористое стекло-жидкий кристалл» // ДАН СССР. – 1980. – Т. 253. – Вып. 3. – С. 598–600.
10. Жданов С.П. Химическая устойчивость щелочносиликатных стекол и ее связь с координатией катионов. Вакансионный механизм выщелачивания // Физика и химия стекла. – 1978. – Т. 4. – Вып. 5. – С. 505–514.
11. Мешковский И.К. О кинетике выщелачивания ликвидированных стекол // ЖПХ АН СССР. – 1984. – Вып. 1. – С. 45–49.
12. ОСТ 3-1899-81.
13. Земский В.И., Колесников Ю.Л., Мешковский И.К. Физика и техника импульсных лазеров на красителях. – СПб: СПб ГУИТМО, 2005. – 176 с.
14. Карслоу Г., Егер Д. Теплопроводность твердых тел. – М.: Наука, 1964. – 488 с.

Мешковский Игорь Касьянович – Санкт-Петербургский государственный электротехнический университет, зав. кафедрой, доктор технических наук, профессор, igorkm@runnet.ru

Олехнович Артем Олегович – Санкт-Петербургский государственный электротехнический университет, ассистент, romanart1@inbox.ru

УДК 535.1

ЗАКОНОМЕРНОСТИ ВЗАИМОДЕЙСТВИЯ ВСТРЕЧНЫХ СВЕТОВЫХ ВОЛН ИЗ МАЛОГО ЧИСЛА КОЛЕБАНИЙ В НЕЛИНЕЙНЫХ ДИЭЛЕКТРИЧЕСКИХ СРЕДАХ

Е.М. Буяновская, С.А. Козлов

Теоретически рассмотрены закономерности взаимодействия встречных световых волн из малого числа колебаний в нелинейных диэлектрических средах. Получено аналитическое решение уравнений динамики поля встречных световых волн из малого числа колебаний, взаимодействующих в диэлектрической среде с безынерционной кубичной по полю нелинейностью. Приведены изменения временной и спектральной структуры светового импульса при его взаимодействии со встречным импульсом. Показано, что эффективность этих явлений определяется энергией встречного импульса и не зависит от его длительности и спектральной структуры.

Ключевые слова: импульсы из малого числа колебаний, взаимодействие встречных волн, нелинейные диэлектрические среды.

Введение

Оптика импульсов из малого числа колебаний светового поля интересна как с фундаментальной, так и с практической точки зрения. Для таких импульсов теряет свое физическое содержание понятие огибающей, поэтому при теоретическом изучении особенностей их распространения в различных оптических средах обычно анализируют динамику непосредственно поля излучения [1–3].

К настоящему времени изучены многие явления нелинейной оптики таких предельно коротких (по числу колебаний) импульсов: их временное и спектральное уширение и сжатие, самофокусировка, нелинейное отражение, взаимодействие при попутном распространении [4–10]. В работе [11], по-видимому, впервые были выведены уравнения динамики поля световых импульсов из малого числа колебаний при их встречном распространении в нелинейной среде. В настоящей работе приведены аналитические интегральные решения этих уравнений. На их основе рассмотрены основные закономерности взаимодействия в нелинейных диэлектрических средах встречных оптических импульсов, содержащих лишь несколько колебаний светового поля.

Уравнения динамики поля плоских встречных световых волн из малого числа колебаний и их решения

В работе [11] нами были выведены уравнения, описывающие динамику полей встречных плоских световых волн из малого числа колебаний при их взаимодействии в диэлектрических средах с безынерционной кубической нелинейностью вида

$$\begin{cases} \frac{\partial E_+}{\partial z} + \frac{N_0}{c} \frac{\partial E_+}{\partial t} - a \frac{\partial^3 E_+}{\partial t^3} + b \int_{-\infty}^t E_+ dt' + \frac{cg}{2N_0} \frac{\partial}{\partial t} (E_+^3 + 3E_+^2 E_- + 3E_+ E_-^2) = 0, \\ \frac{\partial E_-}{\partial z} - \frac{N_0}{c} \frac{\partial E_-}{\partial t} + a \frac{\partial^3 E_-}{\partial t^3} - b \int_{-\infty}^t E_- dt' - \frac{cg}{2N_0} \frac{\partial}{\partial t} (E_-^3 + 3E_-^2 E_+ + 3E_- E_+^2) = 0, \end{cases} \quad (1)$$

где $E_+(z, t)$ – поле волны, распространяющейся в положительном направлении оси z , $E_-(z, t)$ – поле волны, распространяющейся ей навстречу; t – время, c – скорость света в

вакууме; N_0, a, b – параметры, характеризующие типичную нерезонансную зависимость показателя преломления диэлектрической среды в диапазоне ее прозрачности [12],

$$n = N_0 + ca\omega^2 - c \frac{b}{\omega^2}, \quad (2)$$

от частоты ω , $g = \frac{4\pi\chi}{c^2}$ описывает нелинейность ее поляризационного отклика

$P_{\text{нл}} = \chi E^3$ [1, 9], χ – нелинейная восприимчивость среды.

Для получения решений уравнений (1) их удобно нормировать, вводя новые безразмерные переменные $E' = \frac{E}{E_{+0}}$, $z' = \frac{z}{\lambda_{+c}}$, $t' = \frac{t}{T_{+c}}$, где E_{+0} – максимальное значение на границе нелинейной среды поля излучения, например, прямой волны E_+ ; T_{+c} – ее центральный период колебаний на той же границе, $\lambda_{+c} = \frac{cT_{+c}}{N_0}$ – центральная длина волны.

В новых переменных уравнения (1) запишутся как

$$\begin{cases} \frac{\partial E_+}{\partial z} + \frac{\partial E_+}{\partial t} - A \frac{\partial^3 E_+}{\partial t^3} + B \int_{-\infty}^{\infty} E_+ dt' + G \frac{\partial}{\partial t} (E_+^3 + 3E_+^2 E_- + 3E_+ E_-^2) = 0, \\ \frac{\partial E_-}{\partial z} - \frac{\partial E_-}{\partial t} + A \frac{\partial^3 E_-}{\partial t^3} - B \int_{-\infty}^{\infty} E_- dt' - G \frac{\partial}{\partial t} (E_-^3 + 3E_-^2 E_+ + 3E_- E_+^2) = 0, \end{cases} \quad (3)$$

где безразмерные коэффициенты $A = \frac{ac}{N_0 T_{+c}^2}$, $B = \frac{bc T_{+c}^2}{N_0}$, $G = \frac{c^2 g E_{+0}^2}{2N_0^2} = \frac{2\pi\chi E_{+0}^2}{N_0^2}$. В уравнениях (3) и ниже значок «'» для краткости не пишем.

Хотя интенсивность встречных световых волн, при которой не будет наблюдаться разрушение вещества, из-за их предельно коротких длительностей может быть весьма велика [13], но результаты их столкновения все-таки очень сильными быть не могут из-за скоротечности взаимодействия. Поэтому естественным методом решения уравнений (3) является метод последовательных приближений Пикара [14], в котором малым параметром является G .

Далее для простоты, но без ограничения общности, будем рассматривать среду без дисперсии линейного показателя преломления. Во-первых, это приближение выполняется при $G \gg A, B$, а, во-вторых, из-за малости нерезонансной дисперсии показателя преломления диэлектриков (2) ее, при необходимости, в первом приближении не сложно учитывать, включая в итерационное решение уравнений (3) аддитивно к нелинейным слагаемым.

В соответствии с техникой последовательных приближений [14] будем искать решение уравнений (3) в виде

$$E_+ = E_+^{(0)} + G E_+^{(1)} + G^2 E_+^{(2)} + \dots, \quad (4a)$$

$$E_- = E_-^{(0)} + G E_-^{(1)} + G^2 E_-^{(2)} + \dots, \quad (4b)$$

в котором в данной работе и в (4a), и в (4b) ограничимся первыми двумя слагаемыми. Тогда в представлении (4) система уравнений (3) при $A = B = 0$ примет вид

$$\begin{cases} \frac{\partial E_+^{(0)}}{\partial z} + \frac{\partial E_+^{(0)}}{\partial t} = 0, \\ \frac{\partial E_-^{(0)}}{\partial z} - \frac{\partial E_-^{(0)}}{\partial t} = 0, \end{cases} \quad (5)$$

$$\begin{cases} \frac{\partial E_+^{(1)}}{\partial z} + \frac{\partial E_+^{(1)}}{\partial t} + \frac{\partial}{\partial t} \left(E_+^{(0)3} + 3E_+^{(0)2} E_-^{(0)} + 3E_+^{(0)} E_-^{(0)2} \right) \\ \frac{\partial E_-^{(1)}}{\partial z} - \frac{\partial E_-^{(1)}}{\partial t} - \frac{\partial}{\partial t} \left(E_-^{(0)3} + 3E_-^{(0)2} E_+^{(0)} + 3E_-^{(0)} E_+^{(0)2} \right) \end{cases} \quad (6a)$$

Решение системы (5) имеет вид [11, 15]:

$$\begin{cases} E_+^{(0)}(z, t) = E_+^{(0)}(t - z), \\ E_-^{(0)}(z, t) = E_-^{(0)}(t + z) \end{cases} \quad (7)$$

и определяется граничными условиями.

Каждое из уравнений системы (6) также несложно решить в квадратурах. Например, первое уравнение этой системы (6a) можно, как это сделано в [11], переписать в новых переменных $z' = z$, $\tau = t - z$. Тогда его решение примет вид:

$$\begin{aligned} E_+^{(1)}(z', \tau) = & - \int_{z'_0}^{z'} \frac{\partial}{\partial \tau} \left(E_+^{(0)}(\tau) \right)^3 dz'' - 3 \int_{z'_0}^{z'} \frac{\partial}{\partial \tau} \left(\left(E_+^{(0)}(\tau) \right)^2 E_-^{(0)}(\tau + 2z'') \right) dz'' - \\ & - 3 \int_{z'_0}^{z'} \frac{\partial}{\partial \tau} \left(E_+^{(0)}(\tau) \left(E_-^{(0)}(\tau + 2z'') \right)^2 \right) dz'', \end{aligned} \quad (8)$$

где первое слагаемое в правой части соотношения характеризует самовоздействие светового импульса, распространяющегося от границы нелинейной среды z'_0 в положительном направлении оси z' , а второе и третье – взаимодействие встречных импульсов в нелинейной среде. Отметим, что соотношение (8) содержит как частный случай решение задачи о столкновении высоко- и низкоинтенсивной ($E_-^{(0)} \gg E_+^{(0)}$) встречных волн, рассмотренной в [11].

Решение (8) можно привести к более удобному для дальнейшего анализа виду:

$$\begin{aligned} E_+^{(1)}(z', \tau) = & - \left[\frac{\partial}{\partial \tau} \left(E_+^{(0)}(\tau) \right)^3 \right] (z' - z'_0) - 3 \frac{\partial}{\partial \tau} \left[\left(E_+^{(0)}(\tau) \right)^2 \int_{z'_0}^{z'} E_-^{(0)}(\tau + 2z'') dz'' \right] - \\ & - 3 \frac{\partial}{\partial \tau} \left[\left(E_+^{(0)}(\tau) \right) \int_{z'_0}^{z'} \left(E_-^{(0)}(\tau + 2z'') \right)^2 dz'' \right]. \end{aligned} \quad (9)$$

Несложно показать, что аналогичным (8) и (9) будет выглядеть и решение уравнения (6б), но его следует решать в переменных $z' = z$, $\xi = t + z$.

Закономерности взаимодействия плоских встречных световых волн из малого числа колебаний

На рис. 1 приведена иллюстрация решения (7), которое получено в первой итерации метода Пикара для импульса вида:

$$E_+^{(0)}(\tau) = \exp\left(-\frac{\tau^2}{\tau_+^2}\right) \sin(2\pi\tau), \quad (10)$$

и встречной ему волны вида

$$E_-^{(0)}(z', \tau) = \gamma \exp\left(-\frac{(\tau + 2z')^2}{\tau_-^2}\right) \sin\left(2\pi\delta(\tau + 2z')\right) \quad (11)$$

где $\gamma = \frac{E_{-0}}{E_{+0}}$, E_{+0} и E_{-0} – исходные амплитуды взаимодействующих импульсов; $\delta = \frac{T_{+c}}{T_{-c}}$,

T_{+c}, T_{-c} – их центральные периоды колебаний, $\tau_+ = \frac{\tau_{+0}}{T_{+c}}$, $\tau_- = \frac{\tau_{-0}}{T_{-c}}$, τ_{+0}, τ_{-0} – исходные

длительности импульсов.

Распространение встречных волн (10) и (11) представлено в системе координат z' и τ , сопровождающей импульс (10), т.е. в той же системе, в которой получено и решение (9) в следующей второй итерации. Динамика полей встречных импульсов дана для случая $\gamma = 1$, $\delta = 2$, $\tau_+ = 0.5$, $\tau_- = 1$. Из рис. 1 видно, что форма импульсов, полученная при расчете в первой итерации, при их распространении и столкновении не изменяется. Оптическая среда, как следует из уравнений (5), в этом приближении для встречных волн линейна.

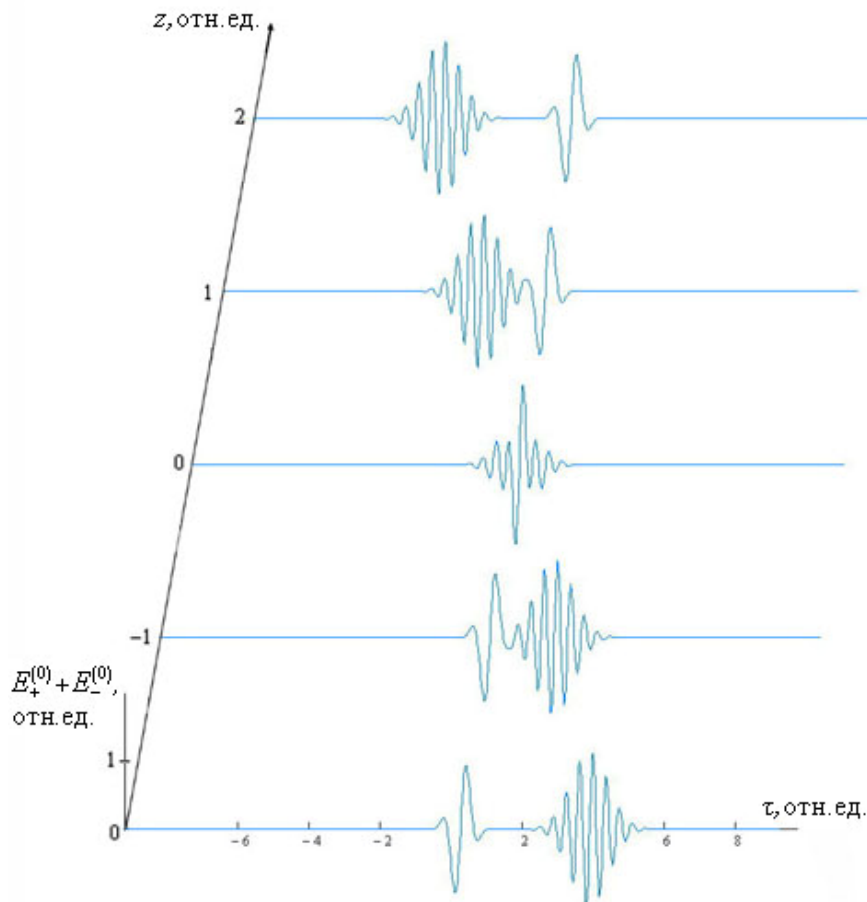


Рис. 1. Распространение встречных световых импульсов из малого числа колебаний в линейной оптической среде

На рис. 2, 3 приведены поправки (9) к решению (7), которые получены во второй итерации, учитывающей нелинейность среды. На них также дано общее решение (4а) исходной системы (3), описывающее самовоздействие импульса (10) и его взаимодействие с волной (11) в нелинейной среде.

На рис. 2 проиллюстрированы результаты самовоздействия импульса (10), описываемые первым слагаемым в правой части выведенного выражения (9). На рис. 2, а, представлена рассчитанная во второй итерации поправка к полю $E_+^{(1)}$, а на рис. 2, б – к

модулю спектра $g_+^{(1)}(\omega) = \int_{-\infty}^{\infty} E_+^{(1)}(\tau) e^{i\omega\tau} d\tau$ при $G(z' - z'_0) = \frac{1}{3}$. Как видно из рис. 2, а, и 2,

б, решение (9) описывает уширение спектра импульса и генерацию излучения на утроенных частотах. Отметим заметное смещение максимума уширяемого спектра импульса и излучения на утроенных частотах в высокочастотную область.

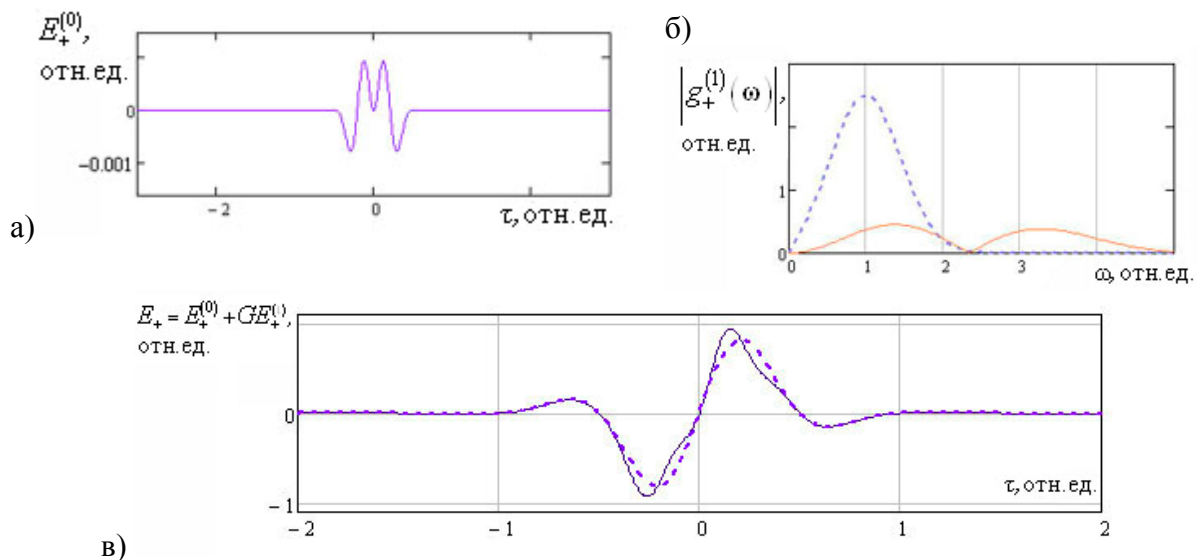


Рис. 2. Изменение временной структуры светового импульса из малого числа колебаний из-за его самовоздействия в нелинейной оптической среде: итерационная добавка к полю $E_+^{(1)}$ (а) и спектру $g_+^{(1)}$ (б), результирующее поле (в). Пунктиром показана временная структура поля до самовоздействия импульса

На рис. 2, в, приведена сумма $E_+^{(0)}$ и первого слагаемого в выражении $GE_+^{(1)}$, которая описывает результат самовоздействия импульса (10) в нелинейной среде, при $G(z' - z'_0) = \frac{1}{3}$. Пунктиром даны формы поля исходного импульса. Из рисунка видно, что основные максимум и минимум «однопериодного» светового импульса из-за самовоздействия излучения в нелинейной среде во времени τ начинают запаздывать. При сохранении положения нулей поля это приводит к искажению временной структуры импульса и описанному выше изменению его спектра.

На рис. 3 проиллюстрированы результаты воздействия на импульс (10) встречной волны (11), описываемые вторым и третьим слагаемым в правой части выражения (9). На рис. 3, а, приведена добавка к полю импульса $E_+^{(1)}$, полученная в результате взаимодействия. На рис. 3, б, представлен ее спектр $g_+^{(1)}$. На рис. 3, в, приведена временная структура электрического поля $E_+ = E_+^{(0)} + GE_+^{(1)}$ импульса до и после воздействия на него встречного импульса.

Из рис. 3 видно, что временная структура поля импульса из-за взаимодействия со встречным импульсом изменяется так, что больше всего смещаются во времени «нули» поля и не смещаются его экстремумы. «Центр тяжести» спектральной плотности при этом смещается в коротковолновую область. Важно, что эффективность этих явлений определяется энергией встречного импульса и не зависит от его длительности и спектральной структуры. Последнее утверждение несложно обосновать, анализируя выра-

жение (9). При $z' \geq 1$ (см. рис. 1), т.е. по окончании взаимодействия встречных импульсов, интеграл во втором слагаемом соотношения (9) для оптических импульсов становится равным нулю [1, 11], а интеграл в третьем слагаемом принимает смысл энергии встречного импульса. Генерации комбинационных частот в таких условиях не происходит. Но при этом отметим, что, если вторая граница нелинейной среды окажется при $z \leq 1$, т.е., когда взаимодействие еще не заканчивается, то генерация на комбинационных и кратных частотах из-за взаимодействия встречных импульсов возможна [11].

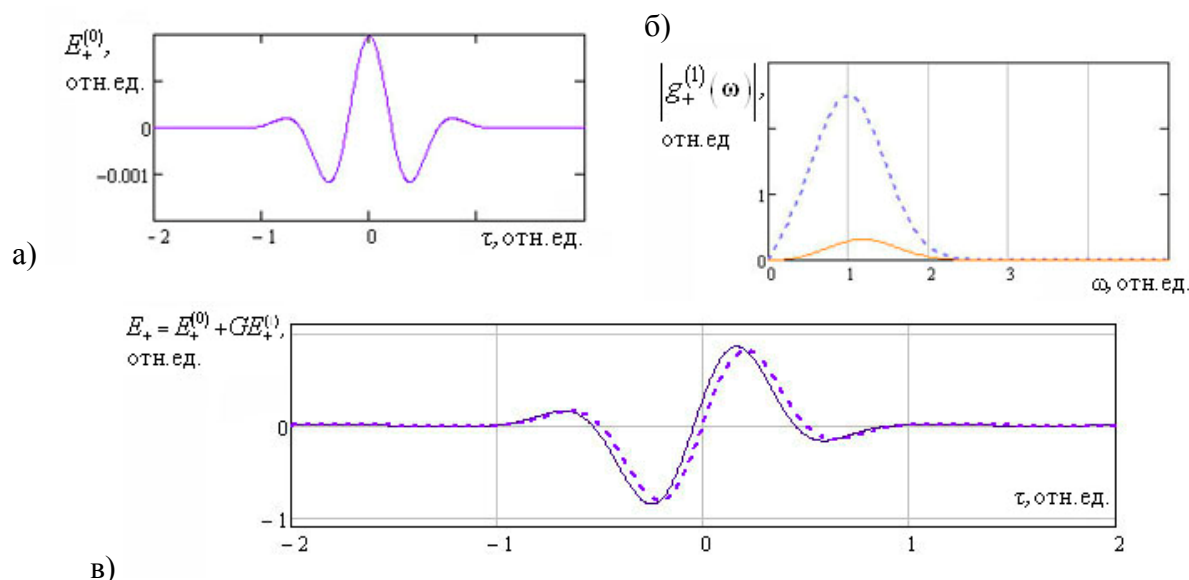


Рис. 3. Изменение временной структуры светового импульса из малого числа колебаний из-за воздействия встречного импульса в нелинейной оптической среде: итерационная добавка к полю $E_+^{(1)}$ (а) и спектру $g_+^{(1)}$ (б), результирующее поле (в).

Пунктиром показана временная структура поля до взаимодействия импульсов

Заключение

В работе теоретически исследованы закономерности взаимодействия встречных световых волн из малого числа колебаний в нелинейных диэлектрических средах. В результате проведенных исследований получено аналитическое решение уравнений динамики поля встречных световых волн из малого числа колебаний, взаимодействующих в диэлектрической среде с безынерционной кубичной по полю нелинейностью. Показано, что временная структура поля импульса из-за взаимодействия со встречным импульсом изменяется так, что больше всего смещаются во времени «нули» поля и не смещаются его экстремумы. «Центр тяжести» спектральной плотности при этом смещается в коротковолновую область, генерации комбинационных частот в таких условиях не происходит. Эффективность этих явлений определяется энергией встречного импульса и не зависит от его длительности и спектральной структуры.

Работа поддержана грантами РНП 2.1.1/4923 и РФФИ 08-02-00902-а.

Литература

1. Козлов С.А., Сазонов С.В. Нелинейное распространение импульсов длительностью в несколько колебаний светового поля в диэлектрических средах // ЖЭТФ. – 1997. – Т. 111. – В. 2. – С. 404–418.

2. Nazarkin A., Korn G. Raman self-conversion of femtosecond laser pulses and generation of single-cycle radiation // *Phys. Rev. A*. –1998. – V. 58. – № 1. –P. R61–R64.
3. Маймистов А.И. Некоторые модели распространения предельно коротких электромагнитных импульсов в нелинейной среде // *Квантовая электроника*. – 2000. – Т. 30. – № 4. – С. 287–304.
4. Казанцева Е.В., Маймистов А.И. Распространение предельно коротких импульсов в нерезонансной квадратично-нелинейной среде // *Квантовая электроника*. – 2000. – Т. 30. – № 7. – С. 623–628.
5. Shpolyanskiy Y.A., Belov D.L., Bakhtin M.A., Kozlov S.A. Analytic study of continuum spectrum pulse dynamics in optical waveguides // *Applied Physics B*. – 2003. – V. 77. – № 2–3. – P. 349–356.
6. Сазонов С.В., Халыпин В.А. О влиянии дифракции на нелинейное распространение оптических импульсов длительностью в несколько колебаний // *Квантовая электроника*. – 2004. – Т. 34. – № 11. – С.1057–1063.
7. Ястребова Н.В., Шполянский Ю.А., Козлов С.А. Нелинейное отражение импульсов из малого числа колебаний светового поля от просветленной границы раздела сред // *Оптический журнал*. – 2004. – Т. 71. – № 6. – С. 78–83.
8. Белов Д.Л., Козлов С.А., Шполянский Ю.А. О самосжатии спектрального суперконтинуума // *Известия РАН, серия физическая*. – 2005. – Т. 69. – № 8. – С. 1128–1130.
9. Berkovsky A.N., Kozlov S.A., Shpolyanskiy Y.A. Self-focusing of few cycle light pulses in dielectric media // *Physical. Review A* 72. – 2005. – 043821 (9 pages).
10. Бахтин М.А., Козлов С.А. Формирование последовательности сверхкоротких сигналов при столкновении импульсов из малого числа колебаний светового поля в нелинейных оптических средах // *Оптика и спектроскопия*. – 2005. – Т. 98. – № 3. – С. 425–430.
11. Буяновская Е.М., Козлов С.А. Динамика полей встречных световых импульсов из малого числа колебаний в нелинейных диэлектрических средах // *Письма в ЖЭТФ*. – 2007. – Т. 86. – В. 5–6. – С. 349–353.
12. Борн М., Вольф Э. Основы оптики. – М.: Наука, 1973. – 720 с.
13. Brabec Th., Krausz F. Intense few-cycle laser fields: Frontiers of nonlinear optics // *Rev. Mod. Phys.* – 2000. – V. 72. – № 2. – P. 545–591.
14. Корн Г., Корн Т, Справочник по математике. – М.: Наука, 1977. – 831 с.
15. Буяновская Е.М, Козлов С.А. Взаимодействие встречных световых импульсов из малого числа колебаний в нелинейных диэлектрических средах // *Научно-технический вестник СПбГУ ИТМО*. – 2006. – № 30. – С. 97–101.
16. Фарлоу С. Уравнения с частными производными для научных работников и инженеров. – М.: Мир, 1985. – 384 с.

Буяновская Елизавета Михайловна	– Санкт-Петербургский государственный университет информационных технологий, механики и оптики, аспирант, lee.buyanovskaya@gmail.com
Козлов Сергей Аркадьевич	– Санкт-Петербургский государственный университет информационных технологий, механики и оптики, доктор физ.-мат. наук, профессор, декан

3

АНАЛИЗ И СИНТЕЗ СЛОЖНЫХ СИСТЕМ.
ЭЛЕМЕНТЫ И УСТРОЙСТВА ТЕХНИЧЕСКИХ
СИСТЕМ

УДК 62.50

АЛГЕБРАИЧЕСКАЯ ОРГАНИЗАЦИЯ УСЛОВИЙ ОБОБЩЕННОЙ
СИНХРОНИЗИРУЕМОСТИ МНОГОАГРЕГАТНЫХ ДИНАМИЧЕСКИХ
ОБЪЕКТОВ

Н.А. Дударенко, М.В. Полякова, А.В. Ушаков

Рассматриваются проблемы алгебраической организации условий обобщенной синхронизируемости многоагрегатных динамических объектов, которые формируются на основе геометрической постановки задачи синхронного движения агрегатов как движения в линейной оболочке, натянутой на задающий вектор.

Ключевые слова: алгебраическая организация, многоагрегатные объекты, обобщенное синхронное движение, собственный вектор, линейная оболочка.

Введение

В современных технологических процессах по обработке материальных и обслуживанию гуманитарных потоков можно выделить задачи, в которых многоагрегатные компоненты технологических ресурсов участвуют в формировании совокупного финального продукта (результата), функционируя как «единое целое». Характерными примерами [1] таких технологических процессов по обработке материальных потоков являются процессы формирования и подачи ленточного материала в листопрокатном производстве, в производстве бумаги и тканей, в организации заготовительных процессов в составе «бесскладовых» технологических производств, процессы согласованного движения автономных больших полноповоротных радиотелескопов в составе радиоинтерферометра типа «Квазар» и т.д. Примерами технологических процессов по обслуживанию гуманитарных потоков являются процессы движения строем подвижных технических средств, управляемых антропокомпонентами-операторами, в виде колонны автомобилей, строя самолетов и вертолетов, и т.п., и автономных антропокомпонентов в виде строя военнослужащих на марше, в забеге и т.п.

Основным требованием к системам управления, встраиваемых в техническую среду отмеченных выше технологических процессов, является обеспечение синхронности функционирования динамических агрегатов в составе многоагрегатного процесса. При этом условие обеспечения синхронного функционирования агрегатов системно не должно быть доминирующим, чтобы не накладывались сильные ограничения на возможность назначения динамики технологического процесса с желаемыми показателями.

Исследования показали, что при геометрической постановке задачи управления многоагрегатными комплексами, доставляющего им синхронное функционирование, она обнаруживает несколько алгебраических формулировок. Так, в первой формулировке в рамках системной парадигмы вырождения сложных динамических систем [2, 3] условие синхронного функционирования агрегатов технологических ресурсов можно трактовать как вырождение многоагрегатной системы, при котором эллипсоид отношения «вход–выход» системы вырождается в отрезок прямой. В рамках второй формулировки задача сводится к алгебраическому требованию обеспечения определенной структуры собственных векторов линейного оператора динамической системы «многомерный вход – многомерный выход» (МВМВ) [2], при которой один из собственных векторов принадлежит линейной оболочке обобщенного синхронного функционирования.

Сравнивая две возможные алгебраические формулировки задачи синхронного функционирования агрегатов по степени развитости алгоритмического обеспечения процедур синтеза законов управления агрегатами, приходится отдать предпочтение второй формулировке. Она имеет развитое алгоритмическое обеспечение, базирующееся на возможностях обобщенного модального управления [4], позволяющего доставлять матрице состояния многоагрегатной системы требуемые собственные значения и собственные векторы. Результаты, полученные в статье, базируются на этом алгебраическом подходе.

Завершая введение в проблему, следует сказать, что попытки решить задачу обеспечения синхронного функционирования многоагрегатных объектов МВМВ-типа были предприняты в работах [5–8]. Они опирались на возможности алгебраических свойств линейного оператора, представленного передаточной матрицей МВМВ-типа. Использование для этих целей векторно-матричных представлений метода пространства состояний существенно расширило возможности решения поставленной задачи.

Постановка задачи

Постановка задачи синтеза управления многоагрегатным техническим объектом (ТО) МВМВ-типа, доставляющего процессам в ТО по выходу синхронное функционирование, формулируется на основе использования векторно-матричного модельного представления этих процессов в форме

$$\dot{x}(t) = Ax(t) + Bu(t), \quad y(t) = Cx(t), \quad (1)$$

где u, x, y – соответственно векторы управления, состояния и выхода ТО; $u, y \in R^m$, $x \in R^n$; A, B, C – соответственно матрицы состояния, управления и выхода, согласованные по размерности с переменными объекта.

Сформулируем геометрическую трактовку задачи обеспечения синхронного функционирования агрегатов многоканального технического объекта (1).

Геометрическая постановка задачи синхронного изменения компонентов $y_l (l = \overline{1, m})$ m -мерного вектора выхода y в линейной по времени и по координатам объекта (1) под действием m -мерного вектора задающего воздействия $g(t)$ имеет вид

$$y(t) \in L_c \{g_0\} \quad \text{для } \forall t; \quad (2)$$

где $L_c \{g_0\}$ – линейная оболочка, натянутая на вектор g_0 , порождающий задающий вектор $g(t)$, параметризованный временем t и обладающий свойством вида (2)

$$g(t) \in L_c \{g_0\} \quad \text{для } \forall t. \quad (3)$$

Требуемая динамика синхронного движения компонентов вектора $y(t)$, заданного в форме (2), будет определяться динамикой изменения задающего воздействия $g(t)$ в линейной оболочке, натянутой на g_0 , задаваемой в форме

$$g(t) = \psi(t)g_0 \quad (4)$$

где $\psi(t)$ – скалярная функция, задающая динамику желаемого развития процессов в ТО (1), удовлетворяющих условию (2).

Введем два определения.

Определение 1. Будем называть задачу синхронизации задачей синхронизации в общепринятом смысле, если все компоненты вектора g_0 равны друг другу. $\square$

В этом случае значения скоростей и ускорений изменения всех компонентов, подчиненных условию (4), оказываются равными друг другу.

Определение 2. Будем называть задачу синхронизации обобщенной задачей синхронизации, если компоненты вектора g_0 произвольны. $\square$

В этом случае значения скоростей и ускорений изменения всех компонентов, подчиненных условию (4), могут отличаться друг от друга.

Для раскрытия в полной мере алгебраических возможностей матричных компонентов векторно-матричного описания объекта (1) и формируемой системы управления, которая строится на агрегировании ТО (1) и формирователя сигнала управления (ФСУ) этим объектом, необходимо направить усилия на анализ алгебраических свойств матричных функций от матриц. Решение задачи синтеза ФСУ будем искать в классе аналитических представлений сигнала управления $u(t)$ в форме

$$u(t) = K_g g(t) - Kx(t), \quad (5)$$

где K_g, K – соответственно матрицы прямой связи по задающему воздействию и отрицательной обратной связи по вектору состояния ТО (1). При построении ФСУ в форме (5) использована априорная гипотеза о полной измеримости вектора задающего воздействия $g(t)$ и вектора состояния ТО $x(t)$. Агрегирование модели ТО (1) и алгоритма функционирования ФСУ в форме (5) образует систему управления:

$$\dot{x}(t) = Fx(t) + Gg(t); \quad x(0), \quad y(t) = Cx(t), \quad (6)$$

где $F = A - BK$, $G = BK_g$.

Для анализа алгебраических свойств матричных функций от матриц целесообразно перейти к описаниям полученных выше векторно-матричных представлений с использованием лапласовых образов их переменных. В этой связи применение преобразования Лапласа к (4) дает

$$g(s) = \psi(s)g_0. \quad (7)$$

В свою очередь, переход в (6) к лапласовым образам позволяет записать:

$$y(s) = C(sI - F)^{-1}x(0) + C(sI - F)^{-1}Gg(s)|_{x(0)=0} = C(sI - F)^{-1}Gg(s); \quad x(0) = x(t)|_{t=0}. \quad (8)$$

Подстановка в (8) условия (7) дает

$$y(s) = C(sI - F)^{-1}Gg(s)|_{g(s)=\psi(s)g_0} = C(sI - F)^{-1}G\psi(s)g_0, \quad (9)$$

где $y(s), g(s)$ – лапласовы образы вектора выхода и вектора задающего входного воздействия; $\psi(t), \psi(s)$ – скалярная функция соответственно времени t и комплексной переменной s , задающая динамику желаемого развития процессов в ТО (1).

Основной результат. Поиск достаточных алгебраических условий обобщенной синхронизируемости многоканальных динамических объектов

Решение задачи конструирования матриц K_g и K в алгоритме (5) формирования сигнала управления, опирающейся на алгебраические свойства матричных компонентов соотношения (9), использует гипотезу о том, что выбором матриц K_g и K выражение (9) можно свести к представлениям вида

$$y(s) = C(sI - F)^{-1}G\psi(s)g_0 = C(sI - F)^{-1}BK_g\psi(s)g_0 = \zeta(s)g_0, \quad (10)$$

где $\zeta(s)$ – скалярная функция, имеющая вещественнозначное обратное преобразование Лапласа, определяющая динамику изменения вектора выхода ТО (1) в линейной оболочке $L_c\{g_0\}$, тем самым доставляет системе вида (6) обобщенное синхронное изменение компонентов вектора выхода $y(t)$ в форме (3).

Решение этой задачи содержится в следующих утверждениях.

Утверждение 1 (У1). Чтобы $y(t) \in L_c \{g_0\}$ или чтобы выполнялось соотношение $y(s) = \zeta(s)g_0$, где $\zeta(s)$ – скалярная функция переменной s , имеющая вещественно-значное обратное преобразование Лапласа, достаточно, чтобы выполнялись условия:

1. вектор $BK_g g_0$ был бы правым собственным вектором матрицы состояния F , т.е. выполнялось соотношение

$$F(BK_g g_0) = \lambda_i(BK_g g_0), \quad (11)$$

где $\lambda_i (i = \overline{1, n})$ – одно из вещественных собственных значений матрицы F ;

2. вектор g_0 был бы правым собственным вектором матрицы CBK_g , т.е. выполнялось равенство

$$(CBK_g)g_0 = \lambda_{gl}g_0, \quad (12)$$

где $\lambda_{gl} (l = \overline{1, m})$ – одно из вещественных собственных значений матрицы CBK_g , так что матрица прямой связи K_g ищется из условия

$$K_g = \arg \{(CBK_g)g_0 = \lambda_{gl}g_0\}. \quad \square (13)$$

Доказательство. Рассмотрим векторно-матричное выражение

$$y(s) = C(sI - F)^{-1}G\psi(s)g_0 = C(sI - F)^{-1}BK_g g_0\psi(s).$$

По свойству [8] матричной функции $f(F)$ сохранять геометрический спектр исходной матрицы F $F\xi_i = \lambda_i\xi_i$, где $\lambda_i : \det(\lambda I - F) = 0$, в форме

$$f(F)\xi_i = f(\lambda_i)\xi_i, \quad (14)$$

(11) и (14) позволяют для матричной функции от матрицы $f(F) = (sI - F)^{-1}$ записать цепочку соотношений

$$(sI - F)^{-1}BK_g g_0 = (s - \lambda_i)^{-1}BK_g g_0. \quad (15)$$

Подстановка (15) во второе представление выражения (10) дает

$$y(s) = CBK_g g_0\psi(s)(s - \lambda_i)^{-1}. \quad (16)$$

Если в (16) учесть (12), то получим:

$$y(s) = \lambda_{gl}\psi(s)(s - \lambda_i)^{-1}g_0 = \zeta(s)g_0, \quad \zeta(s) = \lambda_{gl}\psi(s)(s - \lambda_i)^{-1}. \quad \blacksquare (17)$$

Заметим, что собственные числа λ_i выбираются из требований к динамике развития процессов по выходу в подпространстве $L_c \{g_0\}$. Альтернативой полученному результату являются положения следующего утверждения.

Утверждение 2 (У2). Чтобы $y(t) \in L_c \{g_0\}$ или чтобы выполнялось соотношение $y(s) = \zeta(s)g_0$, где $\zeta(s)$ – скалярная функция переменной s , имеющая вещественно-значное обратное преобразование Лапласа, достаточно, чтобы выполнялись условия:

1. вектор $BK_g g_0$ удовлетворял соотношению (11), а, следовательно, и (15);
2. матрица CBK_g была бы единичной, т.е. применительно к матрице K_g выполнялось соотношение

$$K_g = \arg \{CBK_g = I\}. \quad \square (18)$$

Доказательство. На основании подстановки (15) в (10) получим выражение (16).

Если в (16) учесть (18), то получим

$$y(s) = g_0(s - \lambda_i)^{-1} \psi(s) = (s - \lambda_i)^{-1} \psi(s) g_0 = \zeta(s) g_0,$$

где $\zeta(s) = (s - \lambda_i)^{-1} \psi(s)$. ■(19)

Таким образом, если задачу понимать расширительно, то матрицы C и B при конструировании объекта управления и матрица K_g при синтезе закона управления должны формироваться из условия

$$(C, B, K_g) = \arg \{CBK_g = I\}. \quad (20)$$

Как и в случае утверждения 1, собственные числа λ_i выбираются из требований динамики развития процессов в подпространстве $L_c \{g_0\}$.

Утверждение 3(У3). Матрица M , приводящая произвольную $n \times n$ -квадратную матрицу F простой структуры к диагональной форме $\Lambda = \text{diag}[\lambda_i, i = \overline{1, n}]$ в силу соотношения

$$\Lambda = M^{-1}FM, \quad (21)$$

имеет своими столбцами [8] собственные векторы матрицы F . □

Доказательство. Запишем базовое уравнение матричного подобия для рассматриваемого случая,

$$M\Lambda = FM, \quad (22)$$

в столбцовой форме

$$M[\Lambda_1 \quad \Lambda_2 \quad \dots \quad \Lambda_i \quad \dots \quad \Lambda_n] = F[M_1 \quad M_2 \quad \dots \quad M_i \quad \dots \quad M_n], \quad (23)$$

где Λ_i, M_i – i -ые столбцы соответственно матриц Λ и M ($i = \overline{1, n}$). Перейдем теперь от матричного уравнения (22) к n векторно-матричным уравнениям вида

$$M\Lambda_i = FM_i; \quad i = \overline{1, n}, \quad (24)$$

где i -ый столбец Λ_i диагональной матрицы Λ имеет представление

$$\Lambda_i = [0 \quad \dots \quad 0 \quad \lambda_i \quad 0 \quad \dots \quad 0]^T. \quad (25)$$

Легко видеть, что с учетом (25) векторно-матричное уравнение (24) принимает вид

$$\lambda_i M_i = FM_i; \quad i = \overline{1, n}. \quad (26)$$

Векторно-матричное соотношение (26) является определением собственного вектора матрицы F , откуда следует, что M_i – собственный вектор матрицы F . ■

Положения утверждений У1–У3, дополненные концепцией обобщенного модального управления [9], являются алгоритмической основой для синтеза управления и доставляют системе МВМВ-типа достаточные алгебраические условия обобщенной синхронизируемости многоканальных динамических объектов.

Приведем алгоритмы конструирования матричных компонентов закона управления, присвоив им номера 1 и 2, а также сопроводим их аббревиатурой А1 и А2 соответственно.

Алгоритм 1 (А1) конструирования матричных компонентов закона управления (5), опирающийся на положения утверждения 1.

1. Задать g_0 , определяющий $L_c \{g_0\}$ – линейную оболочку, натянутую на вектор g_0 , единичной размерности, которой в процессе функционирования ОУ должен принадлежать вектор выхода $y(t)$.
2. Сформировать требования к показателям качества в переходном и установившемся режимах процессов, протекающих в подпространстве $L_c \{g_0\}$.
3. Сформировать векторно-матричное представление технического объекта в форме (1).

4. Задать λ_{gl} и решить уравнение $K_g = \arg\{CBK_g = \lambda_{gl}I\}$;
5. Сформировать собственный вектор матрицы F $\xi_i = BK_g g_0$;
6. Сформировать модальную модель с матрицами Λ , H , образующих наблюдаемую пару, где диагональная матрица Λ является носителем желаемого спектра собственных значений матрицы F состояния проектируемой системы, в состав которого входит собственное значение λ_i , удовлетворяющее требованию

$$\lambda_i = \arg\{(\lambda_i I - A)\xi_i \in JmB\}, \quad (27)$$

где JmB – образ матрицы B , т.е. пространство ее столбцов;

7. Записать матричное уравнение Сильвестра $M\Lambda - AM = -BH$ в декомпозированной форме

$$\begin{bmatrix} \tilde{M} & \tilde{M} \end{bmatrix} \begin{bmatrix} \tilde{\Lambda} & 0 \\ 0 & \bar{\Lambda} \end{bmatrix} - A \begin{bmatrix} \tilde{M} & \tilde{M} \end{bmatrix} = -B \begin{bmatrix} \tilde{H} & \bar{H} \end{bmatrix}, \quad (28)$$

где в силу утверждения 3 $\tilde{M} = \xi_i$, $\tilde{\Lambda} = \lambda_i$: $F\xi_i = \lambda_i \xi_i$.

8. Выделить из системы (28) матричное уравнение Сильвестра,

$$(\lambda_i I - A)\xi_i = -B\tilde{H} \quad (29)$$

и найти его решение относительно матрицы $\tilde{H}$, которое в случае $\text{rang} B = m = n = \dim x$ принимает вид $\tilde{H} = B^{-1}(A - \lambda_i I)\xi_i$, а в случае $\text{rang} B = m < n$ находится из соотношения $\tilde{H} = (B^T B)^{-1} B^T (A - \lambda_i I)\xi_i$.

9. Выделить из системы (28) матричное уравнение Сильвестра

$$\bar{M}\bar{\Lambda} - A\bar{M} = -B\bar{H}, \quad (30)$$

в котором $(\bar{\Lambda}, \bar{H})$ – наблюдаемая пара, $\bar{\Lambda} = \text{diag}\{\lambda_j : j \neq i; j = \overline{1, n}\}$, и решить уравнение (30) при известных $\bar{\Lambda}$, A , B и $\bar{H}$ относительно матрицы $\bar{M}$.

10. Сформировать матрицы $M = \begin{bmatrix} \tilde{M} & \tilde{M} \end{bmatrix}$ и $H = \begin{bmatrix} \tilde{H} & \bar{H} \end{bmatrix}$, при этом обеспечить максимальную близость собственных векторов матрицы F $\xi_j = M_j (j \neq i)$ для обеспечения модальной робастности системы к ортогональности с учетом возможной минимизации затрат на управление.

11. Вычислить матрицу K обратных связей с помощью соотношения

$$K = HM^{-1} = \begin{bmatrix} \tilde{H} & \bar{H} \end{bmatrix} \begin{bmatrix} \tilde{M} & \tilde{M} \end{bmatrix}^{-1}. \quad (31)$$

12. Построить реализационную версию закона управления в форме

$$\begin{aligned} u(t) &= K_g g(t) - Kx(t) = K_g g(t) - K_y y(t) - K_x x(t) |_{K_y = K_g} = \\ &= K_g (g(t) - y(t)) - K_x x(t) = K_g \varepsilon(t) - K_x x(t) |_{K_g = K_\varepsilon} = K_\varepsilon \varepsilon(t) - K_x x(t) \end{aligned} \quad (32)$$

где $K_x = K - K_g C$.

Алгоритм 2(A2) конструирования матричных компонентов закона управления (5), опирающийся на положения утверждения 2

1. Выполнить пп. 1–3 алгоритма 1.
2. Выполнить п. 4 алгоритма 1 при условии выполнения равенства $\lambda_{gl} = 1$;
3. Выполнить пп. 5–11 алгоритма 1.

Заключение

Основным результатом работы являются алгоритмы конструирования матричных компонентов закона управления, позволяющего обеспечить обобщенную синхронность функционирования динамических агрегатов, входящих в состав технологического процесса. Однако следует отметить, что при решении поставленной задачи использован только один собственный вектор матрицы отношения «вход–выход». Свобода назначения остальных $(m-1)$ векторов ожидает своего системного использования. Предстоит разработать критерии их назначения и адаптировать разработанную алгоритмическую базу под эту задачу.

Литература

1. Мирошник И.В. Согласованное управление многоканальными системами. – Л.: Энергоатомиздат, 1990.
2. Бочков А.Л., Дударенко Н. А., Ушаков А.В. Синтез многомерных функционально вырожденных динамических систем // Известия вузов. Приборостроение. – 2008. – Т. 51. – № 1. – С. 25–29.
3. Дударенко Н.А., Ушаков А.В. Вырождение сложных динамических систем с равно-темповыми структурными компонентами // Научно-технический вестник СПбГУ ИТМО. – 2005. – № 19. – С. 44–51.
4. Ушаков А.В. Обобщенное модальное управление // Известия вузов. Приборостроение. – 2000. – Т. 43. – № 3. – С. 8–16.
5. Лаврентьев В.В., Ушаков А.В. К вопросу о синтезе перекрестных связей, обеспечивающих синхронную работу параллельно включенных агрегатов // Приборы и системы автоматики. Труды ЛИТМО, вып. 85. – Л.: ЛИТМО, 1975.
6. Мирошник И.В., Ушаков А.В. Синтез алгоритма синхронного управления системой квазиоднотипных объектов // Автоматика и телемеханика. – 1977. – № 11.
7. Никифоров В.О., Ушаков А.В. Управление в условиях неопределенности: чувствительность, адаптация, робастность. – СПб: СПб ГИТМО (ТУ), 2002.
8. Гантмахер Ф.Р. Теория матриц. – М.: Наука, 1973.
9. Власенко В.А., Мирошник И.В., Сабинин Ю.А., Ушаков А.В. и др. Система управления лентопротяжным механизмом // Электротехн. пром. Серия «Электропривод». – 1977. – Вып. 5 (58).

<i>Дударенко Наталия Александровна</i>	–	Санкт-Петербургский государственный университет информационных технологий, механики и оптики, кандидат технических наук, доцент, Dudarenko@yandex.ru
<i>Полякова Майя Вячеславовна</i>	–	Санкт-Петербургский государственный университет информационных технологий, механики и оптики, аспирант, 12noch@mail.ru
<i>Ушаков Анатолий Владимирович</i>	–	Санкт-Петербургский государственный университет информационных технологий, механики и оптики, доктор технических наук, профессор, Ushakov-AVG@yandex.ru

УДК 621.396.6

АЛГОРИТМ УСКОРЕННОГО ПОВТОРНОГО РЕШЕНИЯ СИСТЕМ ЛИНЕЙНЫХ АЛГЕБРАИЧЕСКИХ УРАВНЕНИЙ И ЕГО ИСПОЛЬЗОВАНИЕ ПРИ МАТЕМАТИЧЕСКОМ МОДЕЛИРОВАНИИ ЭЛЕКТРОННЫХ УСТРОЙСТВ

М.С. Ревин, Н.С. Савелов

Излагаются численные эксперименты с разработанной программой, реализующей модификацию метода исключения Гаусса, которая позволяет получить значительные преимущества в скорости вычислений при повторных решениях систем линейных алгебраических уравнений. Приводится анализ нового метода в сравнении с известным методом Шермана–Моррисона. Применение этой модификации для ускорения процесса формирования уравнений состояния позволило разработать микропроцессорную систему, способную в реальном времени контролировать температурный режим работы электротехнического объекта.

Ключевые слова: метод, моделирование, эксперименты, уравнения состояния, микропроцессорная система, электротехнический объект.

Введение

Задача повышения эффективности методов математического моделирования является особо актуальной применительно к синтезу и оптимизации различных технических устройств и систем. Математическое моделирование достаточно сложных объектов уже требует десятков часов машинного времени при использовании высокопроизводительной вычислительной техники. Совершенствование методов приводит к существенному сокращению вычислительных затрат на математическое моделирование. При анализе электрических цепей, например, при нахождении различных зависимостей токов от напряжений или от нелинейных сопротивлений, часто необходимо выполнять повторные решения систем линейных алгебраических уравнений (СЛАУ) вида $Ax = b$ после изменения матрицы A или столбца b . В подобных задачах при повторных решениях в системе изменяются не все элементы, а лишь некоторые, и решать систему заново в этом случае нецелесообразно, так как это требует больших вычислительных затрат. Для их сокращения часто применяют формулу Шермана–Моррисона [1], которая на порядок уменьшает число арифметических операций, необходимых для нахождения повторного решения, но в определенных случаях указанная модификация требует меньших вычислительных затрат.

Математические основы метода

Перейдем к краткому описанию алгоритма ускоренного повторного решения СЛАУ. Рассмотрим СЛАУ вида $Ax = b$, где A – квадратная действительная невырожденная матрица порядка n ; x – столбец неизвестных; b – действительный столбец. Характерной особенностью используемой модификации метода исключения Гаусса является преобразование дополнительной матрицы F , которая в исходном виде является единичной. Используем обозначения: a_i – i -ый столбец матрицы A ; f_i – i -я строка F ; $F^{(m)}, f^{(m)}$ – соответственно F и f после m -го изменения, $m=0,1,\dots,n$; E – единичная матрица порядка n .

Для изменения матрицы F и определения искомых неизвестных (т.е. неизвестных, подлежащих определению) используются приведенные ниже выражения в предположении, что $f^{(m)}_{m+1} \cdot a_{m+1} \neq 0$ (в противном случае достаточно выполнить перестановку двух строк матрицы $F^{(m)}$):

$F^{(0)} = E$; $f^{(m+1)}_i = f^{(m)}_i$ для $i = m+1$ (при $m = 0, 1, \dots, n-1$), а также для $i < m+1$ (при $m = 1, 2, \dots, n-1$) и одновременно при условии, что x_i – неискомое;

$$f^{(m+1)}_i = f^{(m)}_i - (f^{(m)}_i \cdot a_{m+1} / f^{(m)}_{m+1} \cdot a_{m+1}) \cdot f^{(m)}_{m+1} \text{ для остальных } i.$$

Если принять $k1 = f^{(m)}_i \cdot a_{m+1}$, а $k2 = f^{(m)}_{m+1} \cdot a_{m+1}$, то

$$f^{(m+1)}_i = f^{(m)}_i - k1 \cdot (f^{(m)}_{m+1} / k2),$$

$$x_i = (f^{(n)}_i \cdot b) / (f^{(n)}_i \cdot a_i) \text{ (} x_i \text{ – искомое).}$$

Ведущей строкой $f^{(n)}_i$ для столбца a_i названа строка, используемая для изменения других строк при обращении к a_i . Образующей строкой для a_i названа строка $f_{io} = (1 / f^{(n)}_i \cdot a_i) \cdot f^{(n)}_i$.

Пусть для матрицы A ($\det A \neq 0$) сформирована матрица F , а затем столбец a_i такой, что x_i – искомое, заменен на столбец a'_i , в результате чего A преобразована в A' . Пусть f_{io} – образующая для a_i . При этом $f_{io} \cdot a'_i = 0$ тогда и только тогда, когда $\det A' = 0$. Поэтому для невырожденной A' искомые неизвестные могут быть определены следующим образом. Пусть имеется другое, кроме x_i , искомое x_j и f_{jo} – образующая для a_j . Новые значения искомых x'_i и x'_j определяются из следующих выражений:

$$f'_{io} = (1 / (f_{io} \cdot a'_i)) \cdot f_{io}, \quad x'_i = f'_{io} \cdot b,$$

$$f'_{jo} = f_{jo} - (f_{jo} \cdot a'_i) \cdot f_{io}, \quad x'_j = f'_{jo} \cdot b \text{ или } x'_j = x_j - (f_{jo} \cdot a'_i) \cdot x'_i.$$

Подробная реализация алгоритма на конкретном примере изложена в [2].

Результаты экспериментального исследования метода

Обратимся к задаче анализа электрической цепи (рис. 1). Пусть необходимо выяснить, как изменяются токи в цепи при изменении сопротивления $r3$ в пределах от 0 до 10 кОм с шагом 1 Ом.

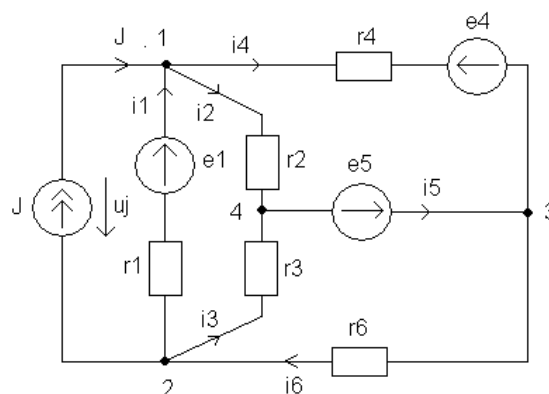


Рис. 1. Электрическая цепь

Полная система уравнений Кирхгофа для решения этой задачи имеет следующий вид:

$$\begin{pmatrix} -1 & 1 & 0 & 1 & 0 & 0 & 0 \\ 0 & -1 & -1 & 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & -1 & -1 & 1 & 0 \\ r1 & r2 & -r3 & 0 & 0 & 0 & 0 \\ 0 & -r2 & 0 & r4 & 0 & 0 & 0 \\ 0 & 0 & r3 & 0 & 0 & r6 & 0 \\ r1 & 0 & 0 & 0 & 0 & 0 & 1 \end{pmatrix} \cdot \begin{pmatrix} i1 \\ i2 \\ i3 \\ i4 \\ i5 \\ i6 \\ u_j \end{pmatrix} = \begin{pmatrix} j \\ 0 \\ 0 \\ e1 \\ -e4 - e5 \\ e5 \\ e1 \end{pmatrix}$$

$$\begin{aligned} a &\leftarrow \begin{pmatrix} -1 & 1 & 0 & 1 & 0 & 0 & 0 \\ 0 & 1 & 1 & 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & -1 & -1 & 1 & 0 \\ 5 & 4 & -6 & 0 & 0 & 0 & 0 \\ 0 & -4 & 0 & 6 & 0 & 0 & 0 \\ 0 & 0 & 6 & 0 & 0 & 10 & 0 \\ 5 & 0 & 0 & 0 & 0 & 0 & 1 \end{pmatrix} \\ b &\leftarrow \begin{pmatrix} 0.1 \\ 0 \\ 0 \\ 30 \\ -50 \\ 20 \\ 30 \end{pmatrix} \\ \text{for } i \in 0, 1 \dots 10 \\ &\left| \begin{aligned} a_{3,2} &\leftarrow -i \\ a_{5,2} &\leftarrow i \\ c^{(i)} &\leftarrow \text{solve}(a, b) \end{aligned} \right. \\ c \end{aligned}$$

	0	1	2	3	4	5	6	7
0	1.319	1.382	1.426	1.457	1.481	1.499	1.514	1.527
1	5.651	5.889	5.915	5.934	5.948	5.96	5.969	5.976
2	0.581	0.47	0.395	0.341	0.299	0.267	0.241	0.22
3	-4.432	-4.407	-4.39	-4.377	-4.368	-4.36	-4.354	-4.349
4	6.432	6.36	6.311	6.275	6.248	6.227	6.21	6.196
5	2	1.953	1.921	1.898	1.88	1.866	1.855	1.846
6	23.405	23.088	22.871	22.715	22.596	22.503	22.428	22.367

Рис. 2. Решение задачи анализа электрической цепи с помощью MathCad 14

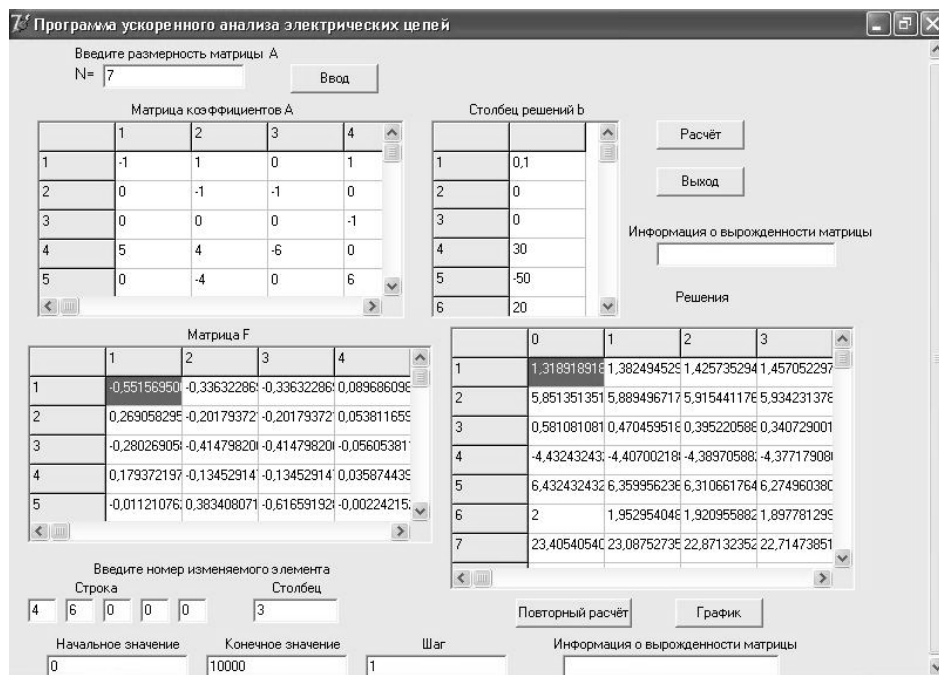


Рис. 3. Решение с помощью разработанной программы

В матрице изменяются только два элемента – в 4-й строке, 3-м столбце и в 6-й строке, 3-м столбце. С помощью пакета программ MathCAD 14 и разработанной программы были получены следующие результаты. MathCAD 14 получил все решения при каждом изменении r_3 спустя 1 мин. 10 с. Разработанная программа решила задачу за время, меньшее 1 с. Если же усложнить задачу и изменять r_3 в пределах от 0 до 1 Ом с шагом 1 Ом, то время решения увеличилось до 8 с. Численные значения решений, полученные с помощью разработанной программы и с помощью пакета программ MathCad 14, во всех случаях совпадали.

Таким образом, разработанная по описанной модификации программа позволяет получить значительные преимущества в скорости вычислений при анализе электрических цепей [3].

Сравнительный анализ формулы Шермана–Моррисона и указанной модификации на примере решения задачи, типичной для анализа электрических цепей, когда в матрице A изменяется только один столбец, приведен в работе [4]. В этой работе получены математические выражения, определяющие число арифметических операций, необходимых для повторного решения при использовании формулы Шермана–Моррисона по алгоритму, описанному в [1], и альтернативного метода. При этом были использованы все потенциальные, не указанные в [1], преимущества формулы Шермана–Моррисона.

Важным преимуществом описанного метода является то, что он позволяет находить значения не всех элементов столбца неизвестных x , а только тех, которые необходимы пользователю, что позволяет значительно снизить вычислительные затраты. Так, при повторном определении одного неизвестного, соответствующего изменяющемуся столбцу, потребуется n умножений, $n-1$ сложение и 1 деление, т.е. $2n$ арифметических операций. Для нахождения любого другого неизвестного потребуется $2n+1$ умножений, $2n-1$ сложение и 1 деление, т.е. $4n+1$ арифметических операций. На нахождение решения по формуле Шермана–Моррисона потребуется $n^2 + 2n + 1$ умножений, 1 деление и $n^2 + n$ сложений. Всего требуется $2n^2 + 3n + 2$ арифметических операций, что на порядок больше, чем с помощью модификации.

При определении всех неизвестных алгоритмы равноценны. Но в практических вычислениях часто необходимо определять не все элементы столбца x , а только те, которые необходимы пользователю. В этом случае описанный метод значительно превосходит формулу Шермана–Моррисона.

Применение метода для моделирования динамических объектов

Описанную модификацию целесообразно использовать и для ускорения процесса формирования уравнений состояния, т.е. для ускорения процесса моделирования динамических объектов.

Уравнения состояния в виде системы обыкновенных дифференциальных уравнений в форме Коши являются одной из наиболее важных форм представления математической модели объектов контроля, управления и диагностики.

Порядок первоначального формирования системы уравнений для переменных состояния был изложен в работе [5].

Матричное уравнение для переменных состояния имеет вид

$$\frac{dx}{dt} = A \cdot x + B_0 \cdot v + B_1 \cdot \frac{dv}{dt} + \dots B_p \cdot \frac{d^p v}{dt^p}, \quad (1)$$

где $A, B_0 \dots B_p$ – действительные матрицы; x – столбец переменных состояния, полученных из x исключением зависимых элементов, v – столбец независимых величин (токов и напряжений).

Применяя алгоритм, основанный на модификации метода Гаусса, можно быстро переформировывать уравнение (1).

На основании описанного алгоритма была разработана микропроцессорная система для контроля температурного режима силового полупроводникового прибора, моделируемого электрической схемой. Эта схема может применяться, в частности, в устройствах индукционной плавки.

Традиционный способ тепловой защиты преобразователя основан на контроле с помощью датчиков температуры охлаждающей среды тиристоров или транзисторов. Вследствие того, что тепловые постоянные времени кристаллов силовых приборов на несколько порядков меньше других постоянных времени, измерение температуры само по себе не способно надежно защитить силовой модуль от больших кратковременных токовых перегрузок.

Тепловую модель тиристора можно представить в виде эквивалентной электрической схемы замещения (рис. 4).

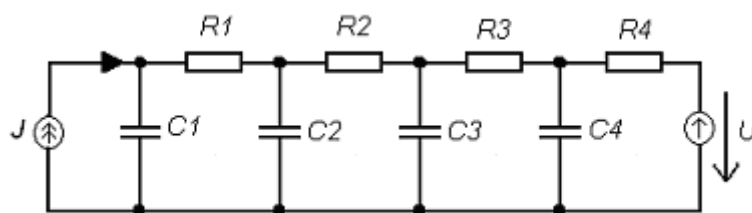


Рис. 4. Тепловая модель кристалла тиристора

Мощностям тепловых потерь в кристалле соответствует ток J , а температуре в характерных точках структуры прибора – напряжения на конденсаторах. Температуре, измеряемой датчиком, соответствует источник ЭДС.

В эквивалентной схеме все сопротивления, за исключением $R4$, постоянны. Микропроцессорная система обеспечивает измерение $R4$ и U и производит переформирование уравнений состояния, а затем, на основании численного решения задачи Коши, определяя температуру кристалла силового тиристора.

Для расчета в память микроконтроллера в соответствии с алгоритмом загружается дополнительная матрица F , которая позволяет с использованием минимального числа арифметических операций рассчитать режим работы устройства.

В качестве основы системы использован микроконтроллер AVR фирмы Atmel.

Применение разработанного алгоритма в микропроцессорной системе позволило организовать защиту температуры кристалла в режиме реального времени.

Моделирование в программе AVR Studio показало, что на расчет температуры кристалла тиристора микропроцессорной системе потребуется около 1,3 мс, что почти в 2 раза меньше наименьшей тепловой постоянной кристалла тиристора, которые лежат в интервале 0,002–0,1 с. Таким образом, можно говорить о том, что разработанная система способна в реальном времени контролировать тепловой режим.

Заключение

Использование описанной модификации метода решения СЛАУ позволяет существенно сократить вычислительные затраты как при предварительном математическом моделировании различных объектов, так и при их моделировании в режиме реального времени. Сравнительный анализ новой модификации метода исключения Гаусса с наилучшим методом для повторного решения СЛАУ – формулой Шермана–Моррисона – показал, что даже в случае использования всех потенциальных преимуществ последней при определении всех неизвестных алгоритмы равноценны. Но в практических вычис-

лениях часто необходимо определять не все элементы столбца x , а только те, которые необходимы пользователю. В этом случае альтернативный метод значительно превосходит формулу Шермана–Моррисона.

Литература

1. Каханер Д., Моулер К., Нэш С. Численные методы и математическое обеспечение: Пер. с англ. – М.: Мир, 1998. – 575 с.
2. Савелов Н.С., Ревин М.С. Численные эксперименты с программным обеспечением для решения систем линейных алгебраических уравнений // Методы и алгоритмы прикладной математики в технике, медицине и экономики: Материалы 7-й междунар. науч.-практ. конф. – Новочеркасск: ЮРГТУ, 2007. – Ч. 1. – С. 85–90.
3. Ревин М.С., Савелов Н.С. Программная реализация метода ускоренного анализа электрических цепей // Труды 50-й научной конференции МФТИ «Современные проблемы фундаментальных и прикладных наук»: Управление и прикладная математика. Том 2. – М.: МФТИ, 2007. – 164 с. – С. 137–138.
4. Савелов Н.С., Ревин М.С. Сравнительный анализ формулы Шермана – Моррисона и альтернативного метода для повторного решения систем линейных алгебраических уравнений // Методы и алгоритмы прикладной математики в технике, медицине и экономики: Материалы 8-й междунар. науч.-практ. конф. – Новочеркасск: ЮРГТУ, 2008. – С. 12–15.
5. Савелов Н.С. Полные исходные системы уравнений электронных схем и формирование частично символьных функций // Изв. высш. учеб. заведений. Электромеханика. – 2007. – № 3. – С. 3–6.

Ревин Михаил Сергеевич

– Южно-Российский государственный технический университет (Новочеркасский политехнический институт), аспирант, misharevin86@mail.ru

Савелов Николай Семенович

– Южно-Российский государственный технический университет (Новочеркасский политехнический институт), кандидат технических наук, доцент, savelovn@mail.ru

УДК 004.89

ТЕОРЕТИКО-МНОЖЕСТВЕННЫЙ ПОДХОД
К ЛОГИЧЕСКОМУ ВЫВОДУ В БАЗАХ ЗНАНИЙ

И.А. Бессмертный

Рассматривается проблема комбинаторной сложности задачи поиска решений при резолюции правил в системах искусственного интеллекта. Предлагается подход к ускорению извлечения знаний путем индексации фактов и сокращения числа используемых фактов с помощью операций над множествами индексов. В отличие от известных алгоритмов, предполагающих предварительный отбор фактов для каждого правила, индексы абстрагируются от правил, что позволяет логически и физически разделить базы фактов и базы правил, а также упростить модификацию базы знаний. Демонстрируется возможность замены логического вывода реляционными операциями над кортежами переменных.

Ключевые слова: искусственный интеллект, индексация фактов, логический вывод.

Состояние проблемы и существующие решения

В большинстве систем искусственного интеллекта (ИИ) имеется база знаний, хранящая факты и правила для предметной области. Применение правил к известным фактам позволяет порождать новые знания. Для извлечения знаний используется интеллектуальный агент (машина вывода), который либо находит существующие факты, либо порождает новые факты путем логического вывода из правил. Вывод из правил представляет собой классическую задачу неинформированного поиска [1] и в случае небольших баз знаний может решаться путем последовательного перебора всех фактов для каждого из условий правила.

Пусть множество фактов базы знаний $F = \{f\}$ образует атомы (триплеты) $f = (s, p, o)$, где s – субъект, p – предикат, o – объект. Мощность множества фактов обозначим n . База знаний содержит также множество правил $R = \{r\}$. Каждое правило r состоит из результирующей части и тела (условия правила). Для простоты будем считать, что результирующая часть содержит один атом, а тело – множество атомов $C = \{c\}$. Пусть среднее число условий в одном правиле равно k , а мощность множества правил равна m . Запуск процедуры применения всех правил к множеству фактов вызывает последовательный перебор всех фактов для каждого из условий правил. Общее число A попыток унификации условий с фактами будет равно

$$A = mn^k.$$

Данная формула показывает линейный рост сложности поиска при добавлении новых правил, степенную зависимость от числа фактов и показательную – от среднего числа условий в правиле. Даже трехзначные значения числа фактов в базе знаний и однозначные числа условий в правилах порождают миллионы вершин дерева поиска. Таким образом, комбинаторная сложность задачи даже для небольших баз знаний не позволяет решать ее путем «буквального» или «наивного» логического вывода.

Наиболее известным методом ускорения резолюции правил является алгоритм *Rete* [2], используемый в экспертных системах *CLIPS*, *Jess*, *Soar* и др. В основу алгоритма положено префиксное дерево, узлами которого являются условия правил. В каждом узле префиксного дерева создается список фактов из базы знаний, которые удовлетворяют условиям правила. Фактически это означает многократное дублирование фактов (рабочей памяти) в области правил (операционной памяти). При перемещении от корня префиксного дерева до листа списки фактов подвергаются операции пересечения,

и в листе находятся только факты, необходимые и достаточные для резолюции правила. Перемещение от корня до листа как раз и означает резолюцию правила.

Узким местом алгоритма *Rete* является изменение префиксного дерева при изменении фактов базы знаний. При добавлении, изменении или удалении фактов префиксное дерево должно строиться заново либо модифицироваться. Большинство модификаций алгоритма *Rete*, например, *Modify-in-place*, *Scaffolding*, *Decision Tree* [3], нацелены именно на изменение префиксного дерева. Между тем, модификация множества фактов происходит в системах ИИ постоянно, поскольку каждое правило дает в качестве результата новые факты, которые должны сразу же использоваться в ходе резолюции.

Индексация и предварительный отбор фактов

Проиндексируем факты следующим образом. Присвоим каждому факту в базе знаний порядковый номер i , тогда нумерованный факт будет выглядеть следующим образом:

$$fn = (i, s, p, o).$$

Для множества термов $T = \{t\}$, встречающегося в фактах, построим индекс в виде

$$X = \{x\} = \{(t, w, \{i_{tw}\})\}, \quad (1)$$

где w – место данного терма в атоме (в качестве субъекта, предиката или объекта), $\{i_{tw}\}$ – множество номеров фактов, имеющих терм t в роли w ; $w = ('s'; 'p'; 'o')$. Резолюция правила заключается в установлении истинности условий и присвоении значений переменным. Теперь при обращении к правилу, тело которого состоит из множества условий $\{c_1, c_2, \dots, c_k\}$, где $c_j = (s_j, p_j, o_j)$, s_j – субъект, p_j – предикат, o_j – объект, $(s; o) = (t, v)$, v – переменная, $p_j = t$. Поскольку в логике первого порядка не допускается использование переменных в качестве предиката, каждое из условий c_j может иметь одно из четырех сочетаний термов и переменных: (t, t, t) ; (t, t, v) ; (v, t, t) ; (v, t, v) .

Для каждого из c_j условий правила извлечение релевантных фактов для перечисленных сочетаний термов заключается в нахождении пересечений множеств индексов:

$$I_j = \{i_{ts}\} \cap \{i_{tp}\} \cap \{i_{to}\}, s_j = const, p_j = const, o_j = const;$$

$$I_j = \{i_{tp}\} \cap \{i_{to}\}, p_j = const, o_j = const;$$

$$I_j = \{i_{ts}\} \cap \{i_{tp}\}, s_j = const, p_j = const;$$

$$I_j = \{i_{tp}\}, p_j = const.$$

Каждой переменной v , используемой в j -м условии, из списков I_j можно поставить в соответствие множество кортежей $\{i, u_i\}$, где i – номер факта, $i \in I_j$, u_i – значение переменной v , извлекаемое из i -го факта. Если переменная v используется более чем в одном условии правила, пересечение

$$U_v = \{u_v\} = \bigcap_{j \in C} \{u\}_j$$

множеств значений переменной во всех условиях C правила, в которых эта переменная используется, позволит сократить число фактов, требуемых для унификации этих условий. Для получения списка фактов I_{vj} , содержащих переменную v для j -го условия правила, где эта переменная встречается, достаточно выполнить операцию реляционного деления

$$I_{vj} = \{i, u_i\} \div \{uv\}.$$

Наконец, если в условии правила c_j участвуют более одной переменной, то пересечение списков

$$I_j = \bigcap_{v \in c_j} I_{vj}$$

для каждой из двух переменных даст окончательный список фактов, которые отвечают j -му условию правила.

Рассмотрим данный алгоритм на простом примере. Пусть имеется база знаний, состоящая из фактов, имеющих сквозную нумерацию (здесь и далее будем придерживаться синтаксиса языка Prolog):

$fn(1, sergey, has_sex, male).$	$fn(8, sergey, parent, nikita).$
$fn(2, nikita, has_sex, male).$	$fn(9, sergey, parent, andrey).$
$fn(3, sergey, is_a, person).$	$fn(10, natalia, parent, nikita).$
$fn(4, natalia, is_a, person).$	$fn(11, natalia, parent, andrey).$
$fn(5, nikita, is_a, person).$	$fn(12, nikita, parent, stepan).$
$fn(6, nikita, has_sex, male).$	$fn(13, andrey, parent, egor).$
$fn(7, natalia, has_sex, female).$	

Построим для этих фактов индекс, как показано в формуле (1):

$x(sergey, s, [1,3,8,9]).$	$x(male, o, [1,2,6]).$
$x(nikita, s, [2,5,6,12]).$	$x(person, o, [3,4,5]).$
$x(natalia, s, [4,7,10,11]).$	$x(female, o, [7]).$
$x(andrey, s, [13]).$	$x(nikita, o, [8,10]).$
$x(has_sex, p, [1,2,6,7]).$	$x(andrey, o, [9,11]).$
$x(is_a, p, [3,4,5]).$	$x(stepan, o, [12]).$
$x(parent, p, [8,9,10,11,12,13]).$	$x(egor, o, [13]).$

Создадим правила в виде $r(conditionList, resultingList)$ с использованием переменных, начинающихся с вопросительного знака, где $conditionList$ – список условий, $resultingList$ – список триплетов результата. Первое правило устанавливает, что субъект является мужчиной, если он является человеком и имеет мужской пол.

$$r([c(“?x”, is_a, person), c(“?x”, has_sex, male)], [f(“?x”, is_a, man)]).$$

	1-е условие правила	2-е условие правила
Исходное условие	$?x, is_a, person$	$?x, has_sex, male$
Индекс для используемых термов	$x(is_a, p, [3,4,5]).$ $x(person, o, [3,4,5]).$	$x(has_sex, p, [1,2,6,7]).$ $x(male, o, [1,2,6]).$
Пересечение по номерам	$?x, s, [3,4,5]$	$?x, s, [1,2,6]$
Значения переменной	$?x = [sergey, natalia, nikita]$	$?x = [sergey, nikita, nikita]$
Пересечение по значениям	$[sergey, nikita]$	
Отфильтрованный индекс	$?x, s, [3, 5]$	$?x, s, [1,2,6]$

Таблица 1. Предварительный отбор фактов для правила $?x is_a man$

Используем индекс для отбора фактов и находим пересечение значений переменной $?x$, используемых во всех условиях.

$$[sergey, natalia, nikita] \cap [sergey, nikita, nikita] = [sergey, nikita]$$

Фильтруем списки фактов, требуемых для резолюции каждого условия правила, включая в них только значения термов, попавшие в пересечение. Для первого условия это факты $[3,5]$, а для второго – $[1,2,6]$. Таким образом, для резолюции правила требуется унифицировать условия правила с пятью фактами. Если не использовать индекс, то каждое условие правила нужно сопоставлять со всеми 13-ю фактами, всего 26 фактов.

Создадим более сложное правило, определяющее отношение «прародитель» (grandparent):

$$r([c(“?x”, parent, “?y”), c(“?y”, parent, “?z”), [f(“?x”, grandparent, “?z”)]).$$

	1-е условие правила	2-е условие правила
Исходное условие	"?x", <i>parent</i> , "?y"	"?y", <i>parent</i> , "?z"
Индекс для используемых термов	$x(\textit{parent}, p, [8,9,10,11,12,13])$.	$x(\textit{parent}, p, [8,9,10,11,12,13])$.
Пересечение по номерам	?x, <i>s</i> , [8,9,10,11,12,13]). ?y, <i>o</i> , [8,9,10,11,12,13]).	?y, <i>s</i> , [8,9,10,11,12,13]). ?z, <i>o</i> , [8,9,10,11,12,13]).
Значения переменной	[<i>nikita, andrey, nikita, andrey, stepan, egor</i>]	[<i>sergey, sergey, natalia, natalia, nikita, andrey</i>]
Пересечение по значениям	[<i>nikita, andrey</i>].	
Отфильтрованный индекс	[8,9,10,11]	[12,13]

Таблица 2. Предварительный отбор фактов для правила ?x *grandparent* ?z

Более одного раза в правиле встречается только переменная ?y; пересечение ее допустимых значений в первом и втором условиях равно

[*nikita, andrey, nikita, andrey, stepan, egor*] $\cap$ [*sergey, sergey, natalia, natalia, nikita, andrey*] = [*nikita, andrey*].

В качестве допустимых значений переменной ?y мы получили список объектов, которые являются одновременно родителями и детьми. Прореживаем список фактов для первого условия [8,9,10,11,12,13], оставляя в нем только факты, имеющие термы [*nikita, andrey*] в качестве объекта. Получаем список [8,9,10,11]. Аналогично получаем список фактов [12,13] для второго условия. Находим пересечение номеров фактов, соответствующих каждой из переменных, в каждом из условий. Для первого условия [8,9,10,11,12,13] $\cap$ [8,9,10,11] = [8,9,10,11], для второго – [12,13] $\cap$ [8,9,10,11,12,13] = [12,13]. Таким образом, мы добиваемся того, что каждому из условий правила в процессе унификации будут предъявляться только факты, гарантированно порождающие результаты.

Добавление нового факта в базу знаний вручную либо как результата резолюции правила не требует сложных вычислений. Для этого нужно присвоить факту номер и модифицировать три индекса, соответствующие субъекту, предикату и объекту, либо создать новые индексы, если данные сущности встречаются в первый раз.

Оценка быстродействия алгоритма

Предлагаемый здесь метод реализован в программе Semantic, разработанной автором специально для изучения принципов построения систем искусственного интеллекта, визуализации знаний, а также для исследований способов построения интеллектуальных агентов. Поскольку прямой логический вывод предполагает независимую обработку каждого правила, длительность вывода имеет линейную зависимость от числа правил. Следовательно, оценку быстродействия можно оценивать на одном правиле. Тестирование данного метода проводилось на базе знаний, описывающей родственные отношения. Факты вида субъект–*is_parent*–объект создавались с помощью генератора случайных чисел. Правило, которое использовалось для измерения времени вывода, описывает отношение типа «прапрародитель» и имеет три условия:

$r([c("?x1", "is_parent", "?x2"), c("?x2", "is_parent", "?x3"), c("?x3", "is_parent", "?x4")], [f("?x1", "is_grandgrandparent", "?x4")])$.

Приведенный ниже график (см. рис.) показывает зависимость времени работы программы логического вывода от числа фактов в базе знаний.

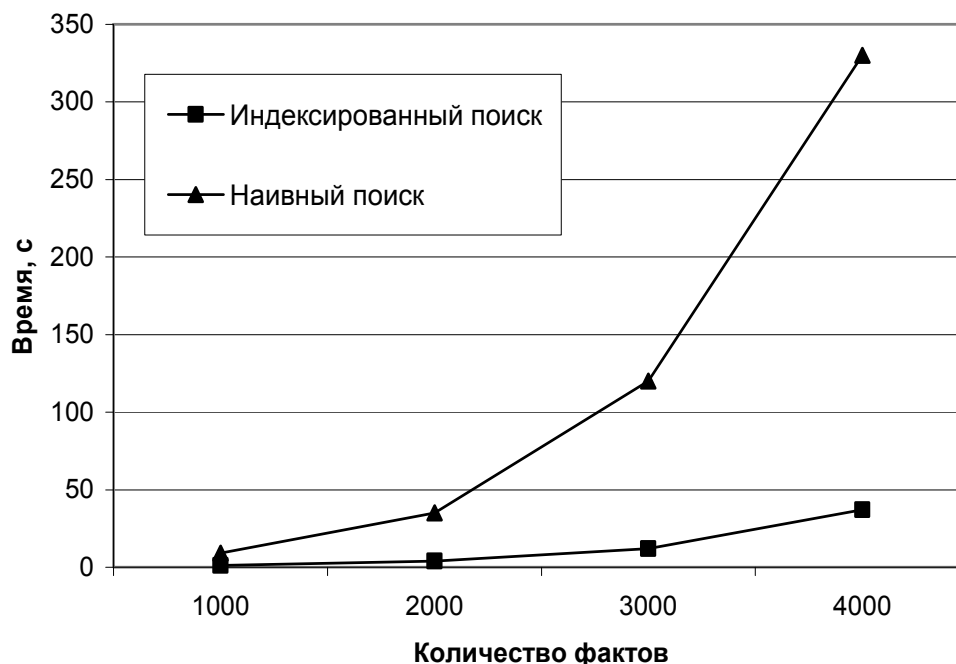


Рис. Зависимость времени логического вывода от числа фактов в базе знаний

Эксперименты на тестовом наборе фактов показали, что эффективность индексации фактов зависит также от результативности правил. Если количество успешных применений правил невелико, то время вывода с использованием индексов стремится к нулю. Если каждая комбинация исходных фактов приводит к успешной резолюции (ситуация, на практике, маловероятная), то использование индексов только увеличивает время за счет издержек на индексацию. В использованном примере успешными были приблизительно 15% всех применений правил.

Замена унификации правил операциями реляционной алгебры

В приведенном примере резолюции правила «субъект является мужчиной» был получен список значений переменной $?x$: $[sergey, nikita]$. Но поскольку эта переменная в правиле единственная, то задача практически уже решена, и новые факты могут быть получены подстановкой значений переменной в заголовок правила: $f(sergey, is_a, man)$; $f(nikita, is_a, man)$.

Рассмотрим, можно ли обойтись без интерпретации правил во втором примере. Извлечем из базы знаний факты $[8,9,10,11]$, а также $[12,13]$ и получим кортежи переменных $?x$ и $?y$ для первого условия и $?y$ и $?z$ для второго условия

$?x$	$?y$		$?y$	$?z$
<i>sergey</i>	<i>nikita</i>		<i>nikita</i>	<i>stepan</i>
<i>sergey</i>	<i>andrey</i>		<i>andrey</i>	<i>egor</i>
<i>natalia</i>	<i>nikita</i>			
<i>natalia</i>	<i>andrey</i>			

Выполнив для этих таблиц операцию соединения по совпадающим значениям $?y$, получим кортеж

$?x$	$?y$	$?z$
<i>sergey</i>	<i>nikita</i>	<i>Stepan</i>
<i>sergey</i>	<i>andrey</i>	<i>Egor</i>
<i>natalia</i>	<i>nikita</i>	<i>Stepan</i>
<i>natalia</i>	<i>andrey</i>	<i>Egor</i>

Результат операции соединения дает допустимые сочетания значений переменных $?x$ и $?z$, т.е. искомое решение:

$f(\text{sergey}, \text{grandparent}, \text{stepan})$.

$f(\text{sergey}, \text{grandparent}, \text{egor})$.

$f(\text{natalia}, \text{grandparent}, \text{stepan})$.

$f(\text{natalia}, \text{grandparent}, \text{egor})$.

В рамках данной работы не рассматриваются отношения отрицания в правилах, а также отношения сравнения, отличные от равенства, хотя интерпретация правил, содержащих отношения «больше», «меньше» и др., приводит всего лишь к модификации операций соединения. Очевидным недостатком является возможное чрезмерное разрастание индексов для отношений объект–свойство–значение в связи с тем, что количество возможных значений свойств обычно существенно больше количества сущностей. Здесь целесообразно использовать хеширование значений свойств, как это делается в некоторых модификациях алгоритма *Rete*.

Заключение

Предлагаемый метод существенно отличается от алгоритма *Rete*. Главное отличие заключается в том, что только индексация фактов выполняется заблаговременно, а операция предварительного отбора фактов выполняется непосредственно перед резолюцией. Это означает, что факты и правила могут существовать отдельно, в том числе на разных сетевых ресурсах, что соответствует концепции Глобальной семантической сети [4]. В ходе предварительного отбора фактов для каждого условия правила создаются множества кортежей значений переменных, используя которые, можно в отдельных случаях отказаться от резолюции правил, а, используя операции реляционной алгебры над кортежами, сразу получить множество решений правила. Тестирование метода индексации фактов показало его работоспособность и ускорение приблизительно на порядок по сравнению с «наивным» логическим выводом. Дальнейшее ускорение возможно путем полноценной реализации замены унификации правил операциями реляционной алгебры.

Литература

1. Рассел С., Норвиг П. Искусственный интеллект: Современный подход. 2-е изд. / пер. с англ. – М.: Изд. дом «Вильямс», 2006.
2. Forgy C.L. RETE: A fast algorithm for the many pattern / many object pattern match problem // Artificial Intelligence. – 1982. – Vol. 19. – P. 17–37.
3. Doorenbos R. B.. Production Matching for Large Learning Systems // PhD Theses. University of South California, 1995. – 208 pp.
4. Berners-Lee T., Hendler J., Lassila O. The Semantic Web // Scientific American Magazine. – May, 2001.

Бессмертный Игорь Александрович – Санкт-Петербургский государственный университет информационных технологий, механики и оптики, кандидат технических наук, доцент, igor_bessmertny@hotmail.com

УДК 519.254, 004.9

ПРОГНОЗИРОВАНИЕ СЛОЖНЫХ СИГНАЛОВ НА ОСНОВЕ ВЫДЕЛЕНИЯ ГРАНИЦЫ РЕАЛИЗАЦИЙ ДИНАМИЧЕСКИХ СИСТЕМ

И.Г. Сидоркина, А.В. Егошин, Д.С. Шумков, П.А.Кудрин

Предлагаются два метода прогнозирования сложного сигнала – на основе многослойного персептрона и метода опорных векторов (SVM), реализованные в программном комплексе. Обучающее множество для предиктора определяется на основе выделения границы реализаций динамических систем в сигнале.

Ключевые слова: прогнозирование, обнаружение разладки, фрактальность, динамические системы, нейронные сети, SVM.

Введение

Сложные цифровые сигналы часто встречаются в деятельности человека, примерами могут служить медицинские показания, сигналы с датчиков на производстве, финансовые и экономические показатели. Сигнал представляется временным рядом значений, наблюдаемых в последовательные моменты времени: $\mathbf{x}(\mathbf{t}) \equiv \{x_1, x_2, \dots, x_N\}$, где $\mathbf{t} \equiv \{t_1, t_2, \dots, t_N\}$, $x_i = \mathbf{x}(t_i)$, $i = 1 \dots N$, где N – число отсчетов в сигнале. Атрибутами сложного сигнала являются следующие характеристики (одна или несколько): нестационарность, непериодичность, функция распределения, наличие сингулярностей, фрактальность, стохастичность и хаотичность. Для решения задачи прогнозирования таких сигналов был предложен подход, реализованный в программном комплексе.

Реализация подхода включает:

1. выделение «стабильного» участка в конце временного ряда, на котором происходит обучение предиктора, учитывающего последнюю динамику исследуемого процесса;
2. выполнение прогнозирования исследуемого временного ряда с использованием методов на основе достижения порога изменения сигнала и SVM.

Использование двух независимых предикторов позволяет повысить стабильность прогнозирования и снизить ошибку.

Метод выделения границ реализации динамической системы

Для выделения границ реализации динамической системы предложена структурная модель, согласно которой сложный сигнал рассматривается в виде последовательности реализаций различных динамических систем (ДС), т.е.

$$\mathbf{x}(t+1) = \mathbf{F}_1(\mathbf{x}(t), \mathbf{x}(t-1), \dots, \mathbf{x}(t-k_1)), \quad t = 0, 1, \dots, r_1,$$

$$\mathbf{x}(t+1) = \mathbf{F}_2(\mathbf{x}(t), \mathbf{x}(t-1), \dots, \mathbf{x}(t-k_2)), \quad t = r_1 + 1, \dots, r_2,$$

...

$$\mathbf{x}(t+1) = \mathbf{F}_n(\mathbf{x}(t), \mathbf{x}(t-1), \dots, \mathbf{x}(t-k_n)), \quad t = r_n, \dots, N-1,$$

где $\mathbf{x}(t)$ – d -мерный вектор состояния системы в момент времени; t , n – число точек смены динамики; $\mathbf{F}_i$ – вектор-функции, определяющие следующее состояние системы в различные периоды времени и имеющие k_i последних значений $\mathbf{x}(t)$ в качестве параметров, $i = 1 \dots n$; r_i – точки смены динамики сигнала; N – число наблюдений сигнала.

Чтобы анализировать и использовать для обучения предиктора последний «стабильный» по некоторой мере фрагмент временного ряда сигнала, необходимо разделить сигнал на участки квазистационарности. Такая задача известна как задача обнаружения «разладки», т.е. изменения любых вероятностных свойств сигнала [1]. Однако для сложных сигналов такое определение может давать неточную границу истинной смены динамики сигнала.

ла. Поэтому предлагается основываться на свойстве, которым характеризуется сложный сигнал и для которого можно получить адекватные численные оценки, используя только наблюдаемый временной ряд. Как показали исследования, таким свойством является фрактальность, точнее, локальная фрактальная размерность – показатель фрактальной размерности для ограниченной области временного ряда сигнала.

В качестве оценки локальной фрактальной размерности предлагается использовать индекс фрактальности [2]. Обнаружение смены динамики сигнала происходит по изменению локальной фрактальной размерности в статистическом смысле, т.е. по наличию разладки в ряде. Это позволяет определять смену динамики сигнала как изменение степени однородности его фрактальной размерности.

Для обнаружения изменения вероятностных свойств производного сигнала использован апостериорный непараметрический алгоритм кумулятивных сумм (АКС) [3]. В результате проведенного исследования выявлено, что он эффективней по скорости и точности обнаружения для сложных стохастических сигналов, чем общий случай алгоритма Бродского–Дарховского [4] и метод на основе принципа минимума информационного рассогласования [5] (точнее в среднем на 5–8%).

Метод обнаружения границы реализаций динамических систем заключается в следующем. По исходному сигналу методом скользящего окна строится производный сигнал оценки локальной фрактальной размерности – индекса фрактальности. В производном и в исходном сигнале методом АКС находятся точки разладки. Положения точек сопоставляются, и берется среднее значение. Это и будет границей между реализациями различных ДС. Для сложных зашумленных стохастических сигналов такой метод обнаружения дает более точную оценку момента разладки (смены динамики), чем обнаружение по исходному ряду методом АКС, в среднем на 9–10%.

Общая последовательность прогнозирования сложного сигнала представляется следующим образом: получение сигнала, получение производного от исходного сигнала ряда локальной фрактальной размерности, поиск точки разладки в исходном сигнале и в ряде локальной фрактальной размерности – получение границы между реализациями различных ДС, выделение для обучения предиктора последнего фрагмента сигнала до найденной границы (фрагмент последней сформировавшейся квазистабильной динамики сигнала), построение комбинированного прогноза, оценка его достоверности. Указанная итерация может повторяться многократно для различных масштабов представления сигнала, чтобы получить многошаговый прогноз. Для прогнозирования сложных сигналов предложены два метода, которые можно объединять в один на основе комитета.

Прогнозирование на основе метода достижения порога изменения

В основу первого метода положено преобразование исходного сигнала в ряд времени достижения порога изменения p [6]. Оно позволяет, во-первых, агрегировать сигнал без потери существенной информации о нем, во-вторых, подойти к задаче прогнозирования с другой точки зрения, в частности, прогнозировать не следующее значение сигнала, а время, через которое изменение сигнала превысит известный порог. Прогнозирование осуществляется двухслойным перцептроном. Следовательно, исходный сигнал преобразуется следующим образом:

$$\begin{aligned} x(t) &\equiv \{x_1, x_2, \dots, x_N\} \rightarrow x' \equiv \{x'_1, x'_2, \dots, x'_{N'}\}, \\ x_i &= x(t_i), \quad i = 1 \dots N, \quad x'_j = x(t'_j), \quad j = 1 \dots N', \\ |(x'_i - x'_{i-1}) / x'_{i-1}| &\geq p, \quad p > 0, \quad i = 2, 3, \dots, N', \\ \tau &= \{\tau_1, \tau_2, \dots, \tau_{N'-1}\}, \quad \tau_i = t'_{i+1} - t'_i, \quad i = 1 \dots N' - 1, \end{aligned}$$

где N – число отсчетов в исходном сигнале, x' – преобразованный сигнал, где оставлены только те значения, относительная разница между которыми превышает порог p , N' – число отсчетов в преобразованном сигнале, τ – ряд времени достижения заданного порога изменения, где каждое значение означает время, которое потребовалось сигналу для того, чтобы превысить порог изменения p . Кроме того, такой метод позволяет выполнять комбинированное прогнозирование, которое включает в себя:

- 1) оценку по x' следующего значения $x'_{N'+1}$;
- 2) оценку по τ следующего значения $\tau_{N'+1}$;
- 3) сравнение $\text{sign}(\tau_{N'+1})$ и $\text{sign}(x'_{N'+1} - x'_{N'})$.

Сопоставление результатов, выполненных по рядам x' и τ , позволяет увеличить степень надежности прогноза, так как совпадение знаков изменения прогнозируемых величин говорит о высокой степени его достоверности.

Прогнозирование с использованием метода опорных векторов (SVM)

В качестве второго подхода, реализованного в разработанном программном комплексе, используется метод опорных векторов (SVM) [7]. Особенностью предложенной реализации является оригинальный метод обучения и подбора входных параметров SVM-модели, который базируется на поиске «похожих» паттернов в исходном временном ряду. Это позволяет учитывать текущее состояние исследуемой динамической системы. Недостатком известных методов обучения SVM является то, что область, используемая для тестирования модели, выбирается случайным образом или в конце временного ряда и не учитывает текущее состояние системы. Для решения этой проблемы предложен оригинальный метод обучения и подбора входных параметров SVM-модели.

В предложенном методе выделяются два этапа. Первый этап дает «грубую» оценку близ оптимальных значений подбора входных параметров путем разбиения исходной выборки на n частей и независимого обучения SVM-моделей при различных значениях оптимизируемого параметра (например, C). При этом остальные входные параметры остаются постоянными. Для каждого участка вычисляется нормализованная среднеквадратическая ошибка E и сложность модели C_M . Далее из полученного набора альтернативных SVM-моделей выбирается m наилучших с помощью критерия эффективности Eff_M , позволяющего сравнивать альтернативные модели в зависимости от ошибок, которые они дают на тестовых данных, и их сложностью. Одной из особенностей данного метода является то, что он может быть легко реализован в алгоритмах параллельной обработки данных, потому что обучение осуществляется на различных участках с разными значениями входных параметров.

Второй этап связан с поиском такой SVM-модели, которая отражала бы не только глобальную динамику процесса, но и учитывала бы и локальную. Для этого тестирование модели происходит на тех участках временного ряда, которые имеют схожую динамику с текущим состоянием. В качестве оценки схожести используется оценка косинуса угла между двумя векторами через их скалярное произведение. Чем ближе значение стремится к единице, тем более схожими считаются вектора. Результирующей моделью выбирается та, у которой нормализованная среднеквадратическая ошибка минимальна как на выборке, участвовавшей в обучении, так и на похожих участках с учетом ее сложности.

В качестве проверки эффективности модели предложен следующий критерий Eff_M . Пусть имеется две альтернативные модели M_1 и M_2 . Тогда $Eff_{M_1} \geq Eff_{M_2}$, если

$\frac{E_1}{C_{M_2}} - \frac{E_2}{C_{M_1}} \leq 0$, где C_M – сложность модели, рассчитанная таким образом, что

$C_M = \frac{N_{SV}}{N}$ (N_{SV} – количество опорных векторов в SVM-модели, N – общее количество векторов при обучении); E – нормализованная среднеквадратическая ошибка, рассчитанная по формуле $E = \frac{1}{\sigma^2 n} \sum_{i=1}^n (x - x')^2$ (σ^2 – несмещенная дисперсия реальных значений временного ряда x , x' – прогнозные значения, n – количество тестовых данных, на которых проверяется модель).

Результаты вычислительных экспериментов

Результаты прогнозирования времени достижения порога изменения сигнала представлены в таблице.

№	Ряд	$p(\%)$	s	$E_{cp}(\%)$	Знак
1a	Авторегрессия с разладкой ($x_0 = 0,02$): 1. $x_i = 0,9 * x_{i-1} + 0,1 + \varepsilon * 1,5$ (3000 отсчетов) 2. $x_i = 0,8 * x_{i-1} + 0,105 + \varepsilon * 1,45$ (2000 отсчетов)	100	2	68,5	+
1b				10,1	+
1c				9,8	+
2a	Посещаемость веб-сервиса – возвраты пользователей с 21.05.08 по 06.06.09 (суточные значения)	20	5	91,54	+
2b				11,32	+
2c				12,1	+
3a	Курс акции «Лукойл» с 01.01.09 по 18.06.09 (часовые значения)	1	3	10,25	+/-
3b				74,58	+
3c				51,23	+

Таблица. Результаты прогнозирования времени достижения порога изменения сигнала

В качестве предиктора использовалась нейронная сеть типа «многослойный персептрон» структуры 3–2–1 для скрытых слоев и линейной задержки на входе на 10. В качестве входных переменных используются время достижения заданного порога изменения и значение сигнала в данной точке. Проводилось 3 подхода на каждый прогноз, за результат прогноза бралось среднее предсказанное значение.

При использовании метода опорных векторов было выбрано гауссовское ядро. В качестве оптимизируемых параметров использованы C и γ , которые подбирались с учетом предложенного выше метода.

В таблице приняты следующие обозначения: s – на сколько отсчетов вперед получился прогноз; E_{cp} – средняя относительная ошибка прогноза; **Знак** – верно или нет определен знак изменения сигнала (+ верно, – нет, +/- часть прогнозов дала неверный знак); ε – случайное число от –1 до 1 с равномерным распределением.

Авторегрессия с разладкой есть двухкомпонентный сигнал, состоящий из двух «склеенных» фрагментов. Каждый фрагмент задается определенным отображением (формулы указаны в таблице). Прогноз под индексом «а» строился на основе всего дос-

тупного множества нейронной сетью, под «b» – с учетом обнаруженной разладки во фрактальной стабильности сигнала нейронной сетью, под «c» – с учетом обнаруженной разладки во фрактальной стабильности сигнала методом опорных векторов.

Выводы

Основным выводом является то, что использование для обучения предиктора последнего фрагмента сигнала с квазистабильной локальной фрактальной размерностью вместо эмпирического выбора обучающего множества дает более точный или надежный прогноз. Программный комплекс, реализующий предложенные методы, может найти свое применение в медицине и биологии, контроле технологических процессов и финансово-экономической деятельности, энергетике и приборостроении – везде, где ведется работа со сложными эволюционными сигналами.

В дальнейшем планируется исследование объединения двух аппаратов прогнозирования, нейронных сетей и метода опорных векторов, в качестве единого предиктора. Есть основания предполагать, что использование нескольких независимых предикторов позволит обеспечить более качественный прогноз (по точности и надежности).

Литература

1. Darkhovski B.S. Nonparametric methods in change-point problems a general approach and some concrete algorithms // IMS Lecture Notes – Monograph Series, Volume 23, 1994. [Электронный ресурс]. – Режим доступа: http://projecteuclid.org/DPubS/Repository/1.0/Disseminate?view=body&id=pdf_1&handle=euclid.lnms/1215463117 (дата обращения 22.10.09).
2. Старченко Н. В. Индекс фрактальности и локальный анализ хаотических временных рядов. Дис. ... канд. физ.-мат. наук: 05.13.18 : защищена 15.02.06. – М. 2005. – С. 34–64. [Электронный ресурс]. – Режим доступа: http://www.mirkin.ru/_docs/kon_diser/diserstarchenko.pdf (дата обращения 26.10.09).
3. Lajos Horvath. Ratio tests for change point detection // IMS Collections, Beyond Parametrics in Interdisciplinary Research: Festschrift in Honor of Professor Pranab K. Sen. – Institute of Mathematical Statistics. – 2008. – Vol. 1. – P. 293–304.
4. Бродский Б.Е., Дарховский Б.С. Асимптотический анализ некоторых оценок в апостериорной задаче о разладке // Теория вероятностей и ее применения. – 1990. – Т. 35. – № 3. – С. 551–557.
5. Акатьев Д.Ю., Савченко В.В. Обнаружение разладки случайного процесса на основе принципа минимума информационного рассогласования // Автотметрия. – 2005. – Т. 41. – № 2. – С. 68–74.
6. Егошин А.В. Анализ времени достижения сигналом порога изменения в задаче нейросетевого прогнозирования временных рядов // Информационно-вычислительные технологии и их приложения: сборник статей VI Международной научно-технической конференции. – Пенза: РИО ПГСХА, 2007. – С. 74–76.
7. Burges C.J.C. A tutorial on Support Vector Machines for classification// Data Mining and Knowledge Discovery. – 1998. – № 2. – P. 121–167.

<i>Сидоркина Ирина Геннадьевна</i>	– Марийский государственный технический университет, доктор технических наук, профессор, декан, dekan_fivt@mail.ru
<i>Егошин Алексей Валерьевич</i>	– Марийский государственный технический университет, аспирант, g12@bk.ru
<i>Шумков Дмитрий Сергеевич</i>	– Марийский государственный технический университет, аспирант, dimavova@mail.ru

УДК 621.397.13.037.372

КЛАССИФИКАЦИЯ НЕСТАЦИОНАРНЫХ ИЗОБРАЖЕНИЙ И РАЗРАБОТКА МЕТОДИКИ ОЦЕНКИ АЛГОРИТМОВ КОДИРОВАНИЯ ИСТОЧНИКА

И.А. Зубакин, Ш.С. Фахми

Предложен алгоритм оценки эффективности различных методов сжатия изображений (методов кодирования источника) на основе введенной меры нестационарности. Сформирована методика оценки эффективности, использующая компактную представительную выборку сюжетов, ранжированных по нестационарности и широкополосности.

Ключевые слова: мера нестационарности, широкополосность, классификация изображений, компактная представительная выборка сюжетов, алгоритм вычисления статистик.

Введение

Разработчики стандартных алгоритмов кодирования источника стремятся создать универсальный алгоритм, обеспечивающий наилучшее субъективное качество передачи сигнала при его неизвестных и переменных статистических свойствах в пространстве и времени и при неформализованных (скрыто учитываемых) ограничениях скорости передачи и сложности кодирующих устройств. Несмотря на то, что создатели метода кодирования источника на основе вейвлет-преобразования отмечают [1], что оно дает хорошие результаты при сжатии нестационарных сигналов, почти все подобные исследования не учитывают в явном виде изменения статистик сигналов в пространстве и времени при постановке задачи, и оговорки о скорости изменения сигнала в пространстве и времени в лучшем случае всплывают где-то в середине или даже в конце изложения. При этом каждый исследователь считает свой вклад в кодирование телевизионных сигналов шагом к такому универсальному кодированию, качество которого определяется экспертной оценкой [2]. Вместе с тем представляется перспективным применение известных результатов теории передачи информации при разработке и объективной оценке эффективности кодеков цифрового телевидения не только в части кодирования канала, но и в части кодирования источника, сигнал которого априори известен приблизительно и нестационарен [3].

Методика оценки алгоритмов кодирования источников

Основные проблемы, стоящие перед разработчиками систем кодирования источника в цифровом вещательном телевидении, для которого проблема зашумленности источника не очень актуальна в силу доступности контента с большим отношением сигнал/шум – это:

- проблема априорной неопределенности статистик сигнала и помех (в силу необозримого разнообразия статистик эта проблема решена лишь для ряда простых наборов априорной информации, и главным достижением теории является выделение двух направлений разработки – устойчивых алгоритмов и адаптивных алгоритмов) [4];
- проблема нестационарности кодируемых сигналов (в силу необозримого разнообразия статистик эта проблема, как и проблема априорной неопределенности, решена лишь для ряда простых наборов нестационарности, и главным достижением теории является выделение класса кусочно-стационарных сигналов и алгоритмов обнаружения интервалов стационарности или сегментации сигналов и изображений) [5];
- проблема сложности кодирования состоит в формализации связи трех величин – точности передачи, скорости передачи и сложности кодера (декодера); решение этой проблемы особо актуально при создании видеосистем на кристалле ограничен-

ной площади, но для которой известны решения лишь для некоторых типов кодирования [3].

Ниже акцент делается на проблеме нестационарности сигналов, без учета сложности алгоритма кодирования источника и его аппаратурной реализации. Разработчики алгоритмов кодирования проверяют эффективность их работы на некоторой выборке различных сюжетов, негласно полагая, что эта выборка является представительной. Так как речь идет о кодировании нестационарных сигналов, то для сравнения эффективности кодирования различных алгоритмов предлагается использовать следующую методику:

1. ввести меру нестационарности сигналов изображений;
2. разработать алгоритм вычисления введенной меры;
3. определить границы изменения этой меры для очень большой выборки сигналов (заведомо представительной);
4. произвести разбиение интервала изменения этой меры на несколько (по возможности равных) интервалов, количество которых должно быть достаточным для уверенности в представительности выборки;
5. отобрать изображения, принадлежащие к выделенным интервалам меры нестационарности;
6. провести кодирование сформированной выборки каждым из исследуемых алгоритмов;
7. для каждого алгоритма и для каждого элемента выборки оценить значения скорости передачи информации при заданной ошибке передачи (иногда возможна инверсия – оценить ошибку передачи при заданной скорости передачи);
8. при отсутствии статистических данных о вероятностях появления сюжетов с различной мерой нестационарности считать их равновероятными;
9. для каждого алгоритма усреднить по сформированной выборке значения скорости передачи при заданной ошибке (или усреднить ошибку передачи при заданной скорости передачи);
10. сравнить результаты усреднения и выбрать алгоритм, обеспечивающий наименьшую скорость передачи при заданной ошибке (наименьшую ошибку при заданной скорости передачи).

Для реализации предлагаемой методики ключевыми вопросами становятся первые два пункта методики: выбор меры нестационарности сигналов и разработка алгоритма вычисления этой меры.

Меры нестационарности и широкополосности

Для оценки нестационарности сигналов изображений представляется целесообразным воспользоваться методологией статистической радиотехники, в которой используются понятия стационарности в широком и узком смысле, т. е. различают процессы с постоянной/переменной дисперсией и с постоянной/переменной автокорреляционной функцией. Это подсказывает, что обобщенная мера нестационарности должна учитывать совокупное изменение различных статистик S_n сигнала изображения. Это могут быть математические ожидания, дисперсии, интервалы (или коэффициенты) корреляции яркости и цвета и др. Каждая из статистик вычисляется для некоторых интервалов: это могут быть оцененные по отдельным адаптивным алгоритмам интервалы стационарности или априори компромиссно выбранные интервалы. Они должны быть не слишком малыми для получения приемлемых методологических ошибок вычисления статистик и не слишком большими для уменьшения вероятности изменения статистик внутри интервала. Можно считать, что именно такие же соображения лежали в основе выбора окна (8×8) элементов при разработке стандартов MPEG.

Общая мера нестационарности M должна вычисляться как сумма (в общем случае взвешенная) мер нестационарности отдельных статистик [6]:

$$M = \sum_{n=1}^K M_n.$$

В статье [7] рассмотрена процедура выделения участков нестационарности, основанная на дискриминантной функции. Здесь же для оценки меры нестационарности каждой статистики сигнала предлагается использовать нормированную статистику, вычисляемую аналогично среднеквадратической полосе частот [8] с тем лишь различием, что в формулу должны входить не сами сигналы, а их статистики S_n , вычисляемые отдельно по каждому аргументу. Ниже для упрощения записан один обобщенный дискретный аргумент $i \in [1, N]$. Формула для непрерывных функций непрерывного аргумента включает отношение интегралов от квадрата производной сигнала к интегралу от квадрата самого сигнала. Для дискретных значений статистик можно использовать не производную, а первую конечную разность [6]:

$$M_n = \frac{\sum_{i=1}^{N-1} (S_{n(i+1)} - S_{ni})^2}{\sum_{i=1}^N S_{ni}^2}.$$

Алгоритмы вычисления меры нестационарности можно разбить на два класса алгоритмов, соответствующих детерминированному и адаптивному выбору интервалов усреднения статистик. Адаптивные алгоритмы (например, входящие в состав пакета *Photoshop*) дают оценки меры нестационарности M с точностью, близкой к потенциальной (характеризуются неоднозначностью результатов в зависимости от назначаемого допустимого отклонения оцениваемой статистики от среднего значения в данном интервале стационарности) и требуют больших вычислительных затрат. Детерминированные алгоритмы оценки многомерных сигналов на порядки проще адаптивных, хотя и обеспечивают несколько большую погрешность оценки меры нестационарности M .

Одним из методов учета нестационарности сигнала при кодировании изображений может быть предварительное выделение зон стационарности, аналогичное применяемому в *Photoshop*. Однако реальные изображения имеют произвольные формы зон стационарности, и передача формы этих зон потребует существенного увеличения сложности кодера и пропускной способности канала связи. Поэтому на практике имеет смысл остановиться на приближенной оценке зон стационарности с точностью до прямоугольных областей (для двумерных сигналов) или параллелепипедных областей для трехмерных сигналов с пропорциональными сторонами. С точки зрения экономии вычислительной сложности перспективными являются размеры таких областей, выражаемые целыми положительными степенями числа 2.

В данной статье для вычисления меры нестационарности рассматривается совокупность статистических характеристик (среднее значение m_i , дисперсия D_i и коэффициент корреляции r_i) исходного изображения, вычисляемые в пределах окна (8×8):

$$M = M_1 + M_2 + M_3 = \frac{\sum_{i=1}^{N-1} (m_{i+1} - m_i)^2}{\sum_{i=1}^N m_i^2} + \frac{\sum_{i=1}^{N-1} (D_{i+1} - D_i)^2}{\sum_{i=1}^N D_i^2} + \frac{\sum_{i=1}^{N-1} (r_{i+1} - r_i)^2}{\sum_{i=1}^N r_i^2}.$$

Важной характеристикой эффективности кодирования изображений различными алгоритмами является скорость передачи информации. В первом приближении она оп-

ределяется полосой частот, занимаемой сигналом, поэтому очевидно, что характеризовать изображение только лишь коэффициентом нестационарности – значит не учитывать скорость передачи информации. В связи с этим, наряду с коэффициентом нестационарности, будем учитывать также полосу частот, занимаемую сигналом изображения. Для оценки полосы частот нестационарного сигнала изображений будем использовать формулу, широко используемую в статистической радиотехнике в качестве меры широкополосности [8]:

$$W = \sqrt{\frac{\sum_{k=1}^{N/2} k^2 |\bar{X}(k)|^2}{\sum_{k=1}^{N/2} |\bar{X}(k)|^2}},$$

где $|\bar{X}(k)|$ – значение спектральной составляющей, полученное путем усреднения спектральных составляющих по всем строкам изображения.

В ходе эксперимента исследовалось большое количество изображений и оценивались параметры M и W . Путем ранжирования всех изображений по нестационарности и широкополосности были выбраны девять изображений, которые составляют компактную представительную выборку сюжетов. Такой подход позволяет разделять все изображения на изображения с малым, средним и большим значениями нестационарности и широкополосности (см. рис.).

В общем случае компактная представительная выборка видеосюжетов должна включать различные сочетания широкополосности и нестационарности по пространству и по времени. В данной статье для компактности изложения приведена лишь методика сравнения кодеров отдельных кадров изображения при использовании сюжетов, отличающихся шириной полосы пространственных частот и пространственной нестационарностью.

Введение меры нестационарности изображений, наряду с занимаемой изображением полосой частот, может быть использовано в задачах классификации изображений на основе их ранжирования по нестационарности и широкополосности.

Задача классификации изображений стоит во многих отраслях прикладной телевизионной техники – от космических систем землепользования и авиационной разведки до гидроакустики и медицины. В терминах инженерного подхода, сформированного в рамках направления «компьютерное зрение», эта задача формулируется как задача сегментации изображений на области с различной текстурой [2]. При этом классификацию текстур относят к алгоритмам, основанным на вычислении статистик и формировании оценок, противопоставляя алгоритмам более сложного класса задач распознавания, основанным на логическом и структурном анализе. В терминах статистической теории связи эти задачи относятся к многоальтернативной проверке гипотез при априорной информации о сигналах, включающей оценки некоторых статистик случайных полей – функции плотности вероятности, корреляционной функции и т. п. [5].

Введенная мера нестационарности может быть использована при формировании блоков из различных кадров видеопотока для последующего адаптивного кодирования с помощью трехмерного дискретного косинусного преобразования, широко используемого сегодня при построении различных алгоритмов сжатия изображений. Выше было отмечено, что для вычисления меры нестационарности может быть использован адаптивный подход формирования областей стационарности. Для формирования таких областей может быть использована методология, основанная на пирамидально-рекурсивном методе кодирования [3, 9].

		
$W = 8, M = 1,00$	$W = 10, M = 0,80$	$W = 34, M = 1,10$
		
$W = 8, M = 0,66$	$W = 13, M = 0,69$	$W = 15, M = 0,50$
		
$W = 9, M = 0,38$	$W = 13, M = 0,42$	$W = 35, M = 0,19$

Рис. Компактная представительная выборка сюжетов

Заключение

Разработана методика объективной оценки эффективности алгоритмов кодирования изображений, основанная на введенном понятии меры нестационарности. Введенная мера нестационарности может быть применена для оценки сигналов различного вида, например, для кусочно-стационарных процессов или с монотонно изменяющимися статистиками (например, винеровского процесса, определяемого как интеграл от белого шума, и имеющего линейно возрастающую дисперсию). Данная методика позволяет сравнивать различные алгоритмы кодирования изображений и выбирать наилучший.

Введенная мера нестационарности позволяет, наряду с мерой широкополосности, классифицировать различные изображения. Сформирована компактная представительная выборка сюжетов, содержащая девять изображений с различными значениями меры нестационарности и широкополосности, на основе которой может быть определена эффективность кодирования изображений различными алгоритмами.

Литература

1. Сэломон Д. Сжатие данных, изображений и звука. – М.: Техносфера, 2006.
2. Прэтт У. Цифровая обработка изображений. – М.: Мир, 1982, кн. 1.

3. Твердотельная революция в телевидении / Под ред. Умбиталиева А.А. и Цыцулина А.К. – М.: Радио и связь, 2006.
4. Репин В.Г., Тартаковский Г.П. Статический синтез при априорной неопределенности и адаптация информационных систем. – М.: Сов. радио, 1977.
5. Миленький А.В. Классификация сигналов в условиях неопределенности. – М.: Сов. радио, 1975.
6. Цыцулин А.К., Фахми Ш.С., Зубакин И.А. Оценка эффективности кодирования нестационарных изображений // Телевидение: передача и обработка изображений. 7-я междунар. конф., 29–30 июня 2009, СПб. – С. 19–22.
7. Парфенов В.И., Сергеева Е.В. Применение дискриминантной процедуры при синтезе и анализе телекоммуникационной системы, основанной на манипуляции статистическими характеристиками случайного процесса // Вестник ВГУ. – Сер. Физика, математика. – 2008. – Вып. 1. – С. 70–74.
8. Вайнштейн Л.А., Зубаков В.Д. Выделение сигналов на фоне случайных помех. – М.: Сов. радио, 1960.
9. Фахми Ш.С. Пирамидально-рекурсивный метод кодирования и декодирования видеоинформации // Вопросы радиоэлектроники, сер. Техника телевидения. – 2007. – Вып. 2. – С. 43–51.

Зубакин Игорь Александрович – Санкт-Петербургский государственный электротехнический университет «ЛЭТИ», аспирант, zubakin_@mail.ru
Фахми Шакиб Субхиевич – Санкт-Петербургский государственный электротехнический университет «ЛЭТИ», кандидат технических наук, доцент, shakeebf@mail.ru

УДК 681.78

ПРОЦЕДУРА ПРОВЕДЕНИЯ ИЗМЕРЕНИЙ ОБЪЕКТОВ ПРИРОДНОЙ СРЕДЫ МЕТОДОМ ГАЗОРАЗРЯДНОЙ ВИЗУАЛИЗАЦИИ

К.Г. Коротков, Д.В. Орлов, Е.Н. Величко, Ю.Ю. Гатчина

Работа направлена на создание процедуры проведения измерений объектов природной среды на оборудовании газоразрядной визуализации. В ходе исследований были выявлены дестабилизирующие факторы измерений и найдены пути их устранения. Предложена процедура проведения измерений, обеспечивающая наиболее стабильные результаты.

Ключевые слова: ГРВ, процедура измерений, природная среда.

Введение

Принципы регистрации цифровых газоразрядных изображений с последующей их компьютерной обработкой и анализом были развиты в работах [1–4]. На этом принципе были созданы различные модификации приборов газоразрядной визуализации (ГРВ). Эти приборы используются для экспресс-диагностики состояния здоровья людей [1, 3], а также для исследования объектов природной среды [1, 4]. С помощью ГРВ-оборудования можно зарегистрировать динамику изменения исследуемого объекта или объектов природной среды, но нельзя оценить абсолютное значение параметров исследуемого объекта, например, электрической емкости или другой физической характеристики.

На получаемые во время проведения измерений значения рассчитываемых параметров может влиять множество факторов, большая часть которых не представляют интереса для исследования и, более того, увеличивают погрешность измерений. К тому же для проведения подобных исследований до сих пор не было сформулировано науч-

но обоснованных методических правил проведения измерений. Это приводило к проблемам при трактовке, статистической обработке и обосновании достоверности полученных результатов.

Целью данной работы является выявление дестабилизирующих факторов, оказывающих влияние на получаемые результаты, выработка путей их устранения и разработка процедуры проведения измерений и обработки получаемых данных.

Экспериментальная установка

Процедура формирования ГРВ-изображений с помощью прибора ГРВ заключается в следующем. Исследуемый объект помещается на кварцевый электрод, на обратную сторону которого нанесено прозрачное токопроводящее покрытие, на которое в течение заданного времени подаются импульсы напряжения от специализированного генератора. Мощность импульсов и длительность воздействия задаются программно в контуре управления. При высокой напряженности поля в пространстве между объектом и пластиной развивается лавинный и/или скользящий газовый разряд, характеристики которого определяются свойствами объекта. Пространственное распределение разряда фиксируется специализированной видеокамерой (на базе ПЗС-матрицы), расположенной непосредственно под прозрачным электродом. Видеопреобразователь осуществляет оцифровку изображения и передачу его в компьютер для дальнейшей обработки. ГРВ-граммы обрабатываются в специально разработанном программном комплексе, где осуществляется расчет параметров изображений, таких как площадь и средняя интенсивность разряда. Эти параметры наиболее сильно коррелируют с физическими характеристиками исследуемого объекта, например, электрической емкостью и сопротивлением [5].

В данной работе использовались приборы ГРВ серии «ГРВ Компакт» с аналоговой видеокамерой (www.kti.spb.ru). Приборы ГРВ с цифровыми камерами требуют более длительного времени разогрева при проведении прецизионных исследований по сравнению с аналоговыми камерами [5]. Эксперименты производились при стандартных условиях: температура воздуха 21–25 °С, относительная влажность 26–30%.

Для проведения измерений на электрод помещался металлический цилиндр (1 см в диаметре и высотой 2 см) – тест-объект, который, в свою очередь, подключался к платиновому электроду, помещенному в сосуд с жидкостью, или к металлической антенне, расположенной в исследуемой области. Антенна используется для исследования газообразных сред. Принципиальная схема экспериментальной установки подробно описывалась в статье [5]. Обработка данных производилась в специально разработанном программном обеспечении «GDV Scientific Laboratory» производства компании «Kirlionics Technologies International»[®].

Результаты

Основные дестабилизирующие факторы

При проведении любых измерений методом газоразрядной визуализации на получаемые результаты оказывают влияние различные факторы, которые мы называли дестабилизирующими. Среди этих факторов можно выделить общие, т.е. относящиеся к измерениям на ГРВ-приборах в целом, и частные, т.е. относящиеся только к измерениям с применением определенной экспериментальной установки, например, измерения с шприцом. Далее приведены общие дестабилизирующие факторы. Полный список с подробным описанием приведен в работе [6].

Плохое заземление. При отсутствии надлежащего заземления прибора и компьютера, подключенного к нему, происходит накопление на них статического заряда, что с течением времени может приводить к разрядке накопившегося заряда и, при длитель-

ных сериях измерений, иногда приводит к самопроизвольному выключению или перезагрузке компьютера. Также при отсутствии заземления сказывается влияние перепадов напряжения в сети, обусловленных другими электрическими приборами, подключенными в одну сеть с прибором ГРВ. Эти факторы влияют на стабильность газового разряда, а, следовательно, и получаемых ГРВ-грамм (рис. 1).

Заземление прибора может помочь избежать влияния других электрических приборов, подключенных в одну сеть с прибором ГРВ и компьютером. От этого влияния можно избавиться альтернативным способом: подключить прибор ГРВ к автономному источнику питания с выходным напряжением 12 В, например, автомобильному аккумулятору, а в качестве компьютера использовать ноутбук, работающий от собственного аккумулятора.

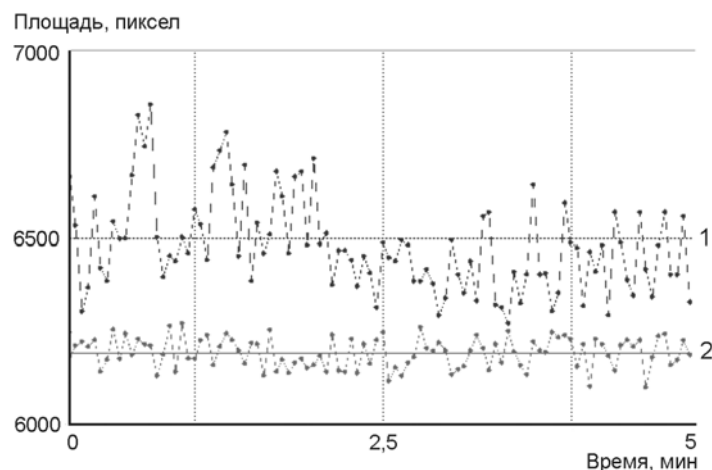


Рис. 1. Влияние заземления на разброс значений площади ГРВ-грамм:
1 – съемка без заземления; 2 – съемка с заземлением

Недостаточная вентиляция. При использовании стандартной для ГРВ-приборов крышки, закрывающей электрод, не происходит необходимой вентиляции воздушного промежутка между крышкой и электродом, в результате чего происходит накопление озона при съемке длительных серий. Накопление озона в значительной степени влияет на получаемые ГРВ-граммы и их параметры.

Исследования показали, что в отсутствии вентиляции в промежутке между крышкой и прозрачной пластиной накапливается озон, что значительно снижает площадь разряда.

Разогрев прибора. Прибору ГРВ требуется некоторое время для выхода на режим работы, при котором параметры ГРВ-грамм наиболее стабильны. Для исследованного нами образца прибора «ГРВ Компакт» стабильному уровню работы соответствует разброс значений средней интенсивности, не превышающий 1,5% от среднего значения в серии, а по площади — 3%. Явление «разогрева» прибора заключается в постепенном, плавном увеличении значения средней интенсивности и уменьшении площади газового разряда после включения (рис. 2).

Для выхода на стабильный режим по средней интенсивности ГРВ-грамм необходимо произвести разогрев один раз после включения прибора.

Однако, поскольку между сериями измерений во время проведения эксперимента зачастую прибор простаивает, то он успевает остыть. По этой причине в начале каждой серии измерений возникает небольшой спад площади свечения на протяжении первых 20–40 разрядов. Если время простоя прибора ГРВ превышает 1 час, то лучше повторно произвести разогревочную серию длительностью 300 кадров.

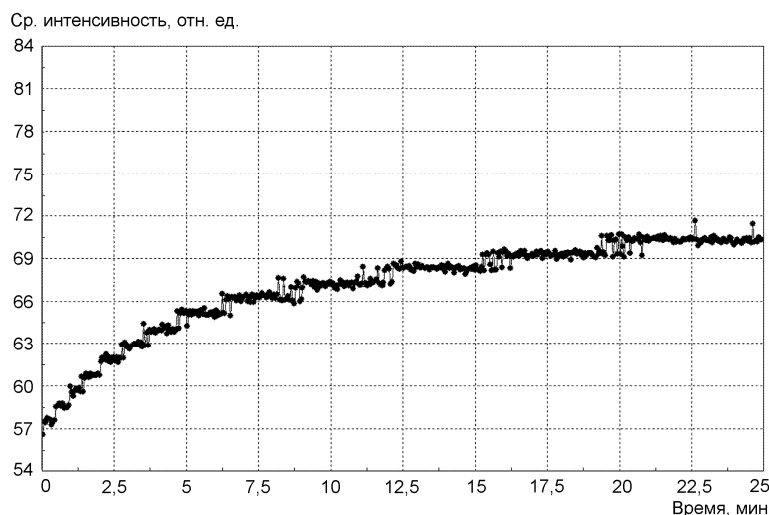


Рис. 2. Явление «разогрева» ГРВ-камеры

Смещение тест-объекта. Одним из дестабилизирующих факторов является изменение положения тест-объекта на стеклянном электроде. Перемещение или переустановка тест-объекта приводит к изменению площади его соприкосновения с электродом. Это выражается в изменении площади свечения или появлении засветки внутри круга, описывающего тест-объект на ГРВ-грамме. Данная засветка появляется из-за неплотного прилегания в некоторых местах поверхности тест-объекта к поверхности стеклянного электрода прибора. Эти явления значительно увеличивают разброс значений параметров ГРВ-грамм внутри серий и между сериями, что снижает стабильность получаемых результатов и не позволяет произвести нормальную статистическую обработку результатов.

Изменение условий окружающей среды. Условия окружающей среды, например, влажность воздуха, температура, электромагнитный фон, оказывают влияние на развитие газового разряда. Состояние воздуха в экспериментальной лаборатории влияет на газовый разряд, поскольку он развивается в окружающем воздухе. Например, включение какого-либо электрического прибора в помещении, в котором производятся измерения на приборе ГРВ, может привести к изменению электромагнитного фона, что, в свою очередь, скажется на получаемых ГРВ-граммах.

Установить влияние каждого из параметров окружающей среды на параметры ГРВ-грамм представляется практически невозможным, поскольку для этого их надо брать по отдельности. Поэтому мы дадим определение одинаковым условиям от противного. Одинаковыми условиями проведения измерений на приборах ГРВ считаются такие, при которых средние значения параметров ГРВ-грамм серий лежат в пределах $\pm 2\%$ по площади и $\pm 1,5\%$ по средней интенсивности. Эти пороговые значения универсальны для проведения любых измерений на ГРВ-приборах в закрытом помещении, однако эти рамки могут расширяться при проведении измерений на открытом воздухе.

Интервал между разрядами. На приборах «ГРВ Компакт» и «ГРВ Камера» минимально возможный интервал между двумя последовательными разрядами в серии составляет 3 с. Этот порог задан программным путем, и изменить его нельзя, поскольку он обусловлен электрическими процессами, происходящими в приборе ГРВ (например, перезарядка конденсатора).

Эксперименты показали, что при временном интервале 3 с между разрядами разброс значений параметров площади и средней интенсивности больше, чем при интервале 5 с. Дальнейшее увеличение интервала между разрядами не приводит к увеличению стабильности получаемых результатов и уменьшению разброса внутри серии.

Протирка электрода. Если между сериями измерений стеклянный электрод прибора ГРВ протереть тканью, то значения параметров ГРВ-грамм в следующей серии изменятся. Это происходит по причине изменения заряда, накопившегося на электроде (рис. 3). Как видно из рисунка, кроме изменения среднего значения площади ГРВ-грамм, произошло сильное увеличение разброса значений внутри серии. Это выразилось в двукратном увеличении среднеквадратического отклонения (изменение высоты прямоугольников) и доверительного интервала («ушки», расположенные сверху и снизу от прямоугольника). То же самое происходит и с параметром средней интенсивности.

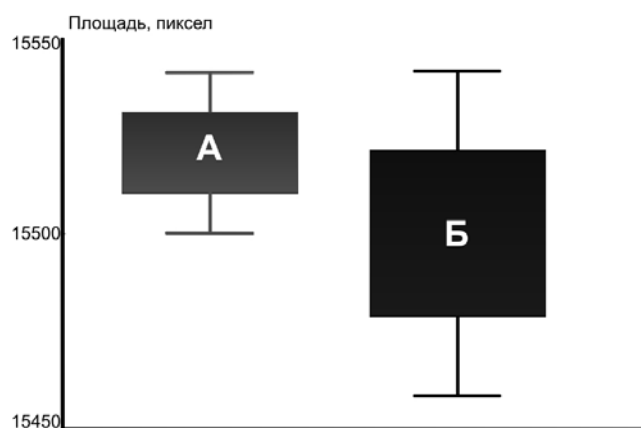


Рис. 3. Влияние протирки электрода сухой тканью на значения площади ГРВ-грамм в сериях: А – серия до протирки; Б – серия после протирки

Засветка изображения. Во время измерений в хорошо освещенных помещениях или на открытом воздухе может происходить засветка получаемого изображения из-за проникновения света в щели между защитной крышкой электрода, держателем тест-объекта или шприца. Подобная засветка будет влиять на параметры ГРВ-грамм и вносить дополнительную погрешность в измерения.

Общие рекомендации

Исходя из перечисленных дестабилизирующих факторов, были сформулированы общие рекомендации проведения прецизионных измерений на приборах ГРВ.

1. Обеспечить заземление прибора ГРВ и компьютера.
2. Во время проведения длительных непрерывных измерений (десятки минут и более) необходимо обеспечивать постоянное вентилирование.
3. Перед началом измерений необходимо производить «разогревочную» серию длительностью 300 разрядов с интервалом 3 с при замыкании тест-объекта на «землю» прибора.
4. Разогрев нужно производить при интервале 3 с, измерения проводить при временном промежутке между разрядами 5 с.
5. Тест-объект следует аккуратно устанавливать перед началом эксперимента и не передвигать, пока эксперимент не будет закончен и все нужные серии сняты.
6. Необходимо следить за условиями окружающей среды:
 - перепады влажности воздуха не должны превышать 5 %;
 - перепады температуры воздуха не должны превышать 5 °C;
 - не включать/выключать электрические приборы во время проведения экспериментов в комнате, в которой стоит прибор ГРВ;
 - не разговаривать по сотовому телефону вблизи прибора ГРВ;
 - количество людей, находящихся рядом с прибором, должно быть постоянным;
 - избегать резких перепадов атмосферного давления.

7. Длительность каждой серии измерений должна быть не менее 140 ГРВ-грамм.
8. При обработке результатов исключать первые 40 снимков из расчета.
9. После каждого изменения условий окружающей среды следует производить калибровку. Для этого можно использовать последние 10 снимков из «разогревочной» серии.

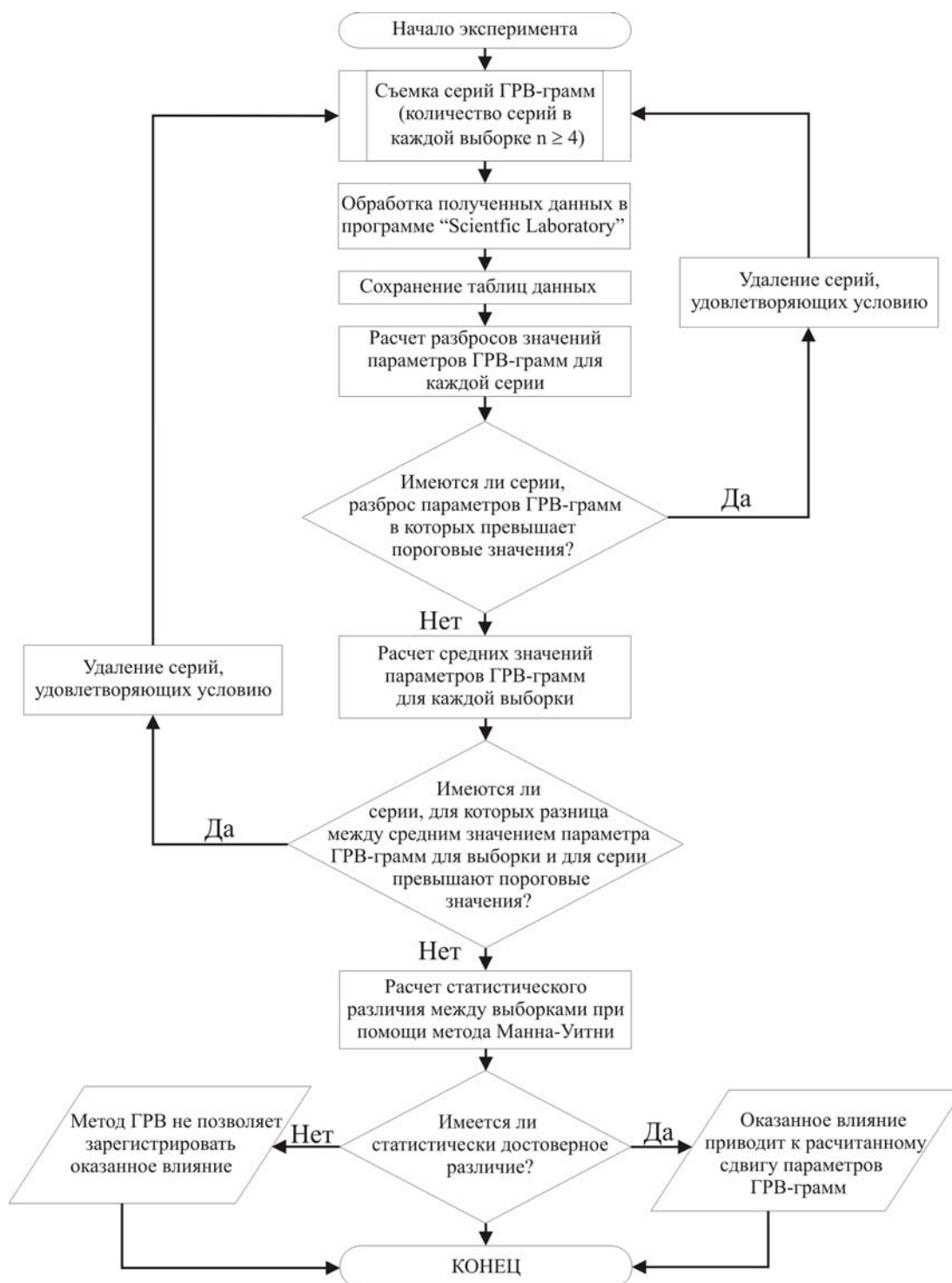


Рис. 4. Алгоритм процедуры проведения измерений

Процедура проведения измерений

В каждом конкретном эксперименте необходимо тщательно продумывать процедуру проведения измерений. Необходимо соблюдать перечисленные выше требования.

Любой эксперимент по исследованию определенного воздействия можно разбить на три этапа: 1 – до воздействия (фон), 2 – воздействие, 3 – после воздействия (последствие). Перед началом непосредственного исследования воздействия следует проверить саму процедуру измерений на стабильность. Для этого надо выполнить приведенную на схеме (рис. 4) последовательность действий, исключив второй этап эксперимента, то есть само воздействие. В случае правильно поставленного эксперимента статистически значимой разницы между значениями параметров ГРВ-грамм не должно быть. Если же разница есть, то следует еще раз перепроверить правильность постановки эксперимента и повторить все шаги.

Как только процедура измерений налажена, проверена и не оказывает значимого влияния на объект исследования, можно переходить к непосредственным измерениям исследуемого воздействия.

Заключение

В результате проведенной работы были выявлены основные дестабилизирующие факторы, оказывающие максимальное влияние на результаты экспериментов, найдены пути устранения этих факторов, а на их основе разработаны общие рекомендации, также был разработан алгоритм процедуры проведения измерений на ГРВ-оборудовании. Все это в совокупности позволяет получать достоверные, воспроизводимые, математически обоснованные результаты при исследовании различных объектов природной среды. В результате проведенных исследований и анализа полученных данных было написано и опубликовано методическое пособие [6], в котором более подробно описана процедура со всеми необходимыми для расчетов математическими формулами.

Литература

1. Коротков К.Г. Основы ГРВ биоэлектрографии. – СПб: ИТМО (ТУ), 2001. – 356 с.
2. Коротков К.Г. Эффект Кирлиан. – СПб: Ольга, 1995. – 218 с.
3. Коротков К.Г. От эффекта Кирлиан к биоэлектрографии. – СПб: Ольга, 1998. – 341 с.
4. Коротков К.Г. Разработка научных основ и практическая реализация биотехнических измерительно-вычислительных систем анализа газоразрядного свечения, индуцированного объектами биологической природы: дис. док. техн. наук. – СПб: СПбИТМО, 1999. – 93 с.
5. Орлов Д.В., Петрова Е.Н., Чайкун К.Е. Параметрические зависимости частотно-резонансных оптоэлектронных контуров // Научно-технический вестник СПбГУ ИТМО. – 2008. – № 48. – С. 225–233.
6. Орлов Д.В. Методика проведения измерений объектов природной среды на программно-аппаратном комплексе газоразрядной визуализации (ГРВ): Методическое пособие / Под ред. д.т.н. Короткова К.Г. – СПб: СПбГУ ИТМО, 2009. – 47 с.

Коротков Константин Георгиевич	– Санкт-Петербургский государственный университет информационных технологий, механики и оптики, доктор технических наук, профессор, korotkov2000@gmail.com
Орлов Дмитрий Владиславович	– Санкт-Петербургский государственный университет информационных технологий, механики и оптики, аспирант, mityaorlov@gmail.com, mitya-9@mail.ru
Величко Елена Николаевна	– Санкт-Петербургский государственный университет информационных технологий, механики и оптики, аспирант, heleonores@mail.ru
Гатчина Юлия Юрьевна	– Санкт-Петербургский государственный университет информационных технологий, механики и оптики, экономист, Gatchina@mail.ifmo.ru

УДК 519.687+004.75

ОЦЕНКА ЭФФЕКТИВНОСТИ РАСПРЕДЕЛЕННЫХ СИСТЕМ
ПРИ РЕШЕНИИ ЗАДАЧ ПЕРЕМЕННОГО РАЗМЕРА

А.С. Хританков

Исследуется проблема описания производительности распределенных систем вычислительных кластеров при решении задач переменной трудоемкости. Для этого используется модель производительности интервальных систем с расписанием, которая в данной работе расширена для учета вероятностного характера трудоемкости задачи. Полученные результаты позволяют применить модель систем с расписанием для расчета верхней границы эффективности и погрешности в ее измерении при решении задач глобальной оптимизации стохастическими алгоритмами.

Ключевые слова: распределенные вычисления, оценка эффективности, глобальная оптимизация.

Введение

Рассмотрим процесс решения задачи глобальной оптимизации эвристическим методом, основанном на подходе Монте-Карло. Важной особенностью является то, что исходная задача A разбивается на M однотипных подзадач $\{A_k\}_1^M$. Например, в методе Basin-Hopping [1] подзадачей является нахождение локального минимума итеративным алгоритмом из начальной точки, выбираемой случайным образом в окрестности предполагаемого минимума. Трудоемкость решения подзадачи напрямую зависит от количества шагов локального алгоритма и сложности вычисления оптимизируемой функции. В широком диапазоне значений параметров алгоритма трудоемкости решения $\{L_k\}_1^M$ подзадач можно полагать независимыми в совокупности случайными величинами, имеющими одинаковое распределение. Пренебрегая затратами на генерацию начальных точек и относя накладные расходы на распараллеливание и коммуникации к решению подзадач, можно положить, что суммарная трудоемкость решения исходной задачи A выражается суммой $L = \sum_{k=1}^M L_k$. Согласно центральной предельной теореме, распределение случайной величины L при большом количестве подзадач близко к нормальному $N(\mu, M\sigma^2)$, где $\mu = EL_i$, $\sigma^2 = DL_i$ в силу единого распределения величин L_i . Трудоемкость решения подзадачи для алгоритмов локальной оптимизации в задачах конформации молекул и докинга протеин-лиганд, согласно экспериментальным данным, имеет положительно асимметричное унимодальное распределение, при этом $\sigma \approx 0,4\mu$. Таким образом, отклонение трудоемкости решения задачи A из $M = 400$ подзадач от среднего значения с вероятностью порядка 95% не превышает $0,04M\mu$. Вследствие того, что точные значения μ и σ неизвестны и должны быть получены в виде оценок, возникает значительная ошибка (около 20%) в определении трудоемкости L , которую необходимо учитывать при сопоставлении экспериментальных данных с моделью работы системы. Кроме того, существенным для модели производительности системы является учет случайного характера изменения трудоемкости подзадач, который приводит к снижению эффективности работы.

В данной работе мы построим модель производительности распределенной системы при решении описанного класса декомпозируемых задач оптимизации A_M , оценим структурную эффективность при решении в рамках интервальной модели, рассмотрим экспериментальный метод оценки трудоемкости решения задачи и получим оценки его точности; сравним построенную модель с работой реального вычислительного комплекса с учетом полученных оценок погрешностей.

Расширение модели систем с расписанием

Для оценки эффективности решения задачи на распределенной вычислительной системе мы будем использовать модель систем с расписанием [2, 3], которую в данном разделе мы кратко изложим. Распределенная система R представляется как совокупность n вычислителей, выделяемых для решения задачи A согласно расписанию $\bar{h}(t)$. В зависимости от цели исследования выбирается модель решения задачи на данной системе. Модель включает некоторый набор предположений относительно структуры задачи, механизма распараллеливания и позволяет рассчитать время решения задачи. Модель такого рода называется эталонной системой P и, по сути, является модельной системой с расписанием, работающей согласно сделанным предположениям по тому же расписанию, что и реальная система. Каждому вычислителю приписывается стоимость использования c_i в единицу времени, а также, в условиях модели P , эталонное время решения T_i – время, которое данный вычислитель затратил бы на решение задачи A . Эталонное время T , т.е. время решения задачи системой P , рассчитывается на основе параметров и предположений модели. Эффективность E_t по времени решения задачи системой R по отношению к системе P определяется как отношение эталонного времени T системы P к реальному времени решения T :

$$E_t = \frac{T}{T}.$$

Ресурсная эффективность E_Φ определяется похожим образом – как отношение стоимости использования системы $\Phi(t)$ за эталонное время T к аналогичной стоимости за реальное время:

$$E_\Phi = \frac{\Phi(T)}{\Phi(T)}.$$

В качестве примера рассмотрим универсальную эталонную систему P , в которой отсутствуют накладные расходы на распараллеливание решения и задача полагается произвольно делимой. Эталонное время решения T в системе P вычисляется как решение задачи минимизации:

$$\min_t t \text{ при условии } \int_0^t \sum_{i=1}^n \pi_i h_i(\tau) d\tau = L, \quad \pi_i = \frac{L}{T_i}.$$

Интервальная эталонная система [3], как и универсальная модель, предполагает, что трудоемкость L остается постоянной при повторном решении задачи и при переходе от последовательного к параллельному решению. При рассмотрении класса задач A_M данное предположение является слишком строгим. Расширим интервальную систему для решения класса задач с переменной трудоемкостью. Обозначим такую систему P_m , в дополнение к ограничениям интервальной системы будем полагать, что $\{L_k\}_1^M$ – независимые в совокупности случайные величины. В дальнейшем будем считать, что их распределение имеет положительный коэффициент асимметрии и конечный коэффициент эксцесса. Напомним, что интервальная система предполагает решение задачи как набора одинаковых по трудоемкости подзадач на системе из кластеров одинаковых вычислителей, управляемых совместно. Подзадачи назначаются на кластер пакетами, промежуток времени решения пакета задач называется этапом решения.

Для эталонной системы P_m можно оценить распределение трудоемкости решения задачи. Действительно, так как $L = \sum_{k=1}^M L_k$, то при больших M , согласно центральной

предельной теореме, распределение L можно полагать приближенно нормальным $N(M\mu, M\sigma^2)$, где μ и σ^2 – математическое ожидание и дисперсия L_k . Итак, мы построили систему P_m , которая отражает основное свойство задач класса A_M – декомпозируемость на множество подзадач, трудоемкости которых являются одинаково распределенными случайными величинами, и учитывает существенные ограничения на управляемость отдельными вычислителями, характерные для современных распределенных систем.

Оценка структурной неэффективности в интервальной модели

Воспользуемся расширением интервальной модели и оценим величину структурной неэффективности, возникающей вследствие различной трудоемкости подзадач. Рассмотрим решение пакета из $M = m \cdot p$ подзадач класса A_M на кластере из p вычислителей. Вследствие того, что подзадачи имеют разную трудоемкость, вычислители, завершившие последовательный расчет m подзадач, будут простаивать, так как следующий пакет подзадач не может быть получен, пока не будет решена последняя подзадача из данного пакета.

Пусть X_i – случайная величина, выражающая суммарную стоимость решения m подзадач на вычислителе i , со средним значением μ_m и дисперсией σ_m^2 . В среднем долю ε_p стоимости вычислений, затраченных на решение подзадач, можно оценить как

$$E\varepsilon_p = E \frac{\sum_{i=1}^p c_i X_i}{\max_{i=1..p} X_i \sum_{i=1}^p c_i}, \Delta_j = \max_{i=1..p} X_i - X_j, X_{i^*} = \max_{i=1..p} X_i,$$

$$E\varepsilon_p = \sum_{i=1}^p \frac{c_i}{c_\Sigma} E \left(\frac{X_i}{X_{i^*}} \right) = \sum_{i \neq i^*} \frac{c_i}{c_\Sigma} \frac{EX_i}{EX_{i^*}} + \frac{c_{i^*}}{c_\Sigma} E \frac{X_{i^*}}{X_{i^*}}.$$

На практике $E\Delta_i$ лежит в пределах от σ_m до $4\sigma_m$ при p от 10 до 50, возрастая как логарифм от числа вычислителей p , $E\Delta_i = a \ln p \cdot \sigma_m$. Дисперсию σ_m^2 можно оценить как дисперсию суммы m случайных величин, $\sigma_m^2 \approx m \cdot \sigma^2$. Для вычислителей из одного кластера логично положить $c_i = c_0 = c_\Sigma / p$, получим:

$$E\varepsilon_p = \frac{p-1}{p} \frac{EX_i}{EX_{i^*}} + \frac{1}{p} = \frac{EX_i}{EX_{i^*}} + \frac{1}{p} \frac{E\Delta_i}{EX_{i^*}}, \quad (1)$$

$$E\varepsilon_p \approx \left(1 + \frac{p-1}{\sqrt{p}} \ln p \frac{C}{\sqrt{M}} \right)^{-1}, \sigma = \theta\mu, EX_{i^*} = EX_i + E\Delta_i, C = a\theta. \quad (2)$$

По построению, величина ε_p является оценкой сверху ресурсной эффективности эталонной системы P_m по отношению к универсальной системе P , усредненной по различным возможным трудоемкостям подзадач.

Из выражения (2) удастся получить функцию изоэффективности $M = F(p, \varepsilon_0)$, указывающую, какого размера (в нашем случае – число подзадач в пакете) должна быть задача, чтобы ее эффективность составляла ε_0 при p вычислителях в кластере:

$$M \approx C^2 \left(\frac{\varepsilon_0}{1 - \varepsilon_0} \right)^2 p \ln^2 p. \quad (3)$$

Оценку средней продолжительности $E\tau$ этапа решения кластера для пакета из M задач получим из выражения для среднего EX_{i*} максимальной трудоемкости задачи на кластере:

$$E\tau = \frac{\mu_m}{c_0} (1 + Cm^{-3/2} \ln p). \quad (4)$$

Следовательно, длительность этапа решения не остается постоянной и при постоянной эффективности возрастает как логарифм количества вычислителей в кластере.

Пусть в распределенной системе имеется N кластеров, долю кластера в полной стоимости решения задачи обозначим через $\xi_j > 0$, $\sum_{j=1}^N \xi_j = 1$. Тогда ресурсная эффективность E_ϕ системы вычисляется как взвешенная сумма эффективностей кластеров E_ϕ^j :

$$E_\phi = \sum_{j=1}^N \xi_j E_\phi^j. \quad (5)$$

Для определения долей ξ_j для конкретного расписания нужно указать алгоритм балансировки, например, представленный в работе [3] с продолжительностью этапа решения (4).

Экспериментальная оценка трудоемкости задачи

Для расчета эффективности решения задачи необходимо знать трудоемкость ее решения L . Метод экспериментальной оценки состоит в последовательном решении ограниченного набора подзадач, измерения трудоемкости их решения и оценки параметров распределения. После этого, используя предположение о нормальности распределения L , рассчитывается среднее значение трудоемкости и дисперсия. Случайным образом выберем K подзадач из M и решим их на вычислителе i последовательно, измеряя время их решения T_i^k . Функцию трудоемкости $L(\cdot)$ определим как $L_k = L(A_k) = T_i^k$. В данной работе мы будем использовать оценки $m_K = \sum_{k=1}^K L_k / K$ для среднего и $s_K^2 = \sum_{k=1}^K (L_k - m)^2 / (K - 1)$ для дисперсии. Отсюда получим оценку среднего значения $\mu_\Sigma = Mm_K$ и стандартного отклонения $\sigma_\Sigma = Ms_K$ для трудоемкости все задачи L . Рассмотрим центральный доверительный интервал для трудоемкости

$$I_{\alpha,\beta} = [E\mu_\Sigma - \alpha\delta_\Sigma; E\mu_\Sigma + \alpha\delta_\Sigma],$$

где отклонение δ_Σ складывается из математического ожидания $E\sigma_\Sigma$ и дисперсий оценок μ_Σ и σ_Σ :

$$\delta_\Sigma = E\sigma_\Sigma + \beta \sqrt{D\sigma_\Sigma + D\mu_\Sigma + 2\text{Cov}(\mu_\Sigma, \sigma_\Sigma)}.$$

Выражения для входящих в формулу величин получены в книге [4]. Обозначая коэффициент эксцесса распределения L_k через $\kappa = \mu_4 / \sigma^4 - 3$, коэффициент корреляции между оценками через ρ и полагая $\sigma = \theta\mu$, получим при $M \approx K^2$:

$$\delta_\Sigma = M\mu \cdot \theta \left(\frac{1}{\sqrt{M}} + \frac{\sqrt{\kappa+6}}{2\sqrt{K}} + \frac{\rho}{K} \sqrt{\frac{\kappa+2}{\kappa+6}} + O(1/K^2) \right). \quad (6)$$

В эксперименте наблюдались значения $\kappa \in [2;4]$, $\theta \in [0,3;0,5]$, следовательно, по порядку величины, при $\alpha = 1$ и $\beta = 1$ (вероятность попадания $P\{L \in I_{1,1}\} \approx 80\%$), и $K = 25$, доверительный интервал для L составляет

$$I_{1,1} = [0,85 \cdot Mm_K; 1,15 \cdot Mm_K]. \quad (7)$$

Основной вклад в точность оценок вносит второе слагаемое в формуле (1), связанное с дисперсией оценки стандартного отклонения. При большом значении коэффициента эксцесса для сохранения точности расчета нужно пропорционально увеличить объем выборки K .

Результаты экспериментов

Для комплекса BnB-Grid при выделении 5 модельных кластеров на MBC-100K в силу схемы управления вычислителями $M = p$. При использовании алгоритма SMBH [5] с максимальным числом итераций локального алгоритма $N_{\max} = 8192$ и радиусе окрестности возмущения $r = 0,9$, экспериментально установлено $\theta \approx 0,4$; исходя из распределения трудоемкости решения подзадач, численно рассчитан коэффициент $a = 0,84$. В данном эксперименте трудоемкость рассчитывалась по $K = 120$ подзадачам. Экспериментально полученные значения ресурсной эффективности, доверительные интервалы $I_{1,1}$ (7), и графики функций (1) и (2) приведены на рис. 1.

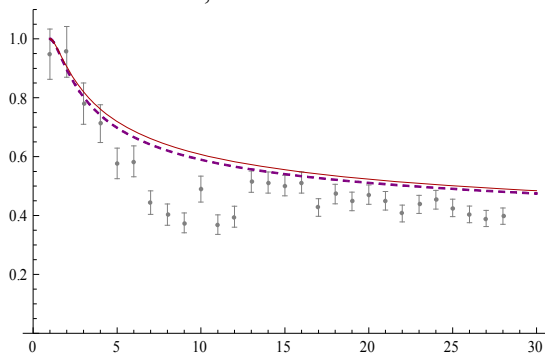


Рис. 1. Сравнение модели структурной неэффективности и экспериментальных данных: приближенная модель (пунктир) и точная модель (сплошная)

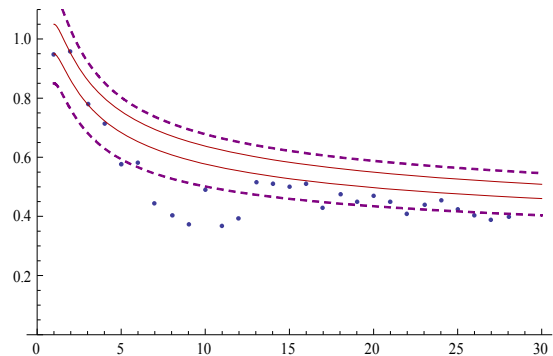


Рис. 2. Возможности обнаружения отклонений результатов экспериментов от модели при оценке по ограниченной выборке (пунктир) и каждой подзадаче (сплошная)

Заметим, что доверительный интервал при $p \leq 12$ больше, так как задача состояла из 120 подзадач, в то время как при $12 < p \leq 28$ – из 512 подзадач. Из графика видно, что различия между (1) и (2) имеют порядок точности измерений, поэтому использование приближения (2) в теоретических выкладках вполне оправдано. Полученная модель оценивает эффективность сверху, отклонения экспериментальных данных, по большей части, объясняются простым вычислителей кластера при недостатке задач, что заметно при $p = 7-12$, и замедлением при использовании нескольких ядер одного процессора. При этом, как видно из рис. 2, определение трудоемкости по выборке размера $K \approx \sqrt{M}$ позволило бы обнаружить снижение эффективности только при $p = 7-12$, что вполне достаточно для реальных задач, когда повторное решение задачи не имеет смысла, а измерение трудоемкости подзадач невозможно в процессе решения. Сплошными линиями на рис. 2 показаны границы при $K \approx M$, видно, что при исследовании

поведения системы этого достаточно для обнаружения небольших отклонений от модели, ошибка определения эффективности составит при $M \approx 500$ порядка 5%.

Заключение

Рассмотрена проблема описания производительности распределенных вычислительных систем при решении задач с переменной трудоемкостью, таких как задачи глобальной оптимизации при использовании методов типа Монте-Карло. Предложенное обобщение модели интервальных систем позволяет оценить сверху эффективность решения на системах вычислительных кластеров (5). Полученные оценки погрешности (6) в определении трудоемкости экспериментальным путем по небольшой доле подзадач на имеющихся экспериментальных данных подтверждают адекватность модели.

В дальнейшем планируется продолжить исследования эффективности интервальных систем, уточнить верхнюю оценку, рассмотрев падение эффективности при недостатке подзадач, обобщить результаты по интервальным системам на случай переменной трудоемкости задачи. Кроме того, необходимо рассмотреть влияние структуры расписания на эффективность по времени.

Литература

1. Wales D.J., Doye J.P.K. Global Optimization by Basin-Hopping and the Lowest Energy Structures of Lennard-Jones Clusters Containing up to 110 Atoms // *Journal of Physical Chemistry*. – 1997. – № 101. – P. 5111–5116.
2. Посыпкин М.А., Хританков А.С. О понятии ускорения и эффективности в распределенных системах // *Научный сервис в сети Интернет'2008: решение больших задач: Труды конференции*. – 2008. – С. 149–156.
3. Хританков А.С. Об одном алгоритме балансировки вычислительной нагрузки в распределенных системах // *Параллельные вычислительные технологии (ПаВТ'2009): Труды международной научной конференции (Нижний Новгород, 30 марта – 3 апреля 2009 г.)*. – Челябинск: Изд. ЮУрГУ, 2009. – 839 с., С. 778–783.
4. Крамер Г. Математические методы статистики. – 2-е изд. – М.: Мир, 1975. – 648 с.
5. Посыпкин М.А. Методы и распределенная программная инфраструктура для численного решения задачи поиска молекулярных кластеров с минимальной энергией // *Параллельные вычислительные технологии (ПаВТ'2009): Труды международной научной конференции (Нижний Новгород, 30 марта – 3 апреля 2009 г.)*. – Челябинск: Изд. ЮУрГУ, 2009. – 839 с., С. 269–281.

Хританков Антон Сергеевич – Московский физико-технический институт (государственный университет), аспирант, anton.khritankov@gmail.com

УДК: 681.3.06

АКТОРНОЕ РАСШИРЕНИЕ ЯЗЫКА JAVA В СРЕДЕ MPS

А.Р. Жукова, М.А. Мазин

Описанное в работе акторное расширение добавляет автоматически распараллеливающиеся функциональные конструкции в универсальный язык Java. Необходимость такого расширения вызвана слабой приспособленностью языка Java к написанию параллельных программ. Акторная модель, поддерживаемая расширением, успешно применяется в функциональном языке параллельного программирования Erlang. В качестве средства создания расширения была выбрана среда мета-программирования MPS, что позволило автоматически получить интегрированные средства разработки для применения расширения и, кроме того, достичь совместимости с другими языковыми расширениями, созданными в среде MPS.

Ключевые слова: акторная модель, проблемно-ориентированное расширение, параллельное программирование, языково-ориентированное программирование

Введение

В настоящее время растет актуальность разработки параллельных программ. Это связано с тем, что дальнейший рост производительности вычислительных систем более не может достигаться увеличением тактовой частоты процессоров, и главной тенденцией становится увеличение числа ядер в системах [1]. Но для того, чтобы программа выполнялась эффективно на нескольких ядрах, она должна быть написана с применением техник параллельного программирования [2].

Универсальные императивные языки программирования [3], наиболее распространенные на сегодняшний день, хорошо подходят для реализации последовательных алгоритмов, но плохо приспособлены для написания алгоритмов параллельных. Это связано с тем, что программа, написанная на императивном языке, представляет собой последовательность инструкций процессору, изменяющих состояние памяти [4]. Из-за того, что порядок этих команд зафиксирован, выполнение программы не может быть распределено между несколькими процессорами. Поэтому программисты вынуждены самостоятельно разбивать свою программу на параллельно выполняющиеся потоки [5].

Автоматическому распараллеливанию хорошо поддаются программы, написанные на функциональных языках программирования [6, 7]. Но функциональные языки, в отличие от императивных, значительно менее распространены. Тем не менее, некоторые функциональные конструкции в последнее время стали активно проникать в универсальные императивные языки программирования. Например, становится популярным использование в императивных языках замыканий [8] для обработки списков, в некоторых случаях [9] это позволяет выполнять автоматическое распараллеливание по данным [10].

В связи с этим естественно продолжить адаптацию абстракций параллельных функциональных программ для императивных языков программирования. К таким абстракциям, в частности, относится акторная модель [11], реализованная в языке программирования *Erlang* [12].

Акторная модель была предложена К. Хьюиттом, П. Бишопом и Р. Штайгером в 1973 г. Они ввели понятие актора – примитива параллельных вычислений, обладающего потоком выполнения и способного обмениваться сообщениями с другими акторами. Актор обрабатывает сообщения из собственной очереди сообщений и в ответ на них выполняет полезные действия.

Разработанное авторами языковое расширение позволяет использовать акторные конструкции при программировании на языке *Java* [13]. На рис. 1 приведены примеры объявления типа актора, объявления сообщения, создания нового актора и отправки сообщения. Из рисунка видно, что такой код синтаксически почти не отличается от обя-

явления класса, объявления метода, создания нового объекта и вызова метода в языке *Java*. Поэтому *Java*-программисту потребуются минимальные усилия для изучения синтаксиса акторного языкового расширения.

```

public actor NamedActor {
    private String name;
    public NamedActor(String name) {
        this.name = name;
    }
    public void printName() {
        System.out.println(this.name);
    }
}

...
public static void main(String[] args) {
    NamedActor na = new NamedActor("Actor");
    na.printName();
    ...
}

```

Рис. 1. Пример синтаксиса

Традиционно создание языка или языкового расширения [14] включает в себя разработку общей инфраструктуры языка: абстрактного и конкретного синтаксиса [15], системы типов [16], операционной семантики и т.д. Для этого реализуются лексический, синтаксический и семантический анализаторы, трансляторы и другие утилиты. Существуют различные инструменты, позволяющие частично автоматизировать эту деятельность [17, 18].

В настоящее время для повышения эффективности труда программисты используют интегрированные среды разработки [19], обладающие знаниями о языке программирования и значительно ускоряющие разработку. Поэтому создание нового языка предполагает также создание инфраструктуры среды разработки [20] с редактором, автодополнением кода, подсветкой ошибок, рефакторингами, системами версионирования и др. Помимо того, что эта задача сама по себе является ресурсоемкой и требует высококвалифицированных разработчиков, возникает проблема совместимости различных языковых расширений при использовании в рамках одной программы. Под совместимостью понимается возможность совместного использования компонентов, даже если они созданы независимо [21]. Например, различные библиотеки для языка *Java* совместимы, если их классы расположены в разных пакетах. Обеспечить совместимость языковых расширений несколько сложнее.

Из существующих средств создания языковых лишь *JetBrains MPS* [14, 22, 23] позволяет одновременно создавать как совместимые языковые расширения, так и языковую инфраструктуру для них. Поэтому именно среда *MPS* была выбрана для создания акторного расширения языка *Java*. Среда *MPS* поставляется с большим количеством уже готовых языков и языковых расширений, поддерживающих дополнительные конструкции: коллекции, замыкания, регулярные выражения, работу с датами и др.

Базовые понятия расширения Actor Language

Созданное авторами языковое расширение позволяет использовать акторы наряду с обычными объектами при программировании на языке *Java*. Декларация типа актора, поддерживаемая расширением, синтаксически совпадает с декларацией *java*-класса с той лишь разницей, что вместо ключевого слова «*class*» используется ключевое слово «*actor*». Для типа актора могут быть определены конструкторы, поля и объявления типов обрабатываемых сообщений. Синтаксис последних полностью совпадает с синтаксисом объявлений обычных методов в языке *Java*. Аналогично, синтаксис отправки сообщений актору не отличается от синтаксиса вызова методов у объекта. Но, в отличие

от вызова методов, отправка сообщений происходит асинхронно: сообщения помещаются в индивидуальную очередь актора и последовательно актором обрабатываются.

Содержание потока в языке *Java* – ресурсоемкий процесс, поэтому для обработки сообщений актору не выделяется отдельный экземпляр класса «*java.lang.Thread*». Вместо этого отправка всех сообщений всем акторам происходит посредством добавления сообщений в очередь общего диспетчера. Диспетчер обладает пулом потоков [24], размер которого либо задается программистом явно, либо вычисляется, исходя из числа доступных ядер в системе. По мере появления новых сообщений в очереди диспетчера и освобождения потоков из пула сообщения последовательно извлекаются из очереди для обработки акторами. При этом гарантируется, что сообщения, направленные одному и тому же актору, не могут обрабатываться параллельно двумя разными потоками. Таким образом, эмулируется обладание актором потоком управления.

На рис. 2 приведен пример объявления двух типов акторов: «*Ping*» и «*Pong*». При отправке актору «*Ping*» сообщения «*start*» между акторами начинается обмен сообщениями «*ping*» и «*pong*». После отправки заданного параметром «*pingCount*» числа сообщений актор «*Ping*» отправляет актору «*Pong*» сообщение «*stop*» и прекращает работу. В ответ на сообщение «*stop*» актор «*Pong*» также завершает работу.

```
public actor Ping {
    private int pingsLeft;
    public Ping() {}
    public void pong(Pong pong) {
        if (this.pingsLeft-- >= 0) {
            System.out.println("Ping: pong");
            pong.ping(this);
        } else {
            System.out.println("Ping: stop");
            pong.stop();
        }
    }
    public void start(int count, Pong pong) {
        System.out.println("Ping: start");
        this.pingsLeft = count;
        this.pong(pong);
    }
}

public actor Pong {
    public Pong() {}
    public void ping(Ping ping) {
        System.out.println("Pong: ping");
        ping.pong(this);
    }
    public void stop() {
        System.out.println("Pong: stop");
    }
}
```

Рис. 2. Пример объявления акторов

Для запуска обмена сообщениями необходимо создать акторы «*Ping*» и «*Pong*» и отправить актору «*Ping*» сообщение «*start*». Код, выполняющий эти действия, и результаты работы этого кода приведены на рис. 3.

```
...
public static void main(String[] args) {
    Pong pong = new Pong();
    Ping ping = new Ping();
    ping.start(3, pong);
}
...
Ping: start
Ping: pong
Pong: ping
Ping: pong
Pong: ping
Ping: pong
Pong: ping
Ping: stop
Pong: stop
```

Рис. 3. Запуск обмена сообщениями между акторами и вывод программы

Обработка сообщений с задержкой

Языковое расширение *Actor Language* позволяет при посылке сообщения указать минимальную задержку, с которой оно должно быть обработано. Для этого в языковое расширение введена конструкция *defer*, которая может быть применена к оператору посылки сообщения. В качестве параметра конструкция *defer* принимает величину минимальной задержки, для задания которой используется встроенное в среду *MPS* языковое расширение *Dates*.

В языковом расширении *Dates* предусмотрены специальные проблемно-ориентированные конструкции для работы с датами, временем и промежутками времени. В частности, оно позволяет записывать периоды времени наглядно, указывая их величину и размерность: «5 minutes», «1 year».

Пример использования конструкции *defer* приведен на рис. 4: код актора *Pong* отредактирован таким образом, чтобы посылка сообщения *pong* происходила не сразу, а с задержкой в пять миллисекунд.

```
public actor Pong {
    public Pong() {}
    public void ping(Ping ping) {
        System.out.println("Pong: ping ");
        ping.pong(this).defer(5 milliseconds);
    }
    public void stop() {
        System.out.println("Pong: stop");
    }
}
```

Рис. 4. Отправка сообщения с задержкой

Отложенный результат

В традиционном объектно-ориентированном программировании некоторые методы в результате своего выполнения возвращают вычисленные значения. Эти значения становятся доступны вызывающему коду сразу по окончании синхронного выполнения таких методов. Так может быть организовано взаимодействие между вызываемым и вызывающим кодом.

При программировании с использованием акторов нет возможности использовать возвращаемые значения, так как к тому моменту, когда операция асинхронной посылки сообщения оказывается выполненной, само сообщение может быть еще не обработано, а возвращаемое значение не вычислено.

Для того чтобы вызываемый код мог оповестить вызывающий код об окончании обработки события и передать вычисленное значение, можно передавать вызывающий актор в качестве параметра сообщения. По окончании обработки сообщения вызываемый актор должен послать вызывающему актору ответное сообщение. Именно так организовано взаимодействие в приведенном выше примере (рис. 2). Однако при таком подходе код обрабатывающего сообщения актора должен «знать» о структуре посылающего сообщения актора, т.е. возникает нежелательная сильная связанность кода [25].

Чтобы обойти эту проблему и позволить использовать возвращаемые обработчиками сообщений значения, в языковом расширении *Actor Language* существует поддержка механизма отложенных результатов. Для любого сообщения актора, так же как и для любого метода в языке *Java*, могут быть заданы тип возвращаемого значения и набор генерируемых исключений. Код обработчика сообщения должен либо вычислить значение соответствующего типа, либо сгенерировать исключение.

Оператор отложенной отправки сообщения имеет специальный тип *deferred*, параметризованный типом возвращаемого сообщения значения. Например, если обработчик сообщения возвращает значение типа *String*, то асинхронная отправка этого сообщения вернет значение типа *deferred<String>*. Используя это значение, код, посылающий сообщение, может определить действия как в связи с успешной обработкой сообщения, так и в связи с возникновением исключительной ситуации. Действия задаются в виде замыканий – анонимных функций, языковая поддержка которых осуществляется встроенным в среду MPS расширением *Closures*.

Использование отложенного значения может быть продемонстрировано на примере программы редактора графов. Одной из функций такой программы является чтение графов из файлов, реализованное актором *GraphReader*. При получении сообщения *read* актор *GraphReader* начинает выполнять длительную операцию чтения и разбора файла. Результатом этой операции является экземпляр класса *GraphModel*, но в процессе выполнения операции может возникнуть исключение *IOException*. Часть программы, использующая актор *GraphReader* (рис. 5), посылает ему сообщение *read*, а полученное отложенное значение сохраняет в переменной *model*. Далее с помощью оператора *addCallback* назначается действие в связи с успешной обработкой сообщения, а с помощью оператора *addErrback* – обработчик исключения *IOException*.

```

...
GraphReader reader = new GraphReader();
deferred< GraphModel > model = reader.read(new File("./graphModel.grph"));
model.addCallback({GraphModel result => frame.setGraphModel(result); });
model.addErrback({IOException e =>
    JOptionPane.showMessageDialog(fileChooser, "Cannot open the file"); });
...

```

Рис. 5. Действия по окончании обработки сообщения

Заключение

Разработанное авторами языковое расширение *Actor Language* позволяет в привычной для Java-программистов манере использовать автоматически распараллеливающиеся функциональные конструкции. Расширение упрощает создание программ, которые эффективно используют возможности многоядерных компьютеров.

Благодаря тому, что расширение создано в среде MPS, программисты имеют возможность использовать его совместно с другими языковыми расширениями этой среды. В дальнейшем планируется разработать языковое расширение *Actor Language* конструкциями для неблокирующего асинхронного ввода / вывода. Это позволит эффективно использовать акторы в разработке веб-приложений [26], а также в приложениях, интенсивно работающих с файловой системой.

Литература

1. Adve S. Parallel Computing Research at Illinois: The UPCRC Agenda. – 2008.
2. Воеводин В.В. Параллельные вычисления / В.В. Воеводин, Вл.В. Воеводин. – БХВ-Петербург, 2004.
3. Sebesta R. Concepts of Programming Languages / Robert W. Sebesta. – Boston, MA, USA: Addison-Wesley Longman Publishing Co., Inc., 2001.
4. von Neumann J. First Draft of a Report on the EDVAC / John von Neumann // IEEE Annals of the History of Computing. – 1993. – Vol. 15. – № 4. – P. 28–75.
5. Эндрюс Г. Основы многопоточного, параллельного и распределенного программирования / Грегори Р. Эндрюс. – Вильямс, 2003.

6. Непейвода Н.Н. Стили и методы программирования. Курс лекций: Учебное пособие. – М.: Интернет-университет информационных технологий, 2005.
7. Trinder P.W. GUM: a portable parallel implementation of Haskell. – 1996.
8. Fowler, M. Closure. – 2004 [Электронный ресурс]. – Режим доступа: <http://martinfowler.com/bliki/Closure.html>, своб.
9. Duy J. PARALLEL LINQ: Running Queries On Multi-Core Processors / Joe Duy, Ed Essey // MSDN Magazine. – 2007.
10. Hillis D. Data Parallel Algorithms / W. Daniel Hillis, Guy L. Jr. Steele // Commun. ACM. – 1986. – Vol. 29. – № 12. – P. 1170–1183.
11. Hewitt C. Universal Modular ACTOR Formalism for Artificial Intelligence / Carl Hewitt, Peter Bishop, Richard Steiger // IJCAI. – 1973. – P. 235–245.
12. Armstrong, J. Programming Erlang: Software for a Concurrent World / Joe Armstrong. – Pragmatic Bookshelf, 2007.
13. Эккель Б. Философия Java. – СПб: Питер, 2009.
14. Дмитриев С. Языково-ориентированное программирование: следующая парадигма // RSDN Magazine. – 2005. – № 5.
15. Компиляторы. Принципы, технологии и инструментарий / А. Ахо, М. Лам, Р. Сети, Дж. Ульман. – Вильямс, 2008.
16. Luo Z. Computation and Reasoning: A Type Theory for Computer Science // International Series of Monographs on Computer Science. № 11. – Oxford University Press, 1994.
17. Bovet J. ANTLRWorks: an ANTLR grammar development environment / Jean Bovet, Terence Parr // Softw. Pract. Exper. – 2008. – Vol. 38. – № 12.
18. Sevenich R. Compiler Construction Tools // Linux Gazette. – 1999. – Vol. 39 [Электронный ресурс]. – Режим доступа: <http://linuxgazette.net/issue39/sevenich.html>, своб.
19. Fowler M. Crossing Refactoring's Rubicon / Martin Fowler // ThoughtWorks. – 2001. [Электронный ресурс]. – Режим доступа: <http://martinfowler.com/articles/refactoringRubicon.html>.
20. Давыдов С. IntelliJ IDEA. Профессиональное программирование на Java. Наиболее полное руководство / С. Давыдов, А. Ефимов. – СПб: БХВ-Петербург, 2005.
21. Solomatov K. DSL Adoption with JetBrains MPS / Konstantin Solomatov // DZone, Architect Zone. – 2009.
22. Fowler M. A Language Workbench in Action. – MPS, 2005.
23. Fowler M. Language Workbenches: The Killer-App for Domain-Specific Languages. — 2005. — jun. [Электронный ресурс]. – Режим доступа: <http://www.martinfowler.com/articles/languageWorkbench.html>.
24. Goets B. Java theory and practice: Thread pools and work queues / Brian Goets // Developer Works, 2002.
25. Приемы объектно-ориентированного проектирования. Паттерны проектирования / Э. Гамма, Р. Хелм, Р. Джонсон, Дж. Влиссидес. Библиотека программиста. – СПб: Питер, 2007.
26. Моделирование контроллера Web-приложений с использованием UML / Е.А. Горшкова, Б.А. Новиков, Д.Д. Белов et al. // Программирование. – 2005. – № 1. – P. 44–51.

Жукова Анна Руслановна	– Санкт-Петербургский государственный университет, студент, anna.zhukova@math.spbu.ru
Мазин Максим Александрович	– Санкт-Петербургский государственный университет информационных технологий, механики и оптики, аспирант, mazine@rain.ifmo.ru

УДК 53.084.2

**РОСТ НАНОВИСКЕРА ПОД ВОЗДЕЙСТВИЕМ
ЭЛЕКТРОННОГО ПУЧКА: МАТЕМАТИЧЕСКАЯ МОДЕЛЬ****С.А. Чивилихин, А.О. Голубок, И.С. Мухин**

Построена математическая модель роста нановискера на поверхности проводника в вакуумной камере электронного сканирующего микроскопа под действием сфокусированного электронного пучка. Положительные ионы углерода, выбиваемые с поверхности камеры, захватываются полем пучка электронов. Двигаясь вдоль пучка, ионы попадают на поверхность нановискера, что и обеспечивает его рост. В работе электрическое поле и потенциал пучка электронов аппроксимированы полем и потенциалом равномерно заряженной нити. Для такой постановки задачи написано уравнение Власова, из которого вытекает уравнение движение поверхности растущего нановискера. С помощью численных методов рассчитан и визуализирован процесс роста, строится динамическая сетка, имитирующая растущей вискер – рассмотрение динамики роста во времени. Полученные данные хорошо согласуются с экспериментом.

Ключевые слова: нановискер, математическая модель роста, уравнение движение поверхности.

Введение

В настоящее время широко распространены такие методы роста вискеро́в, как рост в установках молекулярно-пучковой эпита́ксии, рост из газовой фазы с использованием металлоорганических соединений, выращивание вискеро́в методами электролиза, электропереноса, метод лазерной абляции и т.д. В данной работе предложена математическая модель, описывающая процессы роста расширенного метода неконтролируемого самоорганизованного роста ансамбля вискеро́в под действием электронного пучка. В этом методе [1] часть электронного пучка засвечивает подложку, покрытую углеродом, а часть пучка попадает на поверхность, где планируется рост вискера. В месте экспозиции пучка образуется локально разогретая, отрицательно заряженная область, которая притягивает к себе положительные ионы углерода. Эта область и есть точка зарождения вискера. При перемещении места экспозиции электронного пучка в пространстве, вискер как бы тянется за пучком. При этом растет одиночный (не разветвленный) вискер диаметром 50–100 нм. Предлагаемый метод предполагает контролирование геометрии роста и свойств вискеро́в прямо во время роста.

Экспериментальная установка

На рис. 1 представлена схема экспериментальной установки. Эксперименты проводились в рабочей камере растрового электронного микроскопа (РЭМ) Inspect FEI. Максимальное ускоряющее напряжение составляло 30 кВ, диаметр сфокусированного электронного пучка – 2 нм. Исходный СЗМ-зонд (1) размещается на столике (2) РЭМ. Ниже зонда располагается мишень (3), представляющая собой металлическую поверхность с углеродным покрытием. Электронный пучок (4) фокусируется на вершине зонда, и осуществляется сканирование по области с небольшой площадью. При этом производится экспонирование как участка зонда (1), так и участка расположенной ниже мишени (3). В зоне экспонирования на поверхности зонда происходит зарождение нановискеро́в (НВ) (5), наблюдаемое на экране РЭМ в режиме вторичных электронов. Задавая направление и скорость перемещения пучка электронов относительно неподвижного зонда, а также площадь экспонирования, ускоряющее напряжение и ток пучка, можно управлять ростом НВ.

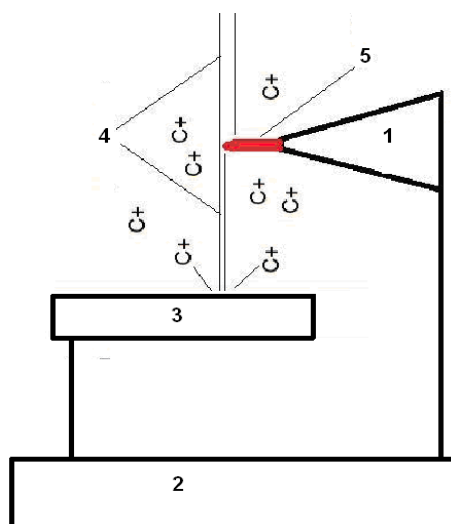


Рис. 1. Схема экспериментальной установки для формирования одиночного НВ на вершине зонда СЗМ: 1 – исходный зонд СЗМ, 2 – столик РЭМ, 3 – мишень с углеродным покрытием, 4 – пучок электронов, 5 – растущий НВ

Математическая модель

Рассмотрим поверхность мишени, покрытую углеродной пленкой, на которую падает пучок электронов с энергией 30 кэВ. Так как энергия ионизации углерода составляет несколько эВ, то с поверхности мишени летят ионизованные атомы углерода. Эти ионы взаимодействуют с пучком электронов, который мы рассматриваем как полубесконечную заряженную нить с линейной плотностью заряда γ_0 . Для описания влияния проводящей мишени воспользуемся методом отраженных зарядов. Для этого введем под поверхностью мишени отраженную полубесконечную нить с зарядом $-\gamma_0$.

Определим потенциал $\varphi(z, r)$ и компоненты вектора напряженности электрического поля E , создаваемого линейным зарядом $\gamma(z)$. Для удобства будем использовать полярную систему координат (r, z, φ) . Из симметрии понятно, что по координате (φ) поле не меняется:

$$\varphi(z, r) = \int_{-\infty}^{\infty} \frac{\gamma(z') dz'}{4\pi\epsilon_0 \sqrt{(z' - z)^2 + r^2}} = \frac{\gamma_0}{4\pi\epsilon_0} \ln \left(\frac{\sqrt{r^2 + h^2} + h}{\sqrt{r^2 + h^2} - h} \right),$$

$$E_r = \int_{-\infty}^{\infty} \frac{r\gamma(z') dz'}{4\pi\epsilon_0 \left((z' - z)^2 + r^2 \right)^{3/2}} = \frac{\gamma_0}{2\pi\epsilon_0 r} \frac{z}{\sqrt{z^2 + r^2}},$$

$$E_z = \int_{-\infty}^{\infty} \frac{(z' - z)\gamma(z') dz'}{4\pi\epsilon_0 \left((z' - z)^2 + r^2 \right)^{3/2}} = -\frac{\gamma_0}{2\pi\epsilon_0} \frac{1}{\sqrt{z^2 + r^2}},$$

где ϵ_0 – диэлектрическая проницаемость.

Отношение перпендикулярной и радиальной компонент напряженности электрического поля равно

$$\left| \frac{E_z}{E_r} \right| = \frac{r}{z},$$

следовательно, вблизи нити, при $r \ll z$, доминирует радиальная компонента электрического поля. Поле нити удерживает ионы углерода и направляет их вдоль оси нити, не меняя их продольную компоненту скорости.

На рис. 2 представлен график потенциала поля в зависимости от радиальной координаты при фиксированной высоте. Видно, что поле для ионов является потенциальной ямой.

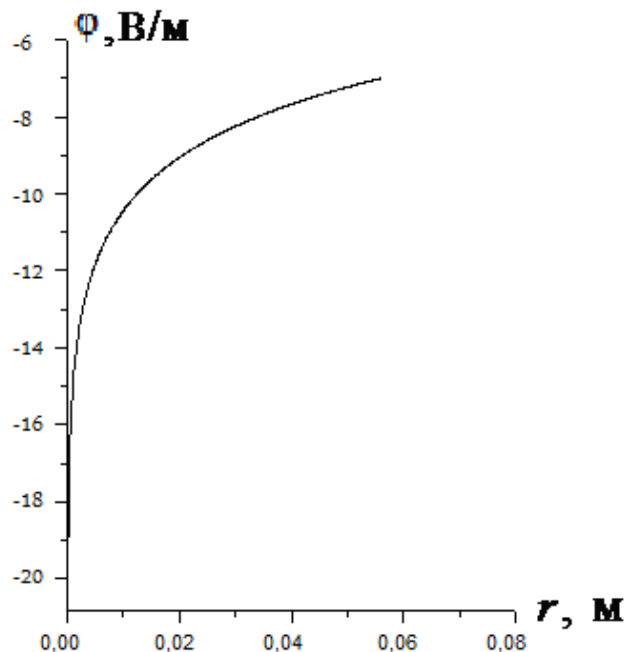


Рис. 2. Зависимость потенциала поля от радиальной координаты

Рассмотрим предельные случаи:

$$\text{а) } r \ll z, \quad \phi = \frac{\gamma_0}{2\pi\epsilon_0} \ln\left(\frac{2z}{r}\right);$$

$$\text{б) } r \gg z, \quad \frac{\gamma_0 z}{2\pi\epsilon_0 r}.$$

Для учета конечного радиуса R пучка электронов вводим внутри пучка квадратичный потенциал, соответствующий равномерно зараженному цилиндру.

Рассмотрим движение ионов углерода вблизи поверхности роста на высоте z . Так как ионы колеблются в потенциальной яме вблизи пучка, то $r \ll z$, и электрическое поле имеет вид

$$E = E_R \begin{cases} \frac{R}{r}, & r > R; \\ \frac{r}{R}, & r < R, \end{cases}$$

где $E_R = \frac{\gamma_0}{2\pi\epsilon_0 R}$. Соответственно, энергия иона W в поле равна

$$W = -W_R \begin{cases} 1 + 2 \ln\left(\frac{r}{R}\right), & r > R; \\ \frac{r^2}{R^2}, & r < R, \end{cases}$$

где $W_R = -\frac{\gamma_0}{4\pi\epsilon_0}$.

Напишем уравнение движения границы растущего вискера. Рассмотрим уравнение Власова [2] движения ионов электрона:

$$\frac{\partial f}{\partial t} + \frac{\partial}{\partial r}(rf) + \frac{\partial}{\partial v}(\dot{v}f) = 0, \quad (1)$$

где f – функция распределения ионов по скоростям и координате, v – скорость движения точек поверхности вискера при его росте, r – их координата.

Мы рассматриваем стационарную задачу, когда все переходные процессы уже прошли, т.е. $\frac{\partial f}{\partial t} = 0$. С учетом ускорения $\dot{v} = \frac{F}{m}$ уравнение (1) принимает вид

$$v \frac{\partial}{\partial r}(f) + \frac{F}{m} \frac{\partial}{\partial v}(f) = 0, \quad (2)$$

где F – действующая сила, m – масса.

Для решения этого уравнения используем метод характеристик [3, 4]. Полная энергия иона равняется $E = \frac{mv^2}{2} + W(r)$. Тогда $f\left(\frac{mv^2}{2} + W(r)\right) = f_0\left(\frac{mv_0^2}{2} + W(r_0)\right)$.

Считаем, что начальное распределение описывается распределением Больцмана,

$$f_0 = A \exp\left(-\frac{\frac{mv_0^2}{2} + W(r_0)}{kT}\right),$$

тогда

$$f = A \exp\left(-\frac{\frac{mv^2}{2} + W(r)}{kT}\right).$$

Соответственно, поток ионов углерода равен $\mathbf{j} = v\mathbf{n}(r)$, где

$n(r) = n_0 \exp\left(-\frac{W(r)}{kT}\right)$ – концентрация ионов углерода у поверхности роста.

Рассмотрим движение поверхности растущего вискера. Так как в общем случае никакой симметрии у растущей поверхности нет, то будем использовать декартову систему координат (x, y, z) . Пусть поверхность задана уравнением $\Phi(x, y, z, t) = 0$, причем внутри вискера $\Phi(x, y, z, t) < 0$, снаружи $\Phi(x, y, z, t) > 0$. Тогда нормаль к поверхности вискера определяется как $\mathbf{n} = \frac{\nabla\Phi}{|\nabla\Phi|}$, соответственно, нормальная скорость движения

границы равна

$$v_n = -\frac{\partial_t \Phi}{|\nabla\Phi|}.$$

С другой стороны, нормальная скорость поверхности определяется выражением $v_n = \frac{\mathbf{j}\mathbf{n}}{\rho}$, где ρ – плотность вещества вискера. Следовательно, уравнение движения границы растущего вискера имеет вид

$$\partial_t \Phi + \frac{\mathbf{j} \mathbf{n}}{\rho} |\nabla \Phi| = 0. \quad (3)$$

Это уравнение решается численно. На рис. 2 представлена поверхность вискера, рассчитанная согласно уравнению (3). По осям отложены декартовы координаты. Вертикальной линией изображен сфокусированный пучок электронов, под действием которого происходит рост вискера. Вискер растет на поверхности, представляющей собой закругленный на вершине конус (имитирующий химически заточенную вольфрамовую иглу).

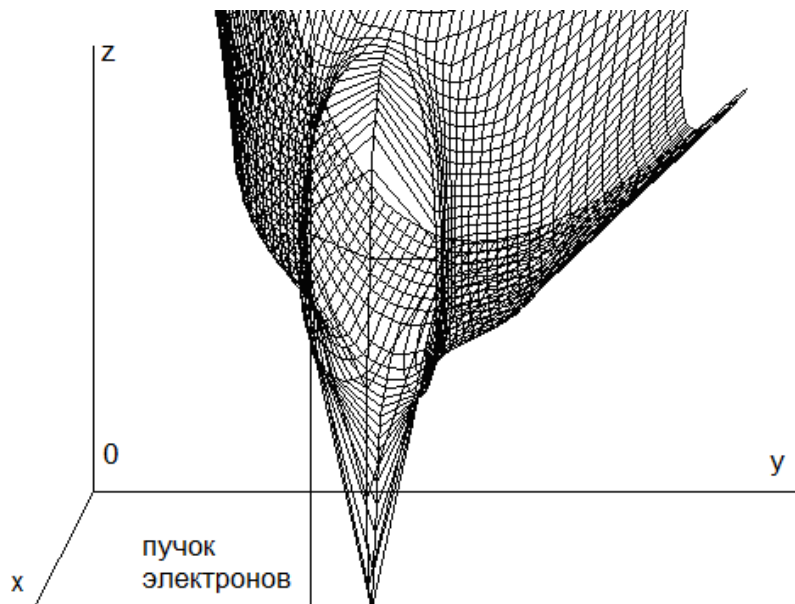


Рис. 2. Результат численного моделирования роста вискера, растущего на острие вольфрамовой иглы

Заключение

Предложена математическая модель роста нановискера на поверхности проводника под действием сфокусированного электронного пучка. Несмотря на упрощающие предположения, построенная модель качественно согласуется с результатами эксперимента [1].

Видно, что вискер является несимметричным в сечении относительно оси роста. Этот факт подтвердился при исследовании вискеров, выращенных в электронном сканирующем микроскопе, при повороте вискеров на 90° вокруг оси роста. Во время роста вискера существует выделенное направление – ось электронного пучка. Ионы углерода, поднимаясь вверх, колеблются вокруг этой оси. Этим и объясняется факт, что вискер имеет несимметричную форму.

Модель позволяет оценивать влияние параметров роста на процессы роста и проводить численные эксперименты по поиску оптимальных условий роста при заданных конечных характеристиках вискера.

Литература

1. Голубок А.О., Ковров А.В., Левичев В.В., Мухин И.С., Приходько О.А. Формирование одиночных нановискеров на вершинах зондов сканирующего зондового микроскопа // Научно-технический вестник СПбГУ ИТМО. – 2009. – № 4(62). – С. 82–87.

2. Веденяпин В.В. Кинетические уравнения Больцмана и Власова. – М.: Физматлит, 2001. –111 с.
3. Курант Р. Уравнения с частными производными. – М.: Мир, 2000. – 843 с.
4. Зайцев В.Ф., Полянин А.Д. Справочник по дифференциальным уравнениям с частными производными первого порядка. – М.: Физматлит, 2003. –416 с.

- Чивилихин Сергей Анатольевич** – Санкт-Петербургский государственный университет информационных технологий, механики и оптики, кандидат физ.-мат. наук, доцент, sergey.chivilikhin@gmail.com
- Голубок Александр Олегович** – Институт аналитического приборостроения РАН, доктор физ.-мат. наук, заместитель директора по научной работе, aogolubok@mail.ru
- Мухин Иван Сергеевич** – Академический физико-технологический университет РАН, аспирант, imukhin@yandex.ru

УДК 004.9; 519.2

**ПРИМЕНЕНИЕ СУПЕРКОМПЬЮТЕРНЫХ ТЕХНОЛОГИЙ
ДЛЯ ПРОВЕДЕНИЯ ТЕПЛОвого АНАЛИЗА
РАДИОЭЛЕКТРОННЫХ УСТРОЙСТВ****И.И. Шибут**

Рассматриваются основные этапы проведения компьютерного теплового анализа на базе суперкомпьютерных технологий. Изложены теоретические и методические подходы к подготовке и проведению теплового анализа радиоэлектронных устройств в целом. Рассмотрен пример проведения теплового анализа телевизора марки «Horizont» с использованием мультипроцессорной версии пакета LS-DYNA на базе суперкомпьютера семейства «СКИФ К – 1000».

Ключевые слова: тепловой анализ, суперкомпьютер, радиоэлектронные устройства.

Введение

В современных условиях радиоэлектронные устройства различного назначения применяются во всех областях жизни человека. Поэтому в связи с постоянным стремлением к миниатюризации, повышению плотности расположения компонент внутри устройств, требований к снижению стоимости и уменьшению жизненного цикла изделия возрастает потребность в быстром и качественном решении такой проблемы, как эффективный теплоотвод в устройстве.

Наиболее приемлемым путем решения проблемы является внедрение компьютерного теплового анализа в процесс проектирования и производства радиоэлектронных устройств. Компьютерный тепловой анализ позволяет сократить не только затраты на проведение натурных испытаний, но и время, затрачиваемое на выявление недостатков и внесение изменений в конструкцию анализируемого устройства. Из-за сложности современных радиоэлектронных устройств и необходимости учета сложного комплекса физических процессов, протекающих в конструкциях, целесообразно применение суперкомпьютерных технологий для проведения теплового анализа.

В Республике Беларусь одним из лидеров в области производства радиоэлектронных устройств бытового назначения является ОАО «Горизонт». Внедрение компьютерного теплового анализа в цикл производства и проектирования продукции ОАО «Горизонт» обеспечит повышение уровня конкурентоспособности предприятия на мировом рынке. В настоящий момент из-за отсутствия квалифицированных кадров и высокой стоимости лицензионного программного обеспечения разработка новых моделей целиком основана на опыте специалистов завода. Это значительно снижает скорость разработки и запуска в производство новых моделей из-за необходимости проведения натурных испытаний и последующей доработки моделей. В то же время ведущие фирмы-изготовители радиоэлектронной продукции (Philips, Sony) успешно применяют компьютерный инженерный анализ в сквозном цикле проектирования и производства.

На базе Объединенного института проблем информатики НАН Беларуси создан и применяется для расчетов суперкомпьютер семейства «СКИФ К-1000». Для внедрения компьютерного теплового анализа на предприятии использовались возможности суперкомпьютера, что позволило значительно ускорить процесс получения результатов и повысить их точность по сравнению с применением персональных компьютеров. Исследования по тепловому анализу моделей телевизионных устройств марки Горизонт проводились на базе мультипроцессорной версии пакета LS-Dyna [1, 2].

Постановка задачи

Для видеотехники одним из важнейших показателей качества является обеспечение оптимального режима тепловыделения в процессе эксплуатации с целью удовлетворения определенным требованиям к надежности и безопасности функционирования устройства.

Традиционно оценка тепловых режимов работы любой радиоэлектронной аппаратуры на отечественных предприятиях проводилась на завершающих этапах ее проектирования посредством стендовых испытаний. Однако сложность современной аппаратуры, высокие требования к надежности требуют уже на ранних этапах разработки учета сложного комплекса физических процессов, протекающих в конструкциях РЭА. Так, отклонение тепловых режимов работы устройства ведет к изменению его эксплуатационных характеристик, вплоть до выхода из строя.

Тепловой расчет работы конструкции радиоэлектронного устройства – это анализ распределения температур в конструкции в зависимости от времени и заданных граничных условий. Источниками тепла в конструкции устройства являются различные электронные элементы. Большая часть мощности, потребляемой электронными элементами, выделяется в виде тепловой энергии. Полезная мощность при этом составляет величину порядка 10%. Тепловые поля всех электронных элементов устройства формируют общую тепловую картину, возникающую при работе устройства. Вид тепловой картины устройства зависит от мощности и распределения источников тепла, конструкции, режима работы устройства и его системы охлаждения, геометрических параметров, физических свойств материалов, из которых оно изготовлено, условий эксплуатации.

Проведение теплового анализа позволяет: определить температурные режимы работы как отдельных элементов, так и изделия в целом; выявить необходимость, параметры и эффективность применения дополнительных теплоотводящих элементов и конструкций (к примеру, введение дополнительных вентиляционных отверстий).

Методика проведения компьютерного теплового анализа радиоэлектронного устройства

Для внедрения технологии теплового анализа радиоэлектронного устройства на базе суперкомпьютерных технологий была проведена серия вычислительных экспериментов по расчету теплового режима модели телевизора в среде системы LS-DYNA [3]. Исследовалось влияние нескольких источников тепла на общую тепловую картину в корпусе телевизора.

При анализе теплового режима внутри корпуса радиоэлектронного устройства учитываются наиболее существенные черты конструкции и протекающие в корпусе физические процессы. Все второстепенные элементы конструкции, которые не влияют существенно на изучаемый объект и физический процесс теплообмена, могут быть проигнорированы. Такой идеализированный объект называют обычно тепловой моделью [4]. Основное требование к тепловой модели может быть сформулировано следующим образом: тепловая модель должна быть адекватна изучаемому явлению и реализуема математически.

Таким образом, на этапе анализа конструкторской модели необходимо выполнить следующие действия:

- провести анализ исходной CAD-модели конструкции радиоэлектронного устройства и технической документации;
- выявить основные конструктивные элементы, которые будут участвовать в процессе тепловыделения и теплообмена;
- на основе технической документации выявить источники тепла и провести сравнительный анализ тепловых характеристик тепловыделяющих элементов.

Удаление источников тепла, которые имеют малую мощность тепловыделения (вторичных), позволяет упростить тепловую модель устройства, снизить общий объем памяти для вычислений, а также время расчета. Удаление вторичных источников не приводит к значительному искажению тепловой модели, так как их тепловое поле локализовано в малой окрестности элемента и имеет лишь местное значение.

Следует также отметить, что в конструкции существуют элементы, не вносящие существенных изменений в общую тепловую картину. Это различные ребра жесткости, выступы, кнопки управления, разъемы для аудио- и видеовходов, внешние элементы, такие как стойки, кронштейны или другие устройства крепления. Удаление таких элементов снижает объем модели и упрощает ее обработку. На рис. 1 приведен пример изменения конструкторской модели телевизора для проведения теплового анализа.

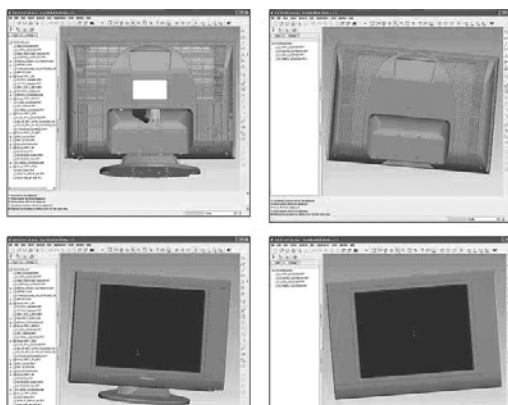


Рис. 1. Исходная (слева) и модифицированная (справа) модели телевизора

Для работы с конструкторской моделью устройства в САЕ-среде необходимо создание конечно-элементной модели (КЭМ), на базе которой и будет проводиться анализ. Существует несколько базовых принципов построения КЭМ, которые приведены ниже.

Конечно-элементная сетка должна генерироваться из конечных элементов, размеры которых находятся в соответствии с размерами геометрической модели. Минимальный размер конечного элемента определяется минимальным размером элементов геометрической модели и наличием в ней мелких конструктивных элементов. При разбиении следует стремиться к равномерности размеров ячеек конечных элементов. Сгенерированная сетка конечных элементов должна отвечать требованию непрерывности: в ней должны отсутствовать «дыры», дублирующие элементы, пересекающиеся элементы. Сгенерированная конечно-элементная сетка не должна содержать конечных элементов низкого качества. Скорость и точность расчетов зависит от качества конечных элементов. При построении сетки конечных элементов необходимо стремиться, чтобы она состояла из правильных треугольников и четырехугольников, также должны отсутствовать треугольники с малыми углами.

Следующим этапом работ по проведению инженерного анализа является создание расчетно-аналитической модели [5]. Под расчетно-аналитической моделью понимается твердотельная и (или) поверхностная модель исследуемой конструкции, представленная конечными элементами и их узлами, к граням которой, или к узлам конечных элементов, прикладываются нагрузки и ограничения.

Для описания задачи теплового анализа радиоэлектронного устройство необходимо:

- задать параметры материалов, из которых изготовлены элементы конструкции устройства;
- определить начальные и граничные условия;

- задать тепловые нагрузки;
- задать тепловой контакт между деталями сборки модели устройства;
- задать параметры расчета на BMBC.

Создание расчетно-аналитической модели телевизора проводилось в среде пре-, постпроцессора LS-PrePost.

После передачи КЭМ устройства в среду LS-PrePost на первом этапе проверяется содержание базовых опций *Part, *Mat и *Section. Опция *Section определяет формулировки конечных элементов, правило их интегрирования, толщину элементов в узлах, а также характеристики поперечного сечения. В некоторых случаях, например, при импорте конечно-элементных моделей из среды HyperMesh, сведения о конечных элементах не передаются. Поэтому следует задавать значения параметров опции *Section вручную. Таким образом, первым шагом при создании расчетно-аналитической модели является задание параметров опции *Section.

Описание параметров материала осуществляется в разделе ключевого слова *Mat. Опция позволяет задавать разные типы материалов для различных видов анализа. В случае теплового анализа используется тип материала *Mat_Thermal{Option}:

- в случае использования при изготовлении устройства изотропных материалов применяется тип *Mat_Thermal_Isotropic. Данный тип используется при описании теплофизических характеристик материалов класса пластмасс, металлов, стекла;
- в разделе задаются такие параметры, как плотность, теплопроводность и теплоемкость материала.

Описание частей модели производится при помощи ключевого слова *Part. При импорте сборки конечно-элементных моделей устройства для каждой отдельной конечно-элементной модели по умолчанию создается карта *Part, которую необходимо заполнить вручную:

- каждой части модели поставить в соответствие идентификатор ключевого слова *Section;
- для каждой части определить идентификатор материала в поле TMID.

Известно, что температура телевизора в начальный момент времени одинакова во всех узлах модели. В связи с этим в набор узлов для задания начальных условий должны входить все узлы модели.

Начальные условия для теплового расчета описываются картой ключевого слова *Initial, задается начальная температура устройства. У телевизора, как было упомянуто выше, все узлы имеют одинаковую начальную температуру. Поэтому удобно вместо идентификатора набора узлов в поле nsid поставить 0, что по умолчанию присваивает значение заданной начальной температуры всем узлам модели.

Граничные условия описываются ключевым словом *Boundary_Convection_Set. Каждому сегменту ставится в соответствие температура окружающей среды и коэффициент конвективной теплоотдачи.

В качестве тепловой нагрузки в модели задаются коэффициенты объемного тепловыделения. Постоянная тепловая нагрузка задается в поле smult ключевого слова *Load_heat_generation_set. Коэффициент объемного тепловыделения рассчитывается по формуле $Q=p/V$, где p – мощность тепловыделения, V – объем детали. В качестве источников тепловыделения в модели телевизора задаются наборы твердотельных элементов конечно-элементных моделей процессора обработки сигналов и трансформатора.

Для описания процессов теплообмена в модели используется ключевое слово *Contact_automatic_surface_to_surface(Thermal)_MPP. С помощью опции задается тепловой контакт между деталями сборки, находящимися в непосредственном контакте, а также между деталями, разделенными воздушной прослойкой.

Для описания контакта частей модели, между которыми расположена воздушная прослойка, используется поле *cf*, в котором задается удельная теплопроводность воздуха. Поле *frad* используется для описания передачи тепла радиацией, однако при температурах менее 100 °С данным видом теплообмена можно пренебречь.

Параметры расчета на суперкомпьютере задаются при помощи ключевого слова **Control*:

- процедура выполнения только теплового анализа задается ключевым словом **Control_Solution*;
- время завершения анализа задается в поле *ENDTIM *Control_Termination*;
- в полях **Control_Thermal_Solver* задаются опции теплового расчета, такие как тип анализа, тип задачи, тип решателя;
- параметры управления шагом решения по времени задаются в ключевом слове **Control_Thermal_Timestep*.

В поле *SOLN *Control_Solution* определяется вид анализа как только тепловой, что соответствует значению *SOLN=1*.

Тип теплового анализа задается в поле *ATYPE*. В случае теплового анализа монитора выбираем переходный тип теплового анализа *ATYPE=1*. Такой анализ проводится с помощью неявных методов.

Результаты исследования

На основе предоставленной информации моделирование процессов распространения тепла в корпусе одной из моделей телевизора проводилось при следующих ограничениях:

- равномерное распределение температур в начальный момент времени, $T_0 = 35^\circ\text{C}$;
- мощность тепловыделения LCD -матрицы телевизора $P=36$ Вт;
- мощность процессора обработки сигналов $P=3$ Вт;
- мощность трансформатора питания $P=12$ Вт;
- коэффициент теплоотдачи конвекцией с поверхности корпуса и кожуха 2,72 Вт/(м²с);
- коэффициент теплоотдачи конвекцией с поверхности источников тепла 6 Вт/(м²с).

Для запуска расчета на ВМБС было задействовано 4 узла по 2 процессора. Следует отметить, что при запуске решения задачи теплового анализа на суперкомпьютере необходимо установить двойную точность расчетов.

Результаты теплового анализа телевизора, проведенного по предложенной методике, представлены на рис. 2.

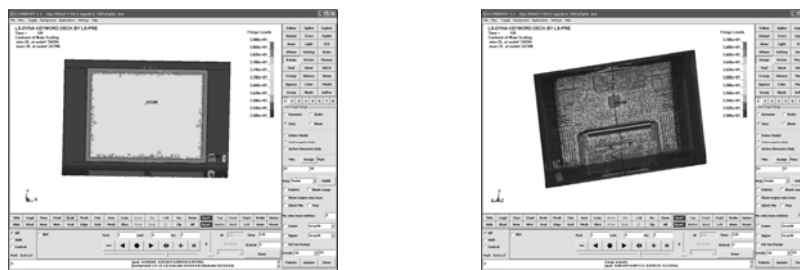


Рис. 2. Результаты анализа теплового состояния корпуса телевизора

В результате проведения вычислительных экспериментов по тепловому расчету в среде Ls-Dyna были получены значения температур наиболее значимых источников тепла, таких как LCD-матрица телевизора, трансформатор питания и процессор обработ-

ки сигналов. Также было получено распределение температур на поверхности корпуса телевизора с учетом тепла, выделяемого перечисленными источниками. При проведении сравнительного анализа полученных данных и данных по натурным испытаниям было установлено, что рассчитанные температуры процессора и трансформатора близки к данным натурных испытаний с погрешностью от 3,5% до 6,7%. Погрешность расчетов связана со сложностью моделирования параметров материала тепловыделяющих элементов.

На основе анализа результатов были вынесены практические рекомендации по оптимизации конструкции телевизионного устройства.

Заключение

Разработанная методика создания расчетно-аналитической модели радиоэлектронного устройства позволяет в полной мере описать процессы теплообмена в устройстве с естественным способом охлаждения и низкой интенсивностью движения среды в корпусе. Методика проведения теплового анализа была успешно внедрена в процесс проектирования телевизионных устройств на предприятии «Горизонт» и позволила разработчикам повысить эффективность подготовки новых моделей телевизионных устройств к запуску в производство. Планируется, что дальнейшая работа в области внедрения компьютерного инженерного анализа на базе суперкомпьютерных технологий в процесс производства радиоэлектронной продукции в Беларуси позволит устранить погрешности, связанные с недостатком информации об импортируемых комплектующих, и снизить процент ошибки компьютерных расчетов.

Результаты, представленные в работе, получены в рамках выполнения задания ПА2.2 «Разработать и внедрить научно-методические основы компьютерного инженерного анализа конструкций аудио-, видеотехники на кластерных мультипроцессорных вычислительных системах с использованием пакетов мирового уровня» проекта ПА2 «Разработка научно-методических основ применения наукоемких информационных САЕ-технологий на базе кластерных высокопроизводительных мультипроцессорных вычислительных систем» научно-технической программы Союзного государства «Триада».

Литература

1. LS-DYNA Keyword User's Manual. Volume 1. – Livermore Software Technology Corporation, 2001. – 853 p.
2. Shapiro Arthur B. Heat Transfer in LS-DYNA // Proc. of the 4th European Users Conference.
3. Губич Л.В., Шибут И.И. Тепловой анализ телевизионного устройства в среде системы LS-DYNA: подходы, методы, решения // Вестник компьютерных и информационных технологий. – 2008. – № 9. – С. 20–25.
4. Дульнев Г.Н. Тепло- и массообмен в радиоэлектронной аппаратуре. – М.: Высшая школа, 1984. – 248 с.
5. Шибут И.И. Формирование расчетно-аналитической модели для расчета теплового состояния радиоэлектронных устройств // Материалы второй Междунар. науч.-техн. конф. «Суперкомпьютерные системы и их применение» (SSA'2008). – Минск: ОИПИ НАН Беларуси, 2008. – С. 306–311.

Шибут Ирина Игоревна – Объединенный институт проблем информатики НАН Беларуси, аспирант, irina_shibut@mail.ru

УДК 681.3

АППАРАТНО-ПРОГРАММНОЕ ПРОЕКТИРОВАНИЕ СЛОЖНЫХ
ФУНКЦИОНАЛЬНЫХ БЛОКОВ С ИСПОЛЬЗОВАНИЕМ
СИСТЕМ НА КРИСТАЛЛЕ

И.А. Зубакин, Ш.С. Фахми, С.С. Шагаров

С увеличением возможностей технологии проектирования СБИС класса «система на кристалле» становится все более привлекательным и необходимым средством создание интегрированных САПР БИС/СБИС с программируемой структурой. В статье рассматриваются некоторые этапы проектирования (аппаратная и программная части проекта схемы) и верификации схемы в целом на примере разработки умножителя с использованием аппаратно-программных средств САПР на базе технологии «система на кристалле».

Ключевые слова: система на кристалле, уровень регистровых передач, сложный функциональный блок, программируемая логическая интегральная схема.

Введение

В настоящее время сложность систем на кристалле (СнК) и разновидности инженерных решений микропроцессорных систем на базе программируемых логические интегральных схем (ПЛИС) привели к бурному развитию методологии проектирования. Методология сопряженного проектирования – методология, которая позволяет архитекторам системы и инженерам верификации работать на более высоком уровне абстракции. Трудности проектирования СнК связаны как с необходимостью выполнять проекты в сжатые сроки, так и с постоянно возрастающей сложностью систем. По существующим оценкам средний объем современных проектов превышает сотни тысяч вентилей и имеет тенденцию постоянного роста, и в окончательной стоимости проекта все большую роль играют разработанные программные и аппаратные ресурсы. Возрастание сложности современных проектов как в части объемов аппаратного обеспечения (количество логических вентилей), так и в части объемов программного обеспечения (объем программного кода) делает проблему технологической, методической и методологической поддержки данного направления актуальной [1, 2].

СнК образуют динамический, быстро растущий сектор рынка. Интеграция все большего числа сложных функциональных блоков (СФ-блоков) в одном кристалле сопровождается сокращением площади плат при реализации проектов, повышением быстродействия систем и их надежности, уменьшением потребляемой мощности и стоимости. Возможность быстрой реконфигурации схем непосредственно в работающем устройстве открывает перспективы не только эффективной отработки прототипа проекта, но и создания принципиально новых структур с динамическим реконфигурированием и многофункциональным использованием аппаратных и программных средств. Все перечисленные реалии и отсутствие материала по предметной области делает проблему методической поддержки сопряженного проектирования СнК востребованной и нуждающейся в популяризации.

Средства проектирования

Для сопровождения проектирования цифровых устройств обработки сигналов на основе перспективной технологии СнК различные производители предлагают интегрированные программные средства проектирования. В данной статье рассмотрена среда FastChip, которая представляет программный пакет, обеспечивающий управление всем

процессом реализации устройств на основе технологии СнК. Среда FastChip имеет характерные черты других подобных САПР, а потому ее возможности представляют интерес для анализа и исследования. На современном этапе наблюдается тенденция специализации средств автоматизированного проектирования электронных изделий, что является важнейшим условием эффективности разработки и ускорения выпуска продукции. Такой пример сочетания специализации, функциональности и открытости современных САПР демонстрирует указанная среда разработки.

Новизна СнК заключается в одновременной аппаратно-программной реализации на одном кристалле значительной (если не всей) функционально законченной части аппаратуры. Поэтому средства моделирования системного уровня как самой аппаратуры, так и кристаллов СнК одни и те же. За рубежом сейчас действуют сотни дизайн-центров при аппаратостроительных фирмах и масса отдельных дизайн-центров. В них соразработчиками элементной базы выступают разработчики аппаратуры. Они либо доводят проект до .edif файла, который передается разработчикам кристаллов, либо делают только системный проект до уровня регистровых передач (RTL-уровня) [3].

Кроме того, поскольку СнК – изделие не только очень сложное, но и дорогое при небольших объемах выпуска, ошибки проектирования должны быть исключены. Поэтому в новой методологии, кроме математического моделирования, необходим этап создания прототипа устройства на базе неких специальных аппаратно-программных платформ, которые сегодня так и называются – алгоритмически ориентированные платформы проектирования. Они содержат набор элементов: ПЛИС, процессорные ядра, память, определенную шинную архитектуру, интерфейсы и т.д., которые позволяют создать прототип схемы до ее воплощения в СБИС. А дальше разработчик решает, остановиться ему на уровне ПЛИС или производить специализированные СБИС на основе СнК [4].

В настоящее время СнК включает в свой состав несколько миллионов эквивалентных вентилях, разнообразные типы процессоров, в том числе и цифровой обработки сигнала, т.е. представляет собой большую сложную систему. Это обстоятельство заставляет при всей общности методологии проектирования уточнять ее на примере конкретной СнК, в качестве которой в данной статье выбрана реконфигурируемая СнК фирмы Triscend (ZyLogic) [5, 6].

Целью данной статьи является описание маршрута сопряженного проектирования аппаратно-программных частей СФ-блоков на основе 32-разрядной СнК, входящей в состав лабораторного макета [7], и отобразить основные проектные процедуры верификации и отладки на примере цифрового устройства умножения целых чисел. Также в статье показано, что весь процесс проектирования проводится как начальный этап, а верификация и соиммуляция – это последующие самостоятельные этапы. Хотя имеются разные трактовки сопряженного проектирования, но реально совместные процедуры выполняются в различных комбинациях на различных этапах проектирования [8].

Следует отметить, что идеология разработки устройств на основе технологии «реконфигурируемая СнК» отличается от проектирования устройств в многокристальном исполнении. Это связано с тем, что реконфигурируемые СнК представляют собой приборы системного уровня интеграции, т.е. сочетают на одном кристалле как разнородные аппаратные структуры (процессор, программируемая логика, оперативная память), так и архитектурные элементы их взаимодействия (адресные селекторы, конфигурируемая системная шина). Специфика разработки цифровых устройств на основе реконфигурируемой СнК заключается в интеграции с помощью САПР двух потоков проектирования:

- реализации аппаратной составляющей системы на основе конфигурируемой системной логики CSL (ячейки матрицы FPGA и средства соединений между ними);

- разработки программной составляющей системы для встроенного микропроцессорного ядра, используя ресурсы как встроенных периферийных элементов, так и возможности интерфейса с конфигурируемой логикой.

Сопряженное проектирование – процесс параллельного проектирования аппаратных и программных средств, при котором оценивается целесообразность выбора аппаратной или программной реализации определенного фрагмента проекта. Этот процесс также позволяет проектировщикам увидеть, как система могла бы работать с данным разделением аппаратных средств ЭВМ и программного обеспечения [9].

Сопряженная верификация – проверка возможности совместного функционирования программного обеспечения и аппаратных средств ЭВМ. С другой стороны, во время этого процесса анализируется, будут ли одинаково решаться специфические задачи проекта при аппаратной или программной реализации. Результатом работы является фиксация выбранного варианта реализации.

Прототипирование аппаратно-программных частей проекта

В основу прототипирования положен простой принцип: необходимо создать поведенческую модель СФ-блока до начала детальной разработки аппаратной и программной частей проекта. При этом прототип должен представлять собой точную (до цикла) поведенческую модель всего СФ-блока, включая его аппаратную и программную части [10].

Рассмотрим цифровое устройство умножения целых чисел (рис. 1), включающее в себя три основных блока:

- а) программная часть проекта для ввода исходных данных в составе двух программно доступных регистров, доступ к которым осуществляется с помощью средств компиляции, позволяющих поддерживать все стадии разработки;
- б) аппаратная часть проекта, реализованная на языке VHDL, для выполнения операции умножения и импортированная в САПР;
- в) блок вывода результатов умножения на индикацию, реализуемый на элементах из библиотеки СФ-блоков.

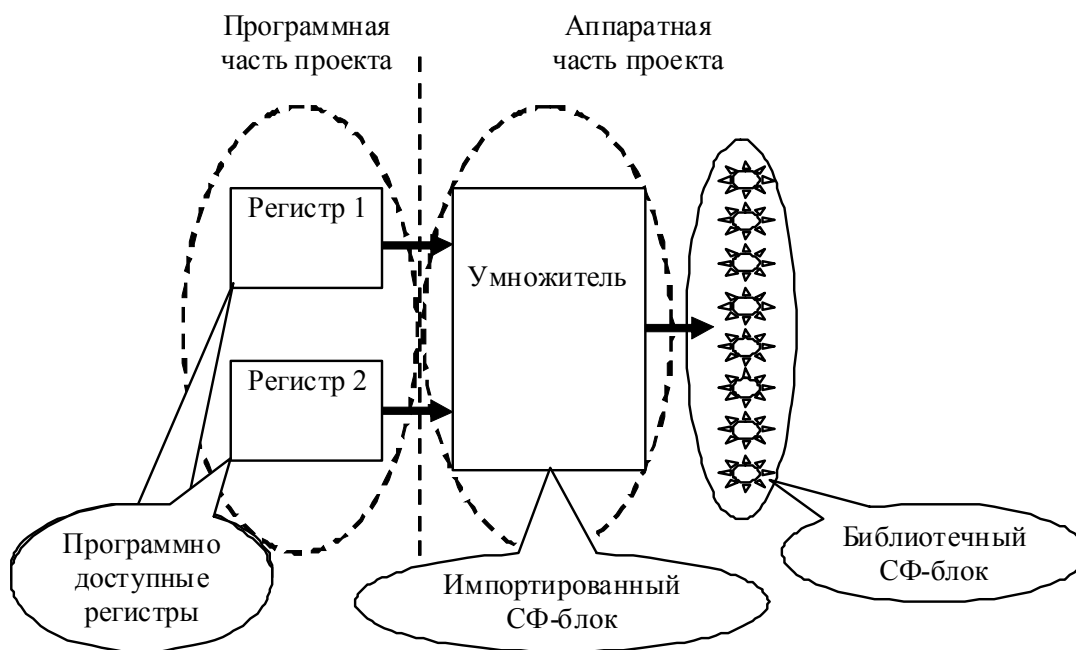


Рис. 1. Цифровое устройство умножения целых чисел

Разработка аппаратной части проекта. Для разработки аппаратной конфигурации САПР FastChip обеспечивает возможность интерактивного выбора составных фрагментов схемы и средств задания соединений между ними. Применение САПР предполагает эффективные, наглядные и контролируемые средства описания проекта, при этом наиболее распространенными являются графический и текстовый (рис. 2).

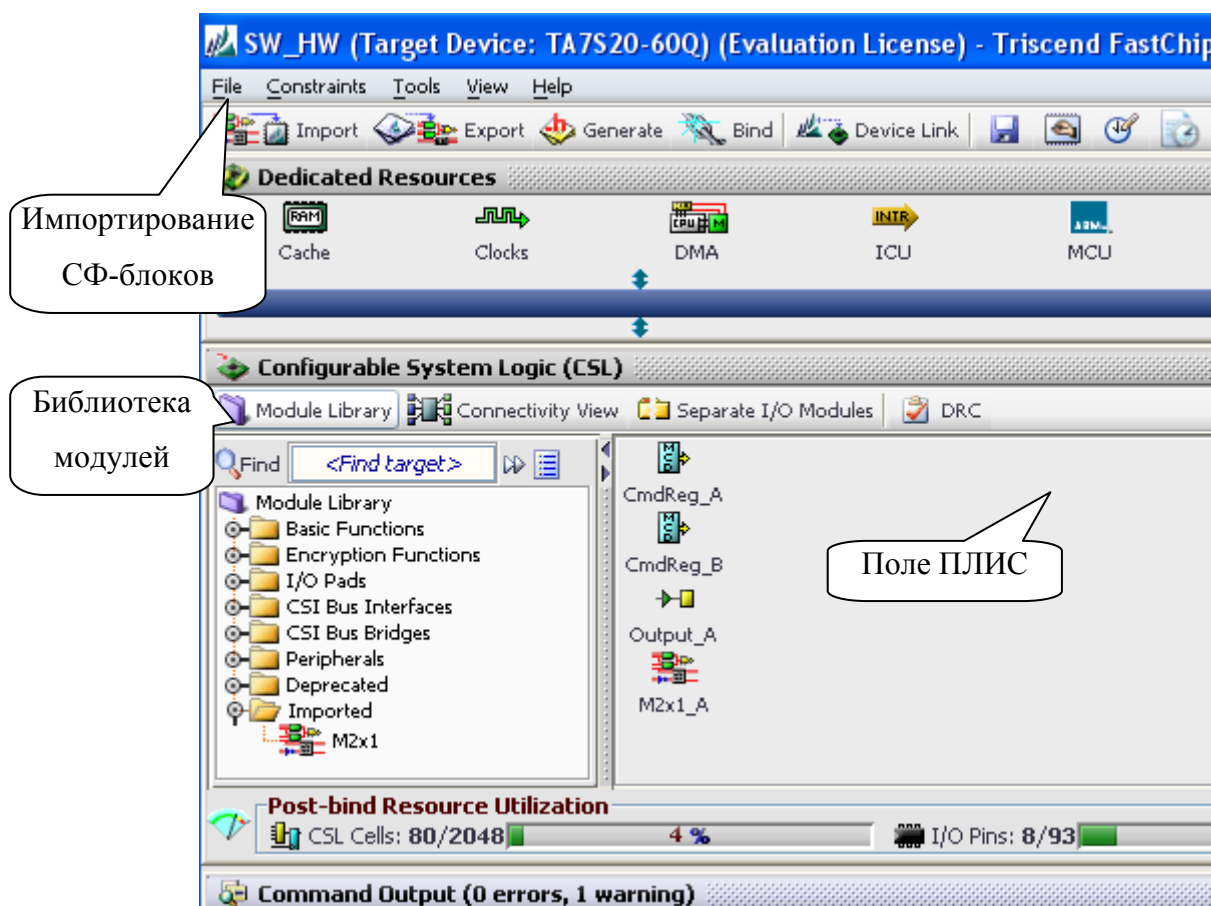


Рис. 2. Графические средства описания проекта

Средства описания проекта САПР FastChip включают как собственные средства (библиотечные элементы), так модули, создаваемые пользователем. Набор предлагаемых средой FastChip предопределенных элементов достаточно широк (около 150), постоянно расширяется, и его можно разделить на следующие группы:

- типовые параметризуемые элементы вычислительной техники (логические элементы, мультиплексоры, счетчики и т. д.), конкретные параметры которых (разрядность, коэффициент пересчета, начальное состояние) могут назначаться разработчиком;
- специализированные элементы взаимодействия микропроцессорной части реконфигурируемой СнК и матрицы программируемой логики (адресные селекторы, элементы системной шины, схемы синхронизации и ожидания);
- специфичные для реконфигурируемой СнК составляющие (буферные элементы, порты ввода-вывода);
- типовые узлы вычислительных систем (контроллеры интерфейсов, графических индикаторов и т. д.), часть параметров которых варьируется проектировщиком.

Библиотечные модули, от простых (логические вентили, сумматоры, умножители, компараторы, счетчики, элементы памяти) до сложных (блок шифрования DES), пред-

ставляют собой готовые к использованию виртуальные элементы принципиальной схемы, являющиеся «блоками интеллектуальной собственности – IP». Такие параметризируемые конструкции в терминологии САПР FastChip называются «Soft Modules», что отражает возможность их легкой модификации разработчиком. Функции модуля выполняются аппаратно с помощью внутрикристальной программируемой логики. Следует отметить, что библиотечные модули не содержат никакой информации о своей компоновке внутри FPGA, следовательно, их быстродействие будет зависеть от конкретного варианта размещения, трассировки и загрузки кристалла.

После определения необходимых для решения задачи библиотечных модулей разработчик размещает выбранные блоки в рабочей области (проектном окне) экрана САПР и при необходимости определяет их символические имена. Далее проектировщик осуществляет настройку изменяемых параметров выбранных модулей, например, разрядность счетчика, начальное состояние регистров и т.д., путем установки конкретных числовых значений в полях конфигурации модуля. Среда FastChip оперативно отражает оценку затрат на реализацию (число требуемых логических ячеек, портов ввода-вывода, адресных селекторов) в зависимости от сложности проекта.

Синтез умножителя на языке VHDL. Наиболее универсальное решение в области синтеза принадлежит компании Synplicity (САПР Synplify), которое покрывает все связанные с этим аспектом области проектирования, а именно логический и физический синтез для ПЛИС любых производителей, макетирование специализированных устройств на ПЛИС, а также средства отладки прошивок на уровне исходного RTL описания и написанные на языке VHDL.

Операция умножения, независимо от способа реализации, требует использования $n \times (m - 1)$ полных одноразрядных сумматоров, где m – число разрядов сомножителей. Реализация умножения с использованием n -разрядных сумматоров порождает широкую номенклатуру параллельных умножителей, в том числе матричную и пирамидальную структуры и смешанные варианты [3]. Устройства этого класса способны выполнить умножение за один такт работы схемы. Подобными характеристиками обладают схемы, построенные на основе набора параллельных умножителей меньшей разрядности с последующим суммированием частичных произведений [3], однако вопросы, связанные с синтезом устройства умножения, не являются целью статьи, а реализация предлагается с использованием встроенной в САПР функции умножения.

Упрощенная программа реализации операции умножения:

```
library IEEE;
use IEEE.std_logic_1164.all;
use IEEE.std_logic_arith.all;
use IEEE.std_logic_unsigned.all;
entity Mul2x1 is
port ( A, B      : in std_logic_vector(7 downto 0);
      PP        : out std_logic_vector(15 downto 0));
end Mul2x1;
architecture AMult of Mul2x1 is
signal sA, sB   : unsigned (7 downto 0);
signal sP       : unsigned(15 downto 0);
begin
    sA <= unsigned(A);
    sB <= unsigned(B);
    sP <= sA * sB;
    PP <= std_logic_vector(sP);
end AMult;
```

По описанию на уровне RTL с помощью программы логического синтеза формируется список цепей (gate level netlist), который учитывает задержки в библиотечных элементах (но, как правило, не учитывает временные задержки между соединениями) и используется для временной верификации проекта (timing verification). Полученную в результате RTL синтеза схему (рис. 3) необходимо импортировать в среду САПР для дальнейшей загрузки в ПЛИС (рис. 2).

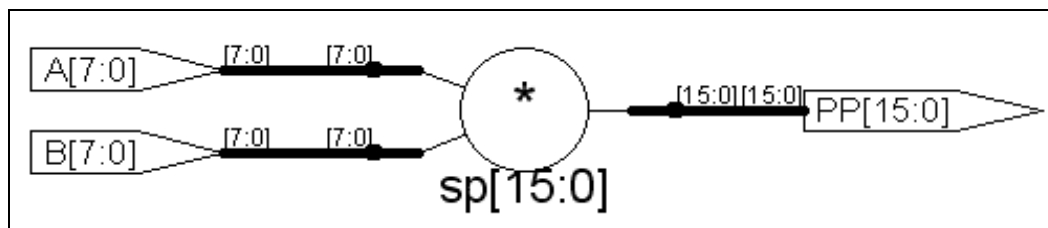


Рис. 3. Схема, получаемая в результате синтеза RTL

Разработка программной части проекта. Программная часть содержит два программно доступных регистра: Command Registr 1 и 2, аппаратная часть включает два модуля (рис. 2): библиотечный модуль Output_A для вывода результата умножения на индикацию и импортированный модуль (в виде .edif файла) M2x1_A, выполняющий операцию умножения целых чисел. Далее выполняются две процедуры, необходимые для верификации проекта: подключение входов и выходов проекта к выводам СнК в редакторе I/O Editor и связывание модулей (Bind) (рис. 2).

Наличие процессора в составе СнК TA7S20-60 позволяет объединить в единый проект аппаратную и программную части. Для этого необходимо воспользоваться средствами программной компиляции, представленными фирмой-производителем САПР. В данном устройстве практически все функции выполнены аппаратно, поэтому код самой программы достаточно прост, и в него входит лишь инициализация программных регистров и записи в них значений.

Первоначально необходимо сгенерировать так называемый заголовочный файл (header-файл). Для этого нужно в панели инструментов в среде FastChip выбрать команду Generate, которая его создаст. Теперь в папке с проектом появился файл с расширением *.h, где * – название проекта.

После этого в папку с проектом нужно скопировать несколько файлов для осуществления компиляции и верификации программной части (main.c, init.s, main.elc и make) из каталога, представленного САПР:

```
#include "HW_SW.h"
#define Reg_A 0x100ffffc /* size in byte(s) = 1, lane = LANEBO */
#define Reg_B 0x100ffff8 /* size in byte(s) = 1, lane = LANEBO */
volatile unsigned int *result1 = ( unsigned int * ) Reg_A;
volatile unsigned int *result2 = ( unsigned int * ) Reg_B;
void init (void) {}
void IRQ_Handler (void) {}
// функция задержки
void wait(unsigned int cnt)
{ volatile unsigned int i;
  for (i = 0; i < cnt; i++) { volatile int ii;
    for(ii=0; ii < 1; ii++) {#ifdef __ARM__ __asm {mov r0, r0
      mov r0, r0
      mov r0, r0
```

```

        mov    r0, r0} #else asm("mov    r0, r0");
                                asm("mov    r0, r0");
                                asm("mov    r0, r0");
                                asm("mov    r0, r0");
        #endif }    }    }

void C_Entry ()
{    init ();    int index=0;    while (1){*result1 =0x02; *result2 =0x03; }}
    
```

Следующий этап процесса разработки программной части – компиляция. Для этого в среде FastChip используется свободно распространяемый компилятор Unix-систем. Чтобы его использовать, надо зайти в эмулятор Unix, который запускается из папки «Triscend SDK 2.6», и в этой папке запустить ярлык «SDK Build Window», после чего появится командная строка. Здесь нужно перейти в каталог с проектом и дать команду make. Если все правильно, то компилятор выдаст сообщение, пример которого представлен на рис. 4.

```

c:/Proj_Shakeeb/HW_SW
$ make
arm-elf-gcc -mcpu=arm7tdmi -I../include -I../ -I"C:/p/include" -ggdb -g2 -mapcs-32 -c -msoft-float -mstructure-size-1
in.o main.c
arm-elf-ld -o main.elf init.o main.o -N -T main.ld -Map main.map
arm-elf-objcopy --set-start 0x00000000 -O ihex main.elf main.hex
Администратор@WS-9 c:/Proj_Shakeeb/HW_SW
$
    
```

Рис. 4. Пример сообщения компилятора

Проектирование и верификация схемы в целом

Концепция прототипирования не может быть реализована без интегрированной среды разработки, позволяющей:

- использовать существующие СФ-блоки независимо от языка или уровня абстракции их описания и разрабатывать новые СФ-блоки;
- собирать прототип из СФ-блоков;
- обеспечить взаимодействие встроенных коммуникационных блоков с прототипом;
- анализировать поведение всей системы.

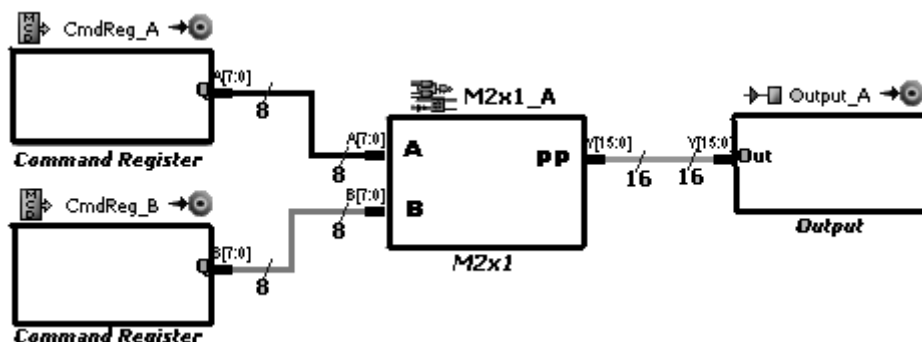


Рис. 5. Схема умножения чисел

Кроме того, поскольку стандартным языком описания системных моделей выступает VHDL, крайне важно, чтобы среда прототипирования избавляла разработчиков от

необходимости вручную переписывать проект системы, а вопросы интегрирования и совместимости были реализуемы в выбранной САПР. В результате сборки всех частей проекта и импортирования .edif файла умножителя в среду САПР FastChip и выполнения процедуры компиляции программным способом были записаны два числа (2 и 3) и произведено их умножение аппаратно с помощью синтезируемой схемы умножителя (рис. 5). Результат выполнения операции выдается на 7-сегментный индикатор через устройство вывода (Output), являющееся элементом библиотеки используемой САПР.

Заключение

Появление наряду с СнК видеосистем на кристалле (ВСнК) не только расширило сферу применения прикладного телевидения (например, благодаря возможности создания телевизионной системы весом несколько граммов, получили распространение беспилотные летательные аппараты, стартующие с ладони), но и сменило метод проектирования прикладных систем, еще более срастив его с проектированием главных систем информатики – компьютеров. Если ранее местом соединений, задаваемых разработчиком, была печатная плата, то современные технологии направлены на реализацию максимально возможного, в пределе – всего многообразия соединений внутри одного кристалла. Это заставляет заново оценивать и пересматривать, дополнять сложившиеся методы проектирования новыми средствами быстрого физического прототипирования ВСнК на различных стадиях процесса проектирования. Такие возможности резко повышают значимость современных САПР и обязательно должны быть поддержаны на этапе подготовки специалистов, что и делается сейчас в рамках программы подготовки специалистов для высокотехнологичных отраслей промышленности.

Перспектива развития высокоуровневого проектирования состоит в том, что Verilog или VHDL способен обеспечить моделирование на системном уровне. Компания Cadence предоставляет все необходимые составляющие для виртуального прототипирования и совместной верификации аппаратной и программной частей на основе виртуального прототипа, причем маршрут проектирования Cadence предусматривает автоматический переход с уровня транзакций к синтезируемым моделям напрямую из VHDL-кода.

На кафедре САПР СПбГЭТУ «ЛЭТИ» разработан макет на базе технологии «система на кристалле», включающий в свой состав 32-разрядный ARM процессор (семейства A7), память, ПЛИС и другие периферийные модули [3]. Макет позволяет осуществить физическое прототипирование систем с применением схем с программируемой логикой, а также выполнить верификацию как на функциональном, так и на вентиляционном уровнях.

Литература

1. Грушвицкий Р.И., Мурсаев А.Х., Угрюмов Е.П. Проектирование систем на микросхемах программируемой логики. – СПб: БХВ-Петербург, 2002.
2. Немудров В.Г., Мартин Г. Системы на кристалле. Проблемы проектирования и развития. – М.: Техносфера, 2004.
3. Березин В.В., Фахми Ш.С. Проектирование устройств обработки сигналов на основе технологии «система на кристалле». – СПб: СПбГЭТУ «ЛЭТИ», 2005.
4. Немудров В. Г. «Прогресс» в проектировании отечественных СБИС // Электроника: наука, технология, бизнес. – 2002. – № 6. – С. 5–9.
5. Березин В.В., Золотуха Р.Н., Фахми Ш. С. Отладка аппаратно-программного обеспечения реконфигурируемых систем на кристалле // Компоненты и технологии. – 2003. – № 7. – С. 118–122.

6. Березин В.В., Фахми Ш.С. Аппаратно-программные средства для проектирования цифровых устройств. – СПб: СПбГЭТУ «ЛЭТИ», 2005.
7. Березин В.В. Золотуха Р.Н. 32-разрядная реконфигурируемая система на кристалле фирмы Triscend // Компоненты и технологии. – 2003. – №2. – С. 14–20.
8. Березин В.В. Методология автоматизированного проектирования с применением технологии «система на кристалле» (программная часть и системная интеграция) // Промышленные контроллеры и АСУ. – 2004. – № 12. – С. 38–40.
9. Березин В.В., Фахми Ш.С. Использование возможностей контроллера динамической памяти как составной части «системы на кристалле» // Неразрушающий контроль и диагностика окружающей среды, материалов и промышленных изделий: Межвузовский сборник СЗТУ. – 2004. – № 9. – С. 213–225.
10. Фахми Ш.С. Автоматизация проектирования БИС на базе «система на кристалле». – СПб: СПбГЭТУ «ЛЭТИ», 2006.

Зубакин Игорь Александрович

– Санкт-Петербургский государственный электротехнический университет «ЛЭТИ», аспирант, zubakin_@mail.ru

Фахми Шакиб Субхиевич

– Санкт-Петербургский государственный электротехнический университет «ЛЭТИ», кандидат технических наук, доцент, shakeebf@mail.ru

Шагаров Сергей Сергеевич

– Санкт-Петербургский государственный электротехнический университет «ЛЭТИ», студент, shagarovss@mail.ru

УДК 004.421.2:519.178

ОПТИМИЗАЦИЯ АЛГОРИТМОВ ИДЕНТИФИКАЦИИ ГРАФОВОГО ИЗОМОРФИЗМА

С.В. Москаленко, Ю.А. Гатчин

Предложен ряд оптимизирующих моментов для современных алгоритмов идентификации графового изоморфизма, призванных увеличить скорость работы данных алгоритмов с сохранением точности вычисления результатов.

Ключевые слова: теория графов, распознавание графических образов, определение межграфового изоморфизма.

Введение

В последнее время растет объем обрабатываемой инженерной документации, в частности, приборостроительных и машиностроительных чертежей. Эта тенденция возникла в связи с активизацией использования электронной версии инженерной документации в САПР [1]. Довольно частой практикой для инженеров и проектировщиков является обращение к уже существующей документации, однако данный процесс всегда оказывается трудоемким, так как требует просмотра всей базы чертежей. Для упрощения поиска часто используется текстовая составляющая документа, когда происходит просмотр по ключевым словам. Такой подход весьма эффективен, но существуют трудности, связанные с получением этой текстовой информации (необходимо вмешательство человека), кроме того, набор ключевых слов не всегда точно может описать содержание чертежа. Поиск требуемого изображения в БД есть не что иное, как проблема сопоставления графических объектов, которая может быть решена путем вычисления степени соответствия входного объекта объекту в базе данных (БД) и сведена к процессу сопоставления графов. Для этого инженерная документация сначала приводится к представлению атрибутивных графов, где вершины графов отображают примитивы, полученные из исходных документов (линии, кривые) [2], а топологические связи

между этими примитивами описаны ребрами графов. Далее вершины графов задаются метками, которые вычисляются по некоторой функции. Метки, как правило, являются целыми числами и используются для сравнения входных графов с заданными в БД графами. Таким образом, сравнивая ребра и метки графов, можно обнаружить чертежи, находящиеся в БД, целиком или частично схожие с текущим чертежом.

В статье рассматривается подкласс процесса сопоставления графов – проверка графов на изоморфизм (далее графовый изоморфизм или ГИ). ГИ можно определить следующим образом: два графа $G=(V_1, E_1)$ и $H=(V_2, E_2)$ изоморфны, если между парами множеств их вершин, ребер и дуг существуют взаимно однозначные соответствия, сохраняющие смежность и ориентацию для дуг.

Понятие ГИ находит применение в различных направлениях – при определении поврежденных клеток на медицинских снимках, поиске изоморфных химических соединений [3], сравнении пар кинематических цепей [4], и т.д. Для решения проблемы поиска ГИ, однако, не существует достаточно эффективных алгоритмов, обладающих полиномиальным временем исполнения, даже если алгоритм ограничен по области применения (например, операции проходят над строго однородными графами). В работе [5] представлен эффективный алгоритм, обладающий сложностью вычисления $O(n^{1/3} \log n)$, но оперирующий лишь со строго однородными графами. В работе [6] приведен еще один алгоритм, являющийся линейным для большинства случайных графов. Однако последний имеет ряд недостатков, которые связаны с невозможностью разрешения неопределенных ситуаций, возникающих при обнаружении множества идентичных друг другу вершин. В рамках данной статьи предложены 3 оптимизационных момента, призванные ускорить поиск ГИ в современных алгоритмах сопоставления графов.

Реализация алгоритма

Сформулируем обобщенный алгоритм работы многих современных систем идентификации ГИ [6, 7]. На старте алгоритм имеет некоторый входной помеченный граф (вершинам которого приписаны метки), а также база данных (БД) эталонных помеченных графов. Требуется найти граф в БД, изоморфный входному графу (см. определение 1, с. 101). Вначале алгоритм строит связи между всеми вершинами входного графа со всеми вершинами эталонного графа в БД, которые имеют одинаковые метки. Ситуацию, когда одна вершина в одном графе идентична с несколькими вершинами в другом, называют неопределенной. На следующих шагах после обнаружения удовлетворяющего эталона алгоритм пытается сократить число ложных связей. Для снижения числа случаев ошибочного определения изоморфизма отдельные алгоритмы, например [6], реализуют вовлечение информации о соседних вершинах в функцию вычисления меток.

Итак, обобщенный алгоритм идентификации ГИ выглядит следующим образом.

- Шаг 1 (предобработка). Для входного графа и для всех графов в БД высчитываются степени у вершин и собирается другая информация, необходимая для последующих шагов, например, тип линии для примитивов инженерных чертежей. Далее полученные данные используются в качестве аргументов функции вычисления меток. На данном шаге также может собираться информация о соседних вершинах (число степеней и пр.) для включения в функцию вычисления меток.
- Шаг 2 (построение связей). Связи строятся между двумя сравниваемыми графами для всех пар вершин, которые имеют одинаковые метки. Пример таких связей приведен на рис. 1, где различные метки показаны черным, белым или серым цветом, вершина u графа G связана с вершиной v графа H . Две вершины соединены потому, что метка от u совпадает с меткой от v , т.е. метка и степень вершины u равны соот-

ветствующим значениям вершины v , а каждый потомок (смежная вершина) от узла u , а именно i, j, k , имеет соответствующего потомка от узла v с такой же меткой и степенью. На рис. 1 потомки i, j, k графа G совпадают с x, y, z на графе H . Однако ясно видно, что графы G и H не изоморфны.

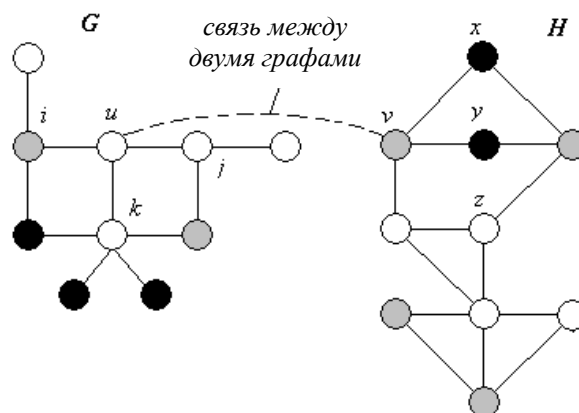


Рис. 1. Построение связи между двумя графами

- Шаг 3 (редукция энтропии). Данный шаг является самой затратной частью алгоритма по количеству вычислений. Происходит попытка разрешить неопределенные ситуации, если таковые возникли на предыдущем шаге, и сделать вывод о наличии изоморфизма между входным графом и эталонным графом.

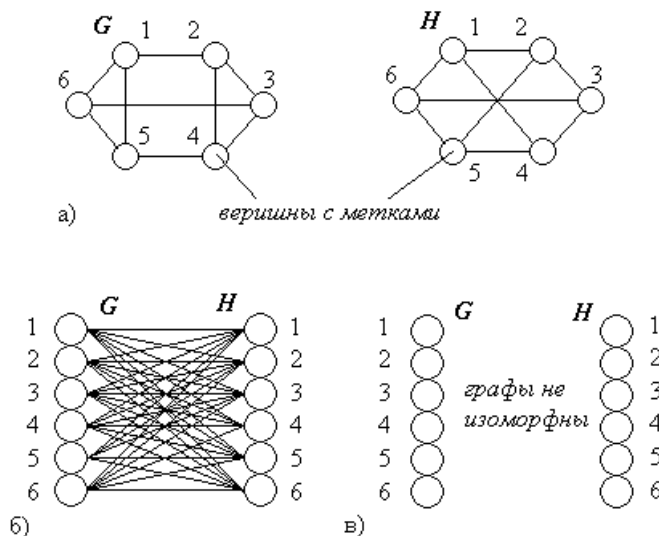


Рис. 2. Иллюстрация работы алгоритма поиска ГИ: а – шаг предобработки, даны два графа G и H ; б – шаг построения связей; в – шаг редукции энтропии

На рис. 2 изображены все шаги представленного алгоритма. На этапе предобработки (рис. 2, а) для обоих графов вычисляются метки для каждой вершины. Здесь каждая вершина содержит метку с информацией о текущем узле, а также о его соседях. Вдобавок следует отметить, что все метки графов G и H одинаковы как внутри графов, так и между G и H . Это равенство обусловлено тем, что метки самих вершин изначально равны (все кружочки отмечены белым цветом), оба графа являются однородными (одинаковое число ребер, смежных с каждым узлом). Данное равенство дает высокую степень неопределенности, когда каждая вершина графа G может быть соединена с каждой вершиной H (рис. 2, б). Неопределенность разрешается на шаге редукции энтропии (рис. 2, в), когда стартует детальный анализ графов.

Пути оптимизации алгоритма

В предыдущем разделе был показан пример (рис. 2), когда при сопоставлении двух графов возникали избыточные неопределенные ситуации. Разрешение таких ситуаций – весьма трудоемкий процесс, требующий привлечения процедуры детального анализа графов, во время работы которой происходит последовательное продвижение по уровням графа (каждый уровень характеризуется набором всех вершин, смежных с любой текущей вершиной). Существует еще ряд проблем, связанных с трудозатратами, которые линейно возрастают по мере добавления в БД эталонных графов. В данном разделе приведены два критерия и один момент, которые могут быть вовлечены в любой алгоритм поиска ГИ в качестве оптимизирующих мер.

Ниже даны несколько ключевых определений.

Определение 1. Графовый изоморфизм. Даны 2 графа $G=(V_1, E_1)$ и $H=(V_2, E_2)$, где $|V_1| = |V_2|$.

Граф G изоморфен H ($G=H$), при $V_1 \rightarrow V_2 \mid \exists (u, v) \in E_1 \ \&\& \ \exists (f(u), f(v)) \in E_2 \ \&\& \ l(u) \in l(f(u)) \mid \forall u \in V_1$, где l – функция вычисления метки для вершины.

Определение 2. Соответствие (идентичность) вершин. Пусть $u \in V_1$ – вершина графа G , а $v \in V_2$ – вершина графа H . Считается, что узел u идентичен узлу v , если метка от u равна метке от v .

Определение 3. Степенью неопределенности узла u , принадлежащего графу G , называется число узлов графа H , с которыми u идентичен. Обозначение: $\deg(u)=m$, где m – число вершин $\in H$, которым соответствует узел u .

Критерий 1. При поиске в БД эталонного графа, изоморфного входному, для каждого графа в БД запускается алгоритм идентификации ГИ, описанный выше. Ясно, что это весьма трудоемкий подход, требующий большого количества вычислений. Для оптимизации поиска ГИ можно ввести условие: во время первого пробега (последовательного сопоставления входного графа с каждым эталонным) алгоритм идентификации ГИ должен запускаться лишь для эталонных графов с числом вершин, равным или меньшим числа вершин входного графа. Данное условие применимо лишь для графов, которые отражают графическую информацию, заложенную в изображениях (инженерная документация, чертежи). Для таких изображений характерно наличие шумов, сказывающихся появлением на графах дополнительных вершин и ребер. Действительно, при разрыве линии на чертеже или при появлении случайного штриха, вызванного недостатками процесса сканирования, появляются шумы, которые, однако, не повлияют на вывод о наличии ГИ.

На вход процедуры сравнения графов подаются G – входной граф, $\{H\}$ – множество эталонных графов в БД, $G=(V_1, E_1)$ и $H=(V_2, E_2)$.

1. Для каждого H
2. если ($|V_1| \geq |V_2|$)
3. выполняется алгоритм поиска ГИ
4. если (ГИ обнаружен)
5. тогда выход из программы с результатом «ГИ найден»
6. Для каждого H
7. если ($|V_1| < |V_2|$)
8. выполняется алгоритм поиска ГИ
9. если (ГИ обнаружен)
10. тогда выход из программы с результатом «ГИ найден»
11. Выход из программы с результатом «ГИ не найден»

Критерий 2. При поиске в БД эталонного графа, изоморфного входному, можно добавить еще одно условие запуска алгоритма идентификации ГИ: запуск должен про-

исходить лишь тогда, когда набор меток входного графа равен набору меток эталонного. Здесь под равенством понимается, что для каждого элемента набора эталонного графа существует идентичный элемент набора входного графа (набор – совокупность элементов, уникальных внутри данной совокупности).

$G=(V_1, E_1)$ – входной граф, $H=(V_2, E_2)$ – эталонный граф, $\{H\} \in \text{БД}$.

$\exists$ набор меток F_1 , заданных функцией $l(u) \rightarrow F_1, \forall u \in V_1$

$\& \exists$ набор меток $F_2, l(v) \rightarrow F_2, \forall v \in V_2$

$\rightarrow \{l(u)\} \geq \{l(v)\}$

Количество элементов набора входного графа может быть больше или равно количеству элементов эталонного. Это допущение основано на природе графических изображений, получение которых может сопровождаться шумами (см. также критерий 1), обуславливающими увеличение числа вершин и ребер соответствующего графа.

1. Для каждого H
2. если $(\{l(u)\} \geq \{l(v)\})$
3. выполняется алгоритм поиска ГИ
4. если (ГИ обнаружен)
5. тогда выход их программы с результатом «ГИ найден»
6. Для каждого H
7. если $(\{l(u)\} < \{l(v)\})$
8. выполняется алгоритм поиска ГИ
9. если (ГИ обнаружен)
10. тогда выход их программы с результатом «ГИ найден»
11. Выход из программы с результатом «ГИ не найден»

Оптимизационный момент. Еще один момент может быть применен непосредственно в самом алгоритме идентификации ГИ. При этом на шаге предобработки графов, когда вычисляются степени для вершин, все узлы со степенью 2 должны быть удалены, а соответствующие две вершины, смежные с только что удаленным узлом, объединены в одну. Во время объединения двух вершин происходит формирование новой метки, связанной с новым узлом, которая хранит атрибуты удаленной вершины.

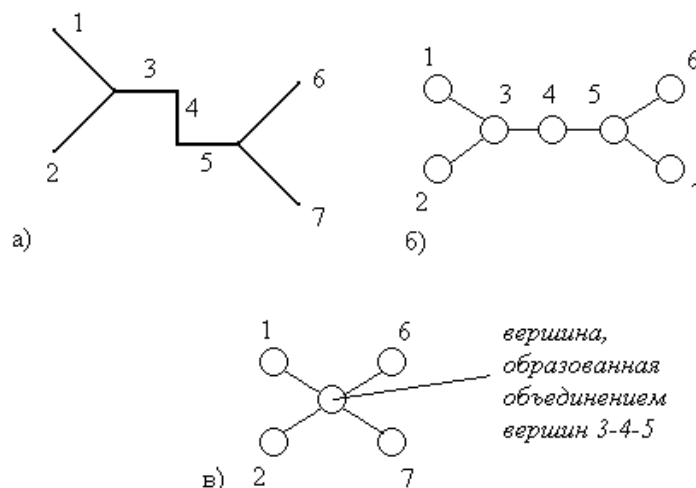


Рис. 3. Пример работы оптимизационного момента: а – входной чертеж с пронумерованными примитивами; б – графовое отображение чертежа; в – объединение вершин графа

На рис. 3 представлен пример работы данного момента. Дан фрагмент инженерного чертежа (рис. 3, а), разбитого на примитивы, каждому из которых приписан свой порядковый номер. Далее происходит переход к графовому представлению чертежа (рис.

3, б). Проанализировав последнее представление, можно сделать вывод о наличии избыточных узлов (узел 4 со степенью, равной 2), которые можно удалить, объединив две смежные вершины 3 и 5 в одну (рис. 3, в).

Использование этой оптимизационной меры сокращает количество узлов в графах, что повышает скорость работы алгоритма, при этом степень информативности графов не снижается. Вдобавок метки приобретают более содержательный характер, что снижает степень неопределенности (см. определение 3).

Заключение

В статье представлены оптимизационные меры (два критерия и один момент), которые могут быть включены в любой современный алгоритм поиска изоморфизма среди графов. Данные меры не снижают точность работы данных алгоритмов и в то же время являются весьма надежными.

Первые два критерия вовлекаются в процесс сопоставления входного графа с эталонными графами в БД. Здесь на ранних стадиях сопоставления происходит сравнение вершин и меток, что обеспечивает пропуск не удовлетворяющих критериям графов без их последующего детального анализа и, тем самым, рост производительности системы. Оптимизационный момент должен быть применен непосредственно в самом алгоритме идентификации ГИ, снижая степень неопределенности графов и повышая информативность меток, что дает более быструю и точную работу процедуры сопоставления графов. В дальнейших работах планируется реализация приведенных оптимизационных мер на языке программирования и проведение нагрузочного тестирования для исследования изменения производительности алгоритмов поиска ГИ.

Литература

1. Tombre K. Analysis of engineering drawings: state of the art and challenges // Graphics recognition. Algorithms and systems. V. 1389. – Lecture Notes. – Computer Science, Springer-Verlag, 1998. – P. 257–264.
2. Москаленко С.В., Гатчин Ю.А. Волновой алгоритм векторизации линейных растровых изображений // Научно-технический вестник СПбГУ ИТМО. – 2008. – № 51.
3. Xu J. A Generic Match Algorithm for structural homomorphism, isomorphism, and maximal common substructure match and its applications // J Chem Infor Comput Sci. – 1996. – № 36(1). – P. 25–34.
4. Rao A.C., Varada Raju D. Application of the Hamming number technique to detect isomorphism among kinematic chains and inversions // Mech. Mach. Theory. – 1991. – № 26 (1). – P. 55–75.
5. Spielman D.A. Faster isomorphism testing of strongly regular graphs: Technical report. – University of California Berkeley, Computer Science Division, Berkeley, California, 1996.
6. Abdulrahim M., Misra M. A Graph Isomorphism Algorithm for Object // Recognition, Pattern Analysis and Applic. – 1998. – V. 1. – № 3. – P. 189–201.
7. von der Malsburg C, Bienenstock E.A. Neural network for retrieval of superimposed connection patterns // Europhysics Let. – 1987. – № 3 (11). – P. 1243–1249.

Москаленко Станислав Владимирович	– Санкт-Петербургский государственный университет информационных технологий, механики и оптики, аспирант, stan.moskalenko@gmail.com
Гатчин Юрий Арменакович	– Санкт-Петербургский государственный университет информационных технологий, механики и оптики, доктор технических наук, профессор, зав. кафедрой, gatchin@mail.ifmo.ru

8

МЕТОДЫ И СИСТЕМЫ ЗАЩИТЫ ИНФОРМАЦИИ

УДК 004.056

**МОДЕЛЬ ОЦЕНКИ УГРОЗ ИНФОРМАЦИОННОЙ
БЕЗОПАСНОСТИ НА ОСНОВЕ МАТЕРИАЛОВ ХАКЕРСКИХ
КОНФЕРЕНЦИЙ****Э.Р. Хусайнова**

Проблема создания модели угроз является мало разработанной и актуальной. Модели угроз должны стать отправными точками для проектирования будущих систем защиты компьютерных и информационных систем. Построение модели угроз на основе материалов хакерских конференций позволит повысить надежность и защищенность информационных систем и программных продуктов.

Ключевые слова: хакерские конференции, модель угроз, защита информации, прогнозирование атак.

Введение

Предотвращение компьютерных атак со стороны злоумышленников – очень актуальная задача для специалистов, работающих в сфере информационных технологий. Согласно данным исследования специалистов из университета Мериленда (США), в Сети каждые 39 секунд происходит новая компьютерная атака [1]. Поэтому очень важно быть в курсе самых последних новостей, касающихся взлома и различного рода атак со стороны хакеров. Получить такую информацию можно из конференций и съездов, организуемых бывшими хакерами, ныне экспертами в области безопасности компьютерных систем и главами соответствующих ИТ-компаний. Попытаемся построить структурную модель оценки угроз информационной безопасности (ИБ) на основе материалов хакерских форумов и съездов.

Цели и задачи конференций хакеров

В литературе моделирование процессов нарушения информационной безопасности предлагается осуществлять на основе рассмотрения логической цепочки: «угроза – источник угрозы – метод реализации – уязвимость – последствия». Требуется, чтобы в ходе анализа все возможные источники угроз были идентифицированы, все возможные уязвимости сопоставлены с идентифицированными источниками угроз, всем источникам угроз и уязвимостям (факторам) сопоставлены методы реализации [2]. На наш взгляд, такая модель в рамках хакерских конференций будет неполной, поскольку не учитываются мотивы и цели субъектов угроз, а также их количественные и качественные характеристики, потоки информации, циркулирующие между источниками угроз и объектами исследований (т.е. уязвимостями).

Рассмотрим, что представляют собой хакерские конференции, каковы цели, задачи, масштабы подобных съездов, набирающих все большую популярность и значимость в области защиты ИБ.

Хакерские конференции – одни из важнейших мероприятий в области ИБ, значительно способствующие повышению уровня защиты информации и квалификации специалистов, а также демонстрирующие новые угрозы и атаки, методы реализации взлома и нападения на компьютерные и информационные системы. География проведения подобных мероприятий достаточно обширна. Функционирует ежегодная конференция Defcon в Лас-Вегасе, где собираются несколько тысяч участников из многих стран мира

– от США до Австралии [3]. С 1989 г. раз в четыре года проходит представительный хакерский форум в Голландии «Hackers At Large». Ежегодно в Германии проходит Всемирный конгресс хакеров под эгидой «Chaos Computer Club». Аналогичные конференции проходят также в таких странах, как Канада, Израиль, Южная Корея, Япония, Малайзия, Саудовская Аравия, Англия и т.д.

На международных съездах хакеров отчетливо прослеживается тенденция взаимодействия хакерского движения с государственными и коммерческими структурами. В них принимают участие представители государственных органов безопасности, администраторы крупнейших фирм. Более того, некоторые из известных хакеров активно участвуют в государственных и международных организациях по информационной безопасности. Так, например, президент и основатель «Chaos Computer Club» (Клуб компьютерного хаоса) Энди Мюллер-Мэган входит в состав всемирной организации «ICANN» (Internet Corporation for Assigned Names and Numbers) [4].

С каждым годом количество и состав участников растет и меняется: от хакеров до профессионалов в компьютерной безопасности, представителей крупнейших корпораций, правительственных организаций и федеральных спецслужб.

Цели, задачи и программа конференций становятся все более серьезными и глобальными, от развлечений до совместной борьбы с кибертерроризмом:

- привлечение внимания к информационной безопасности;
- сбор информации, обмен опытом борьбы с «кибертерроризмом»;
- реализация предупреждающих стратегий;
- демонстрация новейших решений в области управления информационной безопасностью;
- разработка инструментов для защиты системы от возможных угроз;
- обсуждение проблем социальных аспектов хакерства и др.

В программе хакерских конференций обсуждаются не только вопросы защиты ИБ, но и нападения со стороны нарушителей. В частности, в программу прошедших конференций были включены такие темы, как электронное голосование (поиск недостатков и выявление возможных угроз), беспроводные сети (Bluetooth – возможная утечка информации, атака с расстояния, уязвимости), DNS (сетевые атаки через DNS), социальные и технические аспекты (совесть хакера, злоупотребление властью), хакинг (методы против защиты в нецифровом мире), сетевая безопасность (IIS защита баз данных и трояны) и т.д.

Роль конференций в прогнозировании перспективных угроз

Еще одним важным достижением участников конференций является новое программное обеспечение, которое вполне можно использовать в повседневной жизни. Например, хакерская конференция ShmooCon 2008 порадовала целым арсеналом необычных устройств, которые можно было проверить в действии прямо на посетителях. Особую отметку получила оболочка Silica на базе Nokia N800, которая автоматически сканирует все окружающие Windows-компьютеры через WiFi, выявляет уязвимые ПК и позволяет сделать скриншоты с любого из них за пару щелчков мыши [5].

Результаты и подведение итогов конференций носят ценный характер с точки зрения привлечения всеобщего внимания к проблемам компьютерной безопасности и укрепления защиты важных инфраструктур от кибератак. Так, на конференции Defcon 17, прошедшей в июле 2009 года, правительство США открыто приглашало на работу «этичных хакеров». Это связано с тем, что в последние годы атакам подвергаются ключевые компьютерные сети, включая электроэнергетические системы. Специалист по безопасности и пилот Райтер Кункель (Righter Kunkel) продемонстрировал собравшим-

ся, как легко можно проникнуть в управляющую систему движения воздушного транспорта Федерального управления гражданской авиации США. При желании атакующий может, среди прочего, мешать обмену данными с диспетчерской вышкой, блокировать радар, ввести в систему поддельный план полета самолета, выдать себя за пилота, таковым не являясь, или же запретить самолетам вылет из аэропорта. Возможно, именно с этим связано множество авиационных крушений, произошедших в последнее время по всему миру [6].

Как видно, анализ негативных последствий реализации угроз на конференциях предполагает обязательную идентификацию возможных источников угроз, уязвимостей, способствующих их проявлению и методов реализации, т.е. классификацию (рис.)

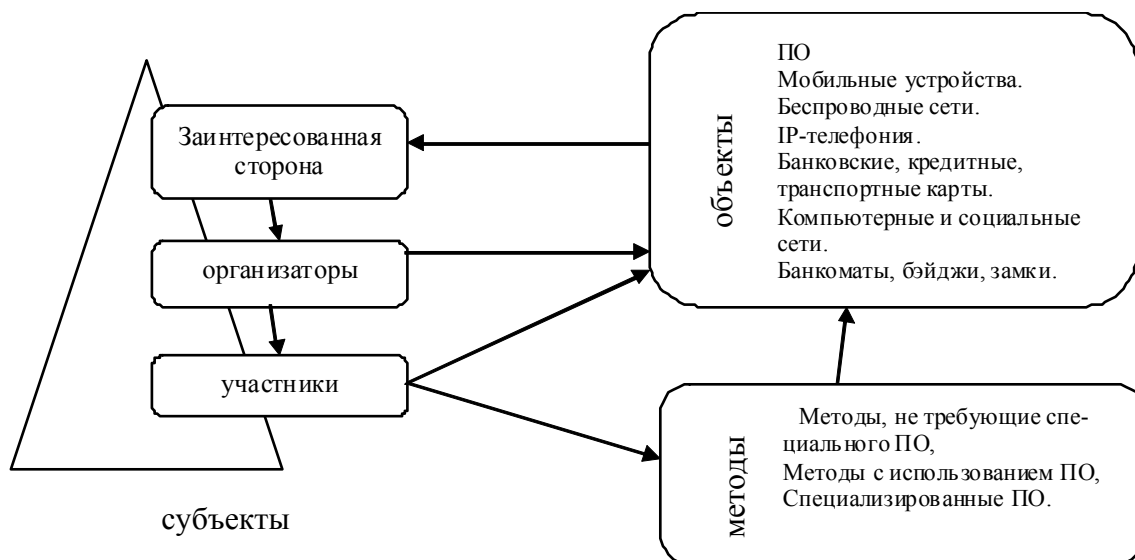


Рис. Структурная модель взаимодействия в среде хакерских конференций

К группе заинтересованной стороны относят представителей государственных структур, спецслужб, корпораций, компаний, банков, ИТ-специалистов, хакеров и агентов, для каждого из которых интерес представляет свой объект исследования. Например, для госструктур такими объектами являются программные устройства, системы контроля доступа, протоколы и т.д.

Заинтересованная сторона выбирает объект исследования, предоставляет полный или ограниченный доступ к ресурсу, ставит цели и задачи перед организаторами и сроки выполнения поставленных задач. Так, например, компания Microsoft предлагает исследовать на предмет безопасности, простоты использования, визуализации новую операционную систему (ОС), допустим, Windows Vista (объект исследования), и ставит перед организаторами следующие задачи: выявить основные области ОС, которые требуют дальнейшей доработки и модернизации, предложить возможные решения выявленных проблем. Организаторам было выделено 100 бета-версий ОС. Организаторы же, в свою очередь, выбирают группу хакеров, определяют информационные и технические ресурсы, подлежащие защите либо атаке. Участники исследуют объект, проводят анализ характеристик угроз и уязвимых мест для информации, оценивают и измеряют риски, подбирают и используют методы и способы защиты или взлома.

Вернемся к нашему примеру. Группа американских исследователей выделила такие объекты: безопасность доступа к ОС, уязвимости с точки зрения Интернет-угроз, защищенность от вирусных атак и т.п. Также были предложены соответствующие методы: интеграция в систему нескольких вирусов и исследование поведения ОС, использование программы генерации пароля для аутентификации. Подобные действия были проведены и другими группами.

Далее участники приступают к практической части конференции. На этом этапе производится весь анализ и составление будущего выступления, презентации своих достижений на конференции:

- проверка методов и способов защиты или взлома на объектах;
- анализ и подведение итогов исследования;
- подготовка к презентации полученных результатов;
- демонстрация и апробация новых утилит и устройств.

Таким образом, модель оценки угроз информационной безопасности – непрерывный цикл информационных потоков между субъектами и объектами съездов. Это совокупность внешних и внутренних факторов и их влияния на состояние информационной безопасности и на сохранение материальных и информационных ресурсов. Модель угроз должна впоследствии использоваться в проектировании будущих систем защиты информационной безопасности от возможных уязвимостей.

Мотивация субъектов конференций

Определив источники угроз и уязвимости для анализа причин и последствий реализации угроз, необходимо выяснить, какими же мотивами и целями руководствовались нарушители. Представители компаний и банков действуют в целях получения личной выгоды, новой информации об уязвимостях, протоколах; государственные и спецслужбы – в целях национальной безопасности, раскрытия конфиденциальной информации, предотвращения утечек баз данных и другой важной информации; хакеры – ради получения порции адреналина от осуществления незаконных действий, а также, чтобы проверить свои возможности и способности в области хакинга и защиты.

В продолжение примера отметим следующее. Американские, русские и французские ученые достигли конкретных результатов и выступили с докладами на конференции. Компания Microsoft, таким образом, получила интересующие ее сведения об ОС Windows Vista и определила для себя дальнейшие пути исследований и модернизации своего продукта. Все субъекты получили свою выгоду: организаторы – денежный чек, участники – премии, вакансии в компании заказчика, известность и способность доказать значимость своих исследований, а Microsoft – новые цели, ценные кадры и рекламу ОС.

В целом такой вид получения информации касательно интересующего заказчика объекта имеет как свои плюсы, так и минусы. Непосредственно сами результаты конференции и конкретных исследований, довольно быстрое достижение результата можно отнести к положительным аспектам, а риск, на который идет заказчик, предлагая участникам изучить сам объект, в силу своей непредсказуемости, безусловно, является отрицательной стороной. Ведь публичное выступление на хакерской конференции коренным образом влияет на общественное мнение, что может как прибавить, так и уменьшить спрос на объект. Несмотря на это, подобные съезды приобретают все большую популярность и охватывают все новые и новые горизонты информационных технологий.

Заключение

Таким образом, исследование материалов хакерских конференций напрямую способствует выявлению уязвимостей программного обеспечения и оборудования, а также предотвращению Интернет-атак. Вышеперечисленные примеры доказывают эффективность подобных мероприятий и вполне адекватную реакцию на это со стороны представителей государственных и коммерческих предприятий. К наиболее важным результатам конференций следует относить последующие изменения в программном и техническом оборудовании, а также глобальное информирование пользователей ПК о возможных путях преодоления защиты и доступа к их персональным данным.

Необходимо проводить хакерские конференции в России для защиты национальных интересов страны, уделять должное внимание вопросам информационной безопасности, каковые и обсуждаются на подобных конференциях.

Результатами проведения подобной конференции в России могут стать:

- привлечение внимания общественности к проблемам защиты информации;
- повышение мирового статуса России в области информационной безопасности;
- привлечение дополнительных государственных и зарубежных инвестиций в область защиты информации.

Это качественно повысит уровень информационной безопасности и даст возможность более интенсивно развивать данную область.

Литература

1. Сайт Интернет-журнала Hackzone [Электронный ресурс] / 2009. – Режим доступа: <http://www.hackzone.ru/news/view/id/5204/>, свободный. – Яз. рус.
2. Вихорев С., Кобцев Р. Как определить источники угроз? [Электронный ресурс] / 2003. – Режим доступа: <http://www.citforum.ru/security/articles/threats/>, свободный. – Яз. рус.
3. Сайт конференции DefCon [Электронный ресурс] / 2009. – Режим доступа: www.defcon.org, свободный. – Яз. англ.
4. Сайт конференции Chaos Communication Congress [Электронный ресурс] / 2009. – Режим доступа: www.ccc.de, свободный. – Яз. англ., нем.
5. Хакерские гаджеты на конференции ShmooCon [Электронный ресурс] / 2008. – Режим доступа: <http://habrahabr.ru/blogs/infosecurity/5893/>, свободный. – Яз. рус.
6. Сайт Интернет-журнала securitylab [Электронный ресурс] / 2009. – Режим доступа: <http://www.securitylab.ru/>, свободный. – Яз. рус.

Хусаинова Эльвира Робертовна

– Санкт-Петербургский государственный университет информационных технологий, механики и оптики, аспирант, elechka.de@gmail.com

УДК 004.056.4

ОЦЕНКА ЭФФЕКТИВНОСТИ АТАК ЗЛОУМЫШЛЕННИКА В ПРОЦЕССЕ ПОСТРОЕНИЯ ЕГО МОДЕЛИ

А.И. Спивак

Проектирование системы защиты информации всегда сопровождается созданием модели нарушителя безопасности. Общеизвестным подходом является использование характеристик злоумышленника в качестве основных параметров для оценки атак на систему защиты. Изменение свойств объекта защиты может быть использовано для выполнения оценки действий нарушителя безопасности.

Ключевые слова: модель, безопасность, злоумышленник, оценка, информация.

Введение

В процессе построения системы информационной безопасности на одном из этапов производится оценка эффективности атак злоумышленника. На основе данных об объекте защиты и возможных атаках строится модель защиты. Полнота описания угроз безопасности, а также правильность модели нарушителя позволяют системе защиты информации учитывать большую часть атак и иметь адекватные меры по препятствованию их реализации.

Сложность оценки обуславливается широким спектром характеристик нарушителя. Наряду с очевидными, такими как техническая оснащенность злоумышленника, его профессиональная подготовка, обеспеченность вычислительными мощностями, методы воздействий, мотивационные факторы, существуют и другие переменные характеристики. Их количество, характер, взаимосвязи, а также количественные показатели являются сложно прогнозируемыми величинами. В связи с этим возникает необходимость в оценке воздействия злоумышленников на систему защиты информации, не учитывающей характеристики нарушителей безопасности.

Существующие подходы к построению модели нарушителя

Основные подходы к решению задачи оценки эффективности атак злоумышленника реализуются в рамках построения модели нарушителя. Остановим свое внимание на описательной части построения модели нарушителя, включающей в себя этап оценки эффективности действий злоумышленника по преодолению системы защиты объекта. Одной из используемых методик построения описательной части такой модели является использование в качестве основополагающего фактора осведомленности [1]. Знания злоумышленника о системе защиты информации существенно повышают вероятность реализации угрозы информационной безопасности. Выделяют несколько уровней осведомленности – от самого высокого до самого низкого. К высокой степени осведомленности относится злоумышленник, которому известна вся проектная документация, и на основе этой информации он обладает данными об уязвимых точках системы. Нарушитель с низким уровнем знаний об объекте имеет информацию только о внешних, общедоступных составляющих системы, и предполагается, что вероятность реализации атаки таким нарушителем ниже, чем злоумышленником с высоким уровнем знаний о системе защиты.

Другим подходом является ранжирование по виду доступа, которым обладает потенциальный нарушитель безопасности [2]. В зависимости от особенностей работы различные категории пользователей системы имеют разные права доступа. Обычный пользователь системы может пользоваться только строго определенным набором средств, список которых и выполняемые действия регламентированы. При этом администратор безопасности обладает доступом ко всем системам обеспечения безопасности и имеет гораздо больше полномочий и, как следствие, возможностей для атаки на систему защиты. Разграничение доступа к системе определяет различные способы для реализации угроз безопасности. Вероятность успешности атаки лежит в прямой зависимости от типа доступа злоумышленника к системе.

Эффективность действий злоумышленника при использовании описанных подходов является производной величиной от характеристик, связанных с нарушителем. В случае с осведомленностью полнота знаний о системе защиты позволяет говорить о большой эффективности действий злоумышленника по реализации угрозы безопасности. Аналогичное значение имеет тип доступа к системе во втором подходе. Данные виды оценки базируются только на одной из характеристик злоумышленника. При учете одновременно двух подходов образуется сложность в выявлении взаимных влияний со стороны различных характеристик. Дальнейшее расширение списка параметров субъекта воздействия вызовет еще большее усложнение процесса построения модели [3].

Оценка эффективности атак на основе изменения вероятности реализации угрозы безопасности

Одним из возможных вариантов решения может служить подход к оценке эффективности атак злоумышленника на основе изменения вероятности достижения цели нападения. Если принять во внимание, что конечной целью является оценка эффектив-

ности действий злоумышленника по реализации угрозы безопасности, то данный подход представляется более универсальным, чем другие. Оценку можно проводить по показателям, определяющим состояние защищаемого объекта, а именно изменение вероятности быть взломанным под воздействием нарушителей. Отличие от существующих подходов заключается в том, что характеристики объекта защиты являются основными при измерении эффективности действий злоумышленника. В таком случае оценка эффективности осуществляется с точки зрения приближения злоумышленника к цели в результате реализации угрозы безопасности. Переориентирование акцента на оценку изменений в свойствах защищаемого объекта позволяет упростить описательную часть построения модели нарушителя. Вместо расчета разнообразных характеристик злоумышленника, учета их взаимосвязей и взаимовлияний достаточно измерить только один параметр объекта воздействия, на который направлены действия нарушителя безопасности.

Одновременно интересен вопрос возможного способа измерения вероятности реализации угрозы безопасности при использовании ориентированного на объект защиты подхода. Характеристики объекта защиты служат тем фактором, изменение которого позволяет связать эффективность действий злоумышленника и приближение его к цели этого воздействия – взлому. Система защиты информации включает в себя некоторое количество уровней безопасности, и злоумышленнику для реализации угрозы нужно их преодолеть. Условимся считать, что рассматриваемые уровни защиты одинаковы по своим характеристикам, а сложность их преодоления злоумышленником не меняется в зависимости от свойств конкретного уровня. Знания о количестве уровней превалируют над информацией об их параметрах, влияющих на их прохождение злоумышленником в процессе взлома. Одним из возможных способов измерения вероятности взлома можно назвать количество пройденных нарушителем за время атаки уровней. Сопоставление данного числа с общим количеством уровней защиты приведет к получению изменения вероятности взлома системы защиты.

Деструктивное воздействие

Введем понятие деструктивного воздействия, которому поставим в соответствие действия злоумышленника, приводящие к снижению общего уровня безопасности объекта. В действительности такого рода воздействия могут представлять собой широкий круг угроз безопасности, таких как нарушение целостности объекта защиты, его конфиденциальности, злонамеренное блокирование доступа к защищаемому объекту.

Формулировка понятия указывает на то, что действия строго направлены на преодоление уровней безопасности объекта защиты и приводят, в конечном счете, к увеличению вероятности реализации угрозы.

Определим способ, который используется для оценки деструктивного воздействия. Очевидно, что деструктивное воздействие направлено на достижение конкретной цели. Деструктивное воздействие подобно понятию ценности информации. Воздействие, также как и получение информации, направлено на достижение цели, т.е. в результате этого воздействия вероятность достижения цели либо не меняется (воздействие ни к чему не привело), либо происходит приближение к цели (вероятность достижения цели возрастает), возможно также уменьшение вероятности вследствие получения неверной информации (дезинформации). Использование соотношения Корогодина [4] для оценки ценности является наиболее удобным способом представления деструктивного воздействия злоумышленника. Такое представление позволяет связать вероятности реализации угроз безопасности до и после момента измерения с действиями злоумышленников между выполняемыми измерениями. Соотношение имеет вид

$$C_i = \frac{P_i - p_i}{1 - p_i},$$

где p_i – вероятность достижения цели до начала деструктивного воздействия для i -го объекта, P_i – вероятность достижения цели после окончания деструктивного воздействия для i -го объекта, C_i – величина деструктивного воздействия для i -го объекта. Значение деструктивного воздействия лежит в пределах от 0 до +1, что позволяет говорить о деструктивном воздействии как о приращении вероятности реализации угрозы безопасности после деструктивного воздействия со стороны злоумышленника.

Достаточно важным представляется момент оценки собственно вероятности реализации угрозы безопасности. Возможным вариантом такой оценки является предположение о том, что реализация угрозы безопасности зависит от количества уровней защиты на конкретном защищаемом объекте. При увеличении количества таких уровней вероятность реализации угрозы безопасности объекта защиты будет стремиться к обратно пропорциональной зависимости от числа уровней защиты. Чем больше уровней защиты фигурирует в соотношении, тем более корректно оно отображает связь между вероятностью реализации угрозы и количеством уровней защиты. Это утверждение очевидно: чем больше уровней защиты нужно преодолеть злоумышленнику для взлома системы, тем меньше вероятность реализации угрозы.

Вероятность реализации угрозы (p_i) для i -го объекта будет равна

$$p_i \sim \frac{1}{n_i},$$

где n_i – количество уровней защиты i -го объекта. Использование данного выражения позволяет соотношению деструктивного воздействия обрабатывать информацию о количестве уровней защиты, которое имеет объект защиты, и на основе этого формировать значение вероятности реализации угрозы. Затем информация об уровнях защиты, которые были преодолены злоумышленником в процессе воздействия, используется для вычисления деструктивного воздействия. Если известно количество уровней защиты, которые злоумышленник в результате деструктивного воздействия смог преодолеть, то можно выполнить оценку величины этого деструктивного воздействия, используя соотношение

$$p_i \sim \frac{1}{n_i - s_i} - \frac{1}{n_i},$$

где n_i – количество уровней защиты i -го объекта, s_i – количество взломанных уровней защиты i -го объекта, p_i – вероятность реализации угрозы безопасности i -го объекта. При $n_i = s_i$ $p_i = 0$, так как деструктивное воздействие ни к чему не привело, приращение вероятности реализации угрозы в этом случае равно нулю.

Использование величины деструктивного воздействия можно связать с эффективностью действий злоумышленника. Деструктивное воздействие, как было показано, характеризуется вероятностью реализации угрозы безопасности, а именно ее изменением. Чем выше вероятность реализации угрозы, тем больше воздействие на систему защиты объекта и тем более эффективными являются действия злоумышленника. Таким образом, величина деструктивного воздействия соответствует эффективности действий злоумышленника, так как отражает успешность действий нарушителя безопасности.

Использование величины деструктивного воздействия

Рассмотренная величина может использоваться для оценки эффективности действий злоумышленника при условии ограниченности или полного отсутствия знаний о

нарушителе безопасности. В таком случае можно опираться при оценке на показатели безопасности объекта защиты, а именно изменение вероятности реализации угрозы. Одной из предложенных для измерения характеристик является количество преодоленных нарушителем уровней защиты системы безопасности. Подход не исключает использование других характеристик, а также их сочетаний.

Заключение

Описанный способ оценки эффективности действий злоумышленника является дополнительным инструментом для построения описательной части модели нарушителя. Совокупность различных подходов к оценке воздействий позволяет получить наиболее полную модель угроз безопасности, а на ее основе спроектировать адекватный набор мер по эффективному противодействию угрозам безопасности.

Возможность базирования в оценке воздействий на систему защиты информации не на характеристики злоумышленника, а на результаты его действий является важным инструментом при построении модели системы защиты информации.

Литература

1. Дровникова И.Г., Буцынская Т.А. Модель нарушителя в системе безопасности // Системы безопасности. – 2008. – №5.
2. Гостехкомиссия России. Руководящий документ. Концепция защиты средств вычислительной техники и автоматизированных систем от несанкционированного доступа к информации. – М.: Военное издательство, 1992. – 12 с.
3. Малюк А.А. Информационная безопасность: концептуальные и методологические основы защиты информации. – М.: Горячая линия–Телеком. 2004. – 280 с.
4. Корогодина В.И., Корогодина В.Л. Информация как основа жизни. – Дубна: Феникс+, 2000. – 208 с.

Спивак Антон Игоревич – Санкт-Петербургский государственный университет информационных технологий, механики и оптики, аспирант, anton.spivak@gmail.com

УДК 338.2

СТРАТЕГИЯ ФОРМИРОВАНИЯ И РАЗВИТИЯ ИННОВАЦИОННОГО ПОТЕНЦИАЛА ПРОМЫШЛЕННОГО ПРЕДПРИЯТИЯ

О.В. Васюхин, Е.А. Павлова

В работе представлена система стратегических решений инновационно-активного предприятия. Определены типы корпоративных стратегий, направленных на инновационное развитие предприятия. Выделены отличительные особенности стратегии создания конкурентных преимуществ, позволяющие определить подходы к обоснованию состава инновационного потенциала предприятия. Раскрыто содержание стратегий инновационной сферы, определяемых масштабами инновационной деятельности. Предложенный методологический подход позволит предприятию целенаправленно развивать свой потенциал в соответствии с индивидуальной спецификой деятельности.

Ключевые слова: стратегии развития предприятия, инновации, инновационный потенциал.

Введение

Анализ производительных сил в современном обществе все больше убеждает в том, что основой развития национальной экономики (особенно в России) является не техника, не технология, не уровень квалификации рабочей силы и даже не научно-технический потенциал, а эффективность используемых систем управления предприятием и, в частности, инновационными процессами. Зарубежный опыт (например, японский конца 70-х годов) свидетельствует о том, что отдельно взятое предприятие способно даже при не очень благоприятных общих внешних условиях добиться положительных хозяйственных результатов, совершенствуя внутрифирменную систему управления на основе активизации инновационных процессов.

Методологические подходы к формированию стратегий деятельности предприятия

В условиях действия достаточно развитых рыночных отношений любое предприятие является экономически мотивированным участником рынка и имеет определенную цель. При этом традиционным измерителем успешности предпринимательской деятельности является прибыль. Предприятие может достичь поставленных целей посредством извлечения прибыли, преобразуя свои ресурсы (материально-технические, трудовые, финансовые, инновационные) в товары (услуги, работы) с последующей их продажей покупателям. Поэтому справедливым становится утверждение о том, что в типовой структуре управления предприятием должны всегда присутствовать, по меньшей мере, такие типы задач, как стратегический анализ и управление внешним окружением предприятия (оценка текущих и перспективных потребностей рынка), а также оперативное управление ресурсами предприятия, в том числе инновационными.

Отсюда следует, что к *стратегическим проблемам* предприятия следует отнести:

- маркетинговые (выбор номенклатуры и ассортимента продукции, определение эффективного сегмента рынка, обоснование методов конкурентирования, оценка реакции рынка на инновации предприятия, формирование набора направлений инновационной деятельности);

- конкурентные (формулировка ориентиров, целей и задач развития предприятия; обоснование моделей расширения, изменения, разнообразия хозяйственной деятельности, выбор методов достижения конкурентного преимущества).

Оперативные решения в общем случае имеют своей целью увеличить отдачу от использования внутренних ресурсов предприятия. Основными сферами принятия таких решений, на наш взгляд, являются:

- распределение ресурсов (построение бюджетов) по функциональным областям и сферам предпринимательской деятельности;
- составление производственных планов, наблюдение и контроль за их исполнением, оперативное регулирование.

В качестве ключевых решений выступают проблемы распределения ресурсов между различными видами деятельности, формирования инновационного потенциала предприятия, финансирования исследований и разработок [1].

Очевидно, что стратегические и оперативные решения взаимосвязаны и дополняют друг друга. С одной стороны, даже если стратегические решения подтверждают, что продукция предприятия и его рынки сбыта выбраны правильно и существует достаточный уровень спроса, то всегда требуются соответствующие оперативные решения, например, в области цен, издержек, объема и времени выпуска продукции. С другой стороны, значительный рост числа и разнообразия оперативных проблем, нарастающих быстрыми темпами, зачастую свидетельствуют о том, что проблемы носят не внутрифирменный характер, а обусловлены динамикой внешней среды, что требует принятия стратегических решений.

Таким образом, в системе управления любого предприятия должен соблюдаться определенный баланс внимания со стороны управленцев как к стратегическим, так и к оперативным решениям. При этом в общем случае приоритет должен отдаваться именно стратегическим решениям, так как они определяются динамикой внешней среды, в которой данное предприятие и функционирует.

Главное отличие в методологии стратегического управления состоит в том, что:

- работа по определению целей и направлений деятельности предприятия ориентирована на будущее, а не описывает текущую ситуацию;
- стратегический план не детерминирован, т.е. в нем нет жесткой регламентации по ресурсам, исполнителям и срокам, он обосновывает общие перспективные направления развития хозяйственной деятельности предприятия;
- в системе стратегического управления отсутствует предположение о том, что будущее конкурентное и экономическое состояние предприятия определяется ретроспективными тенденциями.

Именно эти отличительные особенности стратегического подхода, как доказано в работе [2], наиболее адекватно согласовываются с методологией управления инновационными процессами на предприятии. Ведь сама цель проявления инновационной активности заключается в обеспечении коммерческих и (или) других успехов предприятия в будущем. Отсюда стратегия формирования и развития инновационного потенциала предприятия, обоснование целесообразных направлений инновационной активности должны быть составной частью общекорпоративного стратегического плана.

Система стратегий промышленного предприятия

Основой системы стратегического управления является выработка общей стратегии развития предприятия. На разных уровнях управления стратегия, не меняя своей методологической сущности, должна быть разной. Так, на крупных предприятиях, занимающихся многопрофильной деятельностью, стратегия развития определяет, как вы-

годно размещать общий капитал. В то же время на уровне отдельно взятого функционального подразделения этого предприятия стратегия должна определять, как лучше реализовать данную функцию системы управления [3].

В связи с этим, если предприятие осознает себя как достаточно крупную корпорацию с широко диверсифицированным хозяйственным портфелем, то оно должно выстраивать, поддерживать и реализовывать систему стратегий, имеющую форму иерархии, включающей в себя четыре управленческих уровня. В этой системе на первом уровне определяется наиболее общая, корпоративная стратегия, на втором – совокупность деловых стратегий, на третьем – функциональные стратегии и на четвертом – производственные (операционные) стратегии.

Сущность, цель и задачи каждого из перечисленных уровней стратегий состоят в следующем.

Корпоративная стратегия – это общие направления развития корпорации в целом. Эта стратегия определяет все виды деятельности, которыми корпорация занимается в настоящее время и намерена развивать в будущем. Основная ее цель состоит в том, что она включает действия корпорации, предпринимаемые ею для утверждения своих позиций в различных отраслях национальной (транснациональной) экономики и методов, используемых для управления экономическим и финансовым состоянием корпорации. Это предусматривает решение следующих управленческих задач.

1. Обоснование приоритетных сфер деятельности корпорации и уровня их разнообразия.
2. Определение желаемого уровня конкурентных стратегических и экономических показателей в тех отраслях, где корпорация уже действует.
3. Оценка (нахождение) синергических эффектов среди родственных (входящих в данную корпорацию) предприятий.
4. Обоснование инвестиционной политики и инвестиционных приоритетов в пределах стратегической перспективы.

Следующий уровень системы стратегий – *разработка деловой стратегии* (в мировой практике – «бизнес-стратегия»), суть которой состоит в выработке подходов и методов, достижения желаемых стратегических показателей в каждом виде бизнеса, включенного в общекорпоративный стратегический план. Для достижения поставленных целей в деловой стратегии необходимо отобразить ряд основных позиций.

1. Определение характера реакции корпорации в целом и каждого отдельного ее центра прибыли на динамику внешней макро- и микросреды с целью адаптации деловой стратегии к динамике внешней среды по каждому виду бизнеса.
2. Разработка мер, обеспечивающих конкурентные преимущества в каждом виде бизнеса, на основе объединения стратегических усилий всех функциональных подразделений корпорации (снабжение, новые разработки, текущее производство, продажи).
3. Выработка стратегии в области частных нововведений. Стратегия реализации каждого из таких нововведений может быть специфична. Однако в целом все они должны укладываться в рамки принятой деловой стратегии и, по меньшей мере, не разрушать ее.

В итоге разработка деловой стратегии сводится к выработке и согласованию методов конкурентной борьбы в данном виде бизнеса, рассчитанных на долгосрочную перспективу.

Функциональная стратегия – это план действий отдельного функционального подразделения (функциональной зоны) внутри определенной бизнес-единицы (предприятия, входящего в корпорацию). Например, это может быть стратегия инновационной деятельности, направленной на увеличение доли контролируемого рынка. Роль функциональной стратегии заключается в поддержке деловой стратегии и конкуренто-

способности корпорации в целом. Кроме того, функциональная стратегия создает управленческие ориентиры (вырабатывает цели и методы их достижения) в отдельных функциональных сферах деятельности предприятия (общая политика закупок, стратегия технического и/или технологического перевооружения и т.п.). Очевидно, что каждая функциональная стратегия должна быть согласована со всеми остальными для обеспечения внутреннего синергизма.

Производственная (операционная) стратегия определяет методы управления основными структурными подразделениями (сферой разработок, производственными подразделениями, отделом продаж, отделом закупок) при решении текущих оперативных задач, имеющих стратегическую важность.

Обоснование направлений развития инновационного потенциала предприятия

Приведенный краткий анализ особенностей стратегий разного уровня показывает, что с точки зрения экономического развития и обеспечения долговременной конкурентоспособности особую роль играет выбор того или иного типа общекорпоративной стратегии. Мировая теория и практика выработала несколько качественно различающихся между собой типов корпоративной стратегии.

Стратегия узкой специализации. Большинство крупных корпораций начинает свою деятельность как небольшие узкоспециализированные предприятия, обслуживающие местный или региональный рынок. Однако для обеспечения роста в пределах одной отрасли корпорация может использовать *стратегию вертикальной интеграции* по направлению к поставщикам исходных для производства материально-технических ресурсов и компонентов (*интеграция назад*) и (или) по направлению к конечному потребителю продукции предприятия (*интеграция вперед*). Стратегия вертикальной интеграции предопределяет инновационную активность предприятия на границах освоенного бизнеса, что требует соответствующей адаптации имеющегося инновационного потенциала. Подобное направление развития приводит к созданию «универсального инновационного потенциала».

Стратегия диверсификации в родственные отрасли может осуществляться путем передачи ноу-хау и или других уже освоенных инноваций из одного вида деятельности (отрасли) в другой. Подобный подход позволяет на базе одного и того же инновационного потенциала создавать нововведения по широкому спектру деятельности. Для достижения преимуществ от диверсификации схожие производственные и функциональные подразделения корпорации (особенно в сфере новых разработок) должны быть агрегированы в единый координируемый модуль, обеспечивающий управление развитием и использованием инновационного потенциала корпорации. Данная стратегия предопределяет расширение инновационной активности (распространение и адаптацию инноваций) предприятия на базе имеющегося инновационного потенциала. Подобное направление развития приводит к созданию «адаптивного инновационного потенциала».

Стратегия диверсификации в неродственные отрасли (непрофильная диверсификация). Основным принципом непрофильной диверсификации является то, что любой бизнес, который может быть освоен в короткие сроки, не отвлекает крупные капитальные вложения и имеет хорошие перспективы получения доходов и прибыли, представляет собой выгодное направление для диверсификации. Данная стратегия предопределяет необходимость либо наращивания и реструктуризации имеющегося на предприятии инновационного потенциала, либо кооперации с другими инновационно-активными организациями (отраслевыми НИИ, вузами и т.п.). Подобное направление развития приводит к созданию «пионерного инновационного потенциала».

Таким образом, анализ и обоснование того или иного типа общекорпоративной стратегии позволяет определить общие принципы и направления развития инновационного потенциала предприятия с выделением характерных его особенностей.

Обоснование состава инновационного потенциала предприятия

На основе принятой общекорпоративной стратегии руководство предприятия вырабатывает адекватные конкурентные стратегии для каждого, принятого в стратегическом плане, вида бизнеса. В научных публикациях по исследуемой теме [3, 4] выделяют следующие разновидности конкурентных стратегий.

Стратегия лидерства в издержках обычно применяется крупными корпорациями, занимающими доминирующие позиции и имеющими значительные финансовые, людские и научно-исследовательские ресурсы. Суть данной стратегии состоит в ориентации производства на массовый выпуск стандартной, но качественной продукции и реализации ее по достаточно низким ценам. Она предопределяет такой состав инновационного потенциала предприятия, который будет способен к инновациям в области организации производства, техники и технологии, бизнес-моделирования и т.п., т.е. «потенциал с ориентацией на внутрипроизводственные процессы».

Стратегия дифференциации заключается в том, что предприятие ориентируется на изготовление особой, уникальной продукции для определенного, чаще всего узкого, круга покупателей. Предприятия, реализующие стратегию дифференциации, должны иметь высокий потенциал для проведения научно-исследовательских и опытно-конструкторских работ, прекрасную систему обеспечения высокого качества продукции в процессе производства, а также совершенную систему маркетинга, т.е. «инновационный потенциал с ориентацией на внешнюю среду».

Стратегия фокусирования предполагает концентрацию усилий в ограниченной области, где у предприятия есть явное конкурентное преимущество для достижения главенствующей позиции на выбранном им узком сегменте рынка. Стратегия фокусирования предопределяет такой состав инновационного потенциала предприятия, который будет способен к восприятию инноваций, зарождающихся в недрах крупных корпораций, и реализацию их на базе собственного производства, т.е. «потенциал с ориентацией на лидера».

Таким образом, стратегии создания конкурентных преимуществ имеют существенные отличительные особенности, позволяющие определить общие подходы к обоснованию потребного состава инновационного потенциала предприятия.

Обоснование масштабов развития инновационного потенциала предприятия

Предприятия, занимающие активную позицию в области разработки, использования и распространения инноваций, помимо общекорпоративной стратегии и стратегии создания конкурентных преимуществ, вырабатывают специфическую инновационную стратегию. Анализ научной литературы по данной проблеме показывает, что обычно, говоря о стратегической позиции предприятия в инновационной сфере, выделяют такие альтернативные стратегии.

Стратегия технологического и рыночного лидерства. Основой данной стратегии является нацеленность (и способность) на создание принципиально нового продукта. Цель реализации данной стратегии состоит в том, чтобы добиться признания созданной новой технологии как эталона для рассматриваемой отрасли и тем самым обеспечить рыночное лидерство предприятия-разработчика на длительную перспективу. Стратегия технологического и рыночного лидерства предопределяет крупные масштабы иннова-

ционной деятельности предприятия в форме самостоятельной структурированной функциональной зоны – «инновационный департамент», возможно, в кооперации с другими, внешними научно-исследовательскими центрами.

Стратегия «следования за лидером». Смысл следования за лидером состоит в том, что, с одной стороны, предприятие сосредоточивает свое внимание на новых для рынка товарах, но уже получивших признание у потребителей, с другой стороны, минимизирует затраты на научные исследования и опытно-конструкторские разработки. Стратегия «следования за лидером» предопределяет такие сравнительно небольшие масштабы инновационной деятельности предприятия, которые будут осуществляться на базе собственного инновационного потенциала в форме «отдела по инновациям».

Стратегия имитации базируется на использовании уже известных технологий, а предприятие-производитель осуществляет только лишь их поднастройку в соответствии с требованиями конкретного рынка. Стратегия имитации предопределяет необходимость только лишь модификации известной инновации, что, в общем случае, можно выполнить силами собственного высококвалифицированного инженерно-технического и управленческого персонала в форме «целевой группы», т.е. временного творческого коллектива, состоящего из специалистов различных подразделений предприятия и создаваемого для планирования и реализации конкретного инновационного проекта.

Стратегия зависимости представляет собой частный случай стратегии имитации. Предприятие-производитель вообще не занимается совершенствованием технологии, а сосредоточивает все свои усилия только на отладке производственного процесса и на организации эффективного маркетинга. Стратегия зависимости предопределяет масштабы инновационной деятельности и потребный инновационный потенциал, аналогичный предыдущей стратегии.

Суть *стратегии рыночной ниши* заключается в поиске небольшого, но достаточно привлекательного (в большинстве случаев для малых предприятий) сегмента рынка, где преобладают специфические запросы. При таком подходе внимание предприятия-производителя должно быть целиком сосредоточено на опытно-конструкторских разработках, направленных на создание модифицированной продукции, отвечающей запросам специфического рынка. Отсюда следует, что стратегия рыночной ниши предопределяет масштабы инновационной деятельности и потребный инновационный потенциал, аналогичный стратегии «следования за лидером», но в уменьшенном объеме.

Таким образом, альтернативные стратегии предприятия в инновационной сфере позволяют определить общие подходы к обоснованию масштабов и организационных форм инновационного потенциала.

Заключение

Представленный методологический подход позволяет целенаправленно формировать и развивать инновационный потенциал предприятия в зависимости от индивидуальной комбинации следующих признаков:

- по общекорпоративной стратегии (узкая специализация, диверсификация в родственные отрасли, не профильная диверсификация);
- по стратегии создания конкурентных преимуществ (лидерство в издержках, дифференциация, фокусирование);
- по стратегии предприятия в инновационной сфере (технологическое и рыночное лидерство, следование за лидером, имитация).

Очевидно, что возможны только девять сочетаний этих признаков и, следовательно, столько же различающихся сценариев формирования и развития инновационного потенциала предприятия.

Литература

1. Васюхин О.В., Снетков П.Г. Сущность и роль инноваций в национальной экономике // Научно-технический вестник СПбГУ ИТМО. – 2004. – № 17.
2. Васюхин О.В. Управление инновационной деятельностью предприятия в условиях переходной экономики. – СПб: Нестор, 2001. – 182 с.
3. Павлова Е.А. Система планирования деятельности предприятия. – СПб: Нестор, 2003. – 155 с.
4. Виханский О.С. Стратегическое управление. – М.: Экономистъ, 2006. – 293 с.

Васюхин Олег Валентинович

– Санкт-Петербургский государственный университет информационных технологий, механики и оптики, доктор экономических наук, профессор, prof-vov52@mail.ru

Павлова Елена Александровна

Санкт-Петербургский государственный университет информационных технологий, механики и оптики, кандидат экономических наук, доцент, ea_pavlova@mail.ru

УДК 004.891.2

ИНТЕЛЛЕКТУАЛЬНЫЕ ТЕХНОЛОГИИ ПОДДЕРЖКИ ПОЛЬЗОВАТЕЛЕЙ В ОБЛАСТИ КОМПЬЮТЕРНОГО МОДЕЛИРОВАНИЯ НАНОРАЗМЕРНЫХ СТРУКТУР И НАНОМАТЕРИАЛОВ

А.А. Гуськов, А.В. Дунаев, Ф.В. Подтелкин, А.В. Бухановский

Рассматриваются технологические аспекты разработки экспертной системы для интеллектуальной поддержки в области моделирования наноразмерных структур и наноматериалов.

Ключевые слова: нанотехнологии, экспертные системы, базы знаний.

Активное развитие новых научных областей (к которым, в частности, относятся исследования в области нанотехнологий) связано со сложностями усвоения их результатов сообществом потребителей полученных знаний в условиях отсутствия устоявшейся терминологии. В связи с этим конечный потребитель знаний (ученый или инженер, пользователь предметно-ориентированных программных комплексов) встает перед проблемой выбора наиболее подходящей методики решения поставленной задачи в условиях отсутствия общепризнанных стандартизированных и проверенных подходов. Для обеспечения интеллектуальной поддержки пользователей традиционно используется подход на основе отчуждаемых знаний, интерпретируемых посредством интерактивной экспертной системы (ЭС). ЭС помогает пользователю сформировать терминологически корректную постановку задачи и выбрать для нее адекватное решение (в рамках возможностей программного комплекса), учитывающее не только современное состояние предметной области, но и предпочтения пользователя.

Примером реализации такого подхода является ЭС в составе высокопроизводительного программного комплекса HPC-NASIS для квантовомеханических расчетов и моделирования наноразмерных атомно-молекулярных систем и комплексов [1]. Конструктивно ЭС состоит из четырех компонентов:

- база знаний, записанная в текстовом виде на диалекте языка CLIPS в виде связанного множества продукционных правил;
- компилятор базы знаний из текстового представления во внутреннее объектное дерево с сохранением его как самостоятельной базы знаний для последующего использования;
- скомпилированный бинарный вариант базы знаний, хранящий не только сами знания, но и текущее состояние вычислений: рабочую память, наполненное фактами дерево, активирующие правила;
- интерпретатор, состоящий из машины логического вывода (МЛВ) и обвязки методов, связывающих базу знаний с внешней средой.

База содержит структурированные знания, определяющие приоритеты применения существующих математических моделей, численных методов, алгоритмов и их программных реализаций для выполнения квантово-механических расчетов из первых принципов и для компьютерного моделирования и расчета наноструктур и наноматериалов, включая электронные и фононные характеристики молекулярных систем, оптические и фотоэлектрические свойства ансамблей наночастиц, перенос возбуждения, транспортные свойства нанотрубок, формирование наночастиц, нанотрубок и наносвистков. Экспертные знания, представленные в форме продукций, ассоциированы с гипертекстовыми элементами электронной библиотеки и используют последнюю для предоставления пользователю кратких объяснений касательно методов, инструментов, программных и инфраструктурных решений в области вычислительных нанотехнологий. Для текстового представления базы знаний использован собственный язык описания знаний, совместимый с экспертной системой CLIPS. Диалект записан в расширенной форме Бэкуса–Наура, а для программной реализации использована среда генерирования анализаторов ANTLR [2]. Разработка собственной реализации МЛВ была обусловлена необходимостью активного взаимодействия с другими подсистемами комплекса HPC-NASIS через веб-сервисы, а также лицензионной чистотой используемых библиотек и готовых решений. МЛВ реализована на основе RETE-алгоритма поиска активированных правил [3], использующего двухслойную сеть фильтрации неактивирующих фактов рабочей памяти.

1. Ядро высокопроизводительного программного комплекса для квантово-механических расчетов и моделирования наноразмерных атомно-молекулярных систем и комплексов “HPC-NASIS” // Свидетельство о государственной регистрации программы для ЭВМ № 20010610161 – 2010.
2. The Definitive Antlr Reference: Building Domain-Specific Languages, Terence Parr – Pragmatic Bookshelf, 2007. – 376 с.
3. Forgy C. RETE: A Fast Algorithm for the Many Pattern/Many Object Pattern Match Problem // Artificial Intelligence. – 1982. – № 19. – Р. 17–37.

Гуськов Александр Александрович – НИИ НКТ СПбГУ ИТМО, мл. научн. сотр., i@garfa.spb.ru; Дунаев Антон Валентинович – НИИ НКТ СПбГУ ИТМО, к.т.н., ст. н.с., anton.dunaev@gmail.com; Подтелкин Федор Васильевич – СПбГУ ИТМО, студент, podtelkin@rain.ifmo.ru; Бухановский Александр Валерьевич – д.т.н., профессор, avb_mail@mail.ru

ИССЛЕДОВАНИЕ МЕХАНИЗМОВ ФОРМИРОВАНИЯ КОМПОНЕНТНОГО СОСТАВА МОЛЕКУЛЯРНЫХ СЛОЕВ

Е.Н. Калитеевская, В.П. Крутякова, Т.К. Разумова, А.А. Старовойтов

Работа посвящена изучению равновесного компонентного состава молекулярных слоев на стекле в зависимости от строения концевых групп и длины цепи сопряжения полиметинового красителя.

Ключевые слова: тонкие пленки, полиметиновые красители, стереоизомер.

Гибридные материалы на основе органических красителей, нанесенных на диэлектрические подложки, внедренные в слоистые или пористые материалы, являются потенциальными кандидатами в качестве оптических материалов нового поколения.

На основе симметричных полиметиновых красителей, нанесенных на стекло К8, изучены молекулярные формы, образующиеся в слоях, полученных методом центрифугирования (spin-coating) из концентрированного раствора. Сопоставление спектров поглощения разных слоев показало, что ширина спектра растет с увеличением толщины слоя (эта зависимость имеет характер насыщения), длины полиметиновой цепи сопряжения и электронодонорности концевых групп (Ф). Эти изменения ширины отражают изменения компонентного состава слоя.

В спектрах поглощения тонких слоев присутствуют изначально только *cis*-изомеры (С), при увеличении толщины слоя вначале проявляется *all-trans*-изомер (А), а затем – ассоциаты в виде димеров (d) и J-агрегатов (J) (рис. а). Наблюдаемые зависимости компонентного состава слоя от толщины связаны с асимметрией взаимодействия катионных молекулярных компонентов с отрицательными зарядами стеклянной подложки, приводящей к асимметрии внутримолекулярного распределения электронной плотности молекул. Степень такой асимметрии растет по мере уменьшения толщины слоя и увеличения величины электронодонорности концевых групп. Последняя характеризует степень смещения электронов с концевой группы на полиметиновую цепь и влияет на порядки связи этой цепи, определяющей поглощение в длинноволновой полосе перехода $S_0 \rightarrow S_1$. Влияние электронодонорности на состав молекулярных слоев симметричных полиметиновых красителей близко к закономерностям влияния электронной асимметрии несимметричных полиметиновых молекул на компонентный состав растворов [1]. В растворах симметричных молекул преимущественно присутствует *all-trans*-изомер. Спектр поглощения растворов несимметричных полиметинов заметно уширен за счет существования большого количества *cis*-изомерных форм, относительная концентрация которых растет при увеличении электронной асимметрии (рис. б).

Удлинение цепи сопряжения (увеличение числа *n* карбоцианиновых групп) приводит к уширению спектра поглощения за счет существования большего числа связей цепи, вокруг которых возможны стереоизомерные повороты фрагментов молекул, приводящие к образованию различных *cis*-изомеров (рис. б).

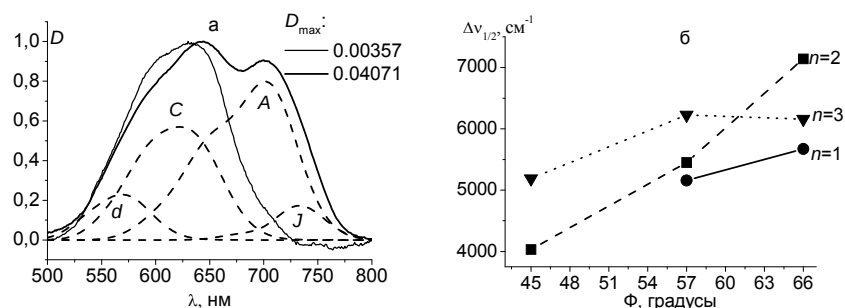


Рис. а – нормированные спектры поглощения и спектры молекулярных компонентов слоев дикарбоцианинового красителя со средней электронодонорностью ($\Phi = 45^\circ$); б – полуширины спектров поглощения молекулярных слоев с $D > 0,03$ полиметинов-гомологов ($n = 1$ – монокарбоцианин, $n = 2$ – дикарбоцианин, $n = 3$ – трикарбоцианин) с разной электронодонорностью концевых групп

Полученные результаты показывают, что спектрально-поляризационные характеристики молекулярных слоев полиметиновых красителей можно варьировать, изменяя степень внутримолекулярной асимметрии электронной плотности, возникающей под влиянием взаимодействия с подложкой.

1. Гадонас Р., Калитеевская Е.Н., Красуаскас В., Пялакаускас А., Разумова Т.К., Тарновский А.Н., Щелкина Е.П. О природе уширения спектров поглощения несимметричных дикарбоцианинов // Оптика и спектроскопия. – 1990. – Т. 68. – № 2. – С. 313–319.

Калитеевская Елена Николаевна – СПбГУ ИТМО, к.ф.-м.н., ст.н.с, lenakalit@yandex.ru; Крутякова Валентина Павловна – СПбГУ ИТМО, к.ф.-м.н., ст.н.с, krutyakova@mail.ru; Разумова Татьяна Константиновна – СПбГУ ИТМО, д.ф.-м.н., вед. н.с., tatiana.razumova@gmail.com; Старовойтов Антон Андреевич – СПбГУ ИТМО, аспирант, anton.starovoytov@gmail.com

РАЗРАБОТКА ФОТОННО-КРИСТАЛЛИЧЕСКОГО ОПТИЧЕСКОГО ВОЛОКНА С БОЛЬШОЙ СЕРДЦЕВИНОЙ

А.Г. Коробейников, К.В. Дукельский, Е.В. Тер-Нерсисянц, Ю.А. Гатчин

Предлагается технология получения фотонно-кристаллического оптического волокна с большой сердцевинной.

Ключевые слова: микроструктурированные световоды, оптическое волокно.

В последнее время большое внимание уделяется разработке и исследованию специальных волокон для передачи мощного лазерного излучения в многомодовом режиме, с повышенной радиационной стойкостью и высокой числовой апертурой [1–3]. Такие световодные волокна, имеющие диаметр световодящей сердцевинной более 10 мкм, называются световодами с большой сердцевинной. Для снижения затухания излучения, проходящего по фотонно-кристаллическому оптическому волокну с большой сердцевинной (микроструктурированному световоду), применяются специальные технологические приемы. Одномодовые и многомодовые световоды могут быть реализованы путем последовательного получения и исследования образцов с различными значениями диаметров сердцевинной, шага структуры и количества образующих сердцевинну элементов. В ходе исследований были отработаны завершающие технологические операции, позволившие получить необходимые микроструктурированные световоды.

Чтобы обеспечить простоту и удобство (технологичность) исходной сборки для вытяжки световода, был применен двухстадийный метод перетяжки поликапиллярной структуры. Надо учитывать, что диаметр световода, удобного для эксплуатации, составляет в среднем 200 мкм, а диаметр воздушных отверстий в укладке, расположенной вокруг большой сердцевинной, составляет в среднем около 3 мкм, что в несколько раз меньше диаметра сердцевинной (в среднем 20 мкм). Такое же соотношение между размерами структурных элементов и наружным диаметром должно иметь место и в исходной заготовке для вытяжки световода. Если внешний диаметр заготовки будет составлять около 20 мм, то диаметр центрального кварцевого штабика будет составлять около 2 мм, а диаметры окружающих его капилляров будут иметь размер около 300 мкм, что затрудняет ручную сборку структуры в заготовке. Для одностадийного способа получения волокна необходимо иметь толщину исходной заготовки менее 10 мм, поэтому применялся двухстадийный метод перетяжки поликапиллярной структуры, несмотря на то, что при одностадийном методе достигается хорошее спекание.

Вначале в опорной кварцевой трубе диаметром 20 мм формировалась структура, образованная центральным кварцевым штабиком и периферийными капиллярами, собранными в плотную гексагональную структуру с некоторым числом порядков окружения и значением диаметра структурных элементов в пределах 1–2 мм (рис. а). Из такой заготовки вытягивалось первичное волокно (предволокну) диаметром 3 мм с сечением, изображенным на рис. б. Диаметр сердцевинной в предволокне составлял 10–20 мкм. Полученное предволокну помещалось в качестве сердцевинной в центр сборки, аналогичной изображенной на рис. а, и происходило вторичное перетягивание заготовки с получением конечного элемента (рис. в).

Таким способом были получены световоды с большим диаметром сердцевинной и с воздушной оболочкой, образованной одним циклом отверстий (волокна с подвешенной сердцевинной), и с оболочкой, имеющей от 2 до 7 циклов отверстий (собственно световоды с большой сердцевинной).

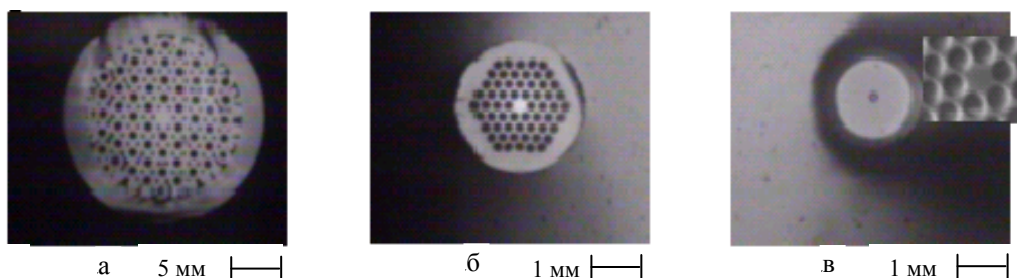


Рис. Двухстадийная перетяжка для поликапиллярной структуры: а – исходная заготовка, б – предволокну, в – микроструктурированное волокно. На вставке показана центральная часть волокна с дополнительным увеличением

1. Issa N.A. High numerical aperture in multimode microstructured optical fibers // Appl. Opt. – 2004. – V. 43. – P. 6191–6197.
2. Ranka J.K., Windeler R.S., Stentz A.J. Optical properties of high-delta air silica microstructure optical fibers // Opt. Lett. 2000. – V. 25. – №. 11. – P. 796–798.
3. Wadsworth W.J., Percival R.M., Bouwmans G., Knight J.C., Russell P.S.J. High power air-clad photonic crystal fibre laser // Opt. Express. – 2003. – № 11. – P. 48–53

Коробейников Анатолий Григорьевич – СПбГУ ИТМО, д.т.н., проф., Korobeynikov_A_G@mail.ru; Дукельский Константин Владимирович – СПбГУ ИТМО, к.т.н., доц., KDukel@GOI.ru; Тер-Нерсисянц Егизе Вавикович – СПбГУ ИТМО, аспирант, volokno@goi.ru; Гатчин Юрий Арменакович – СПбГУ ИТМО, д.т.н., проф., gatchin@mail.ifmo.ru

СТАТИСТИЧЕСКИЕ СВОЙСТВА ГИПОТЕЗ, ПОРОЖДАЕМЫХ ДВУСЛОЙНОЙ НЕЙРОННОЙ СЕТЬЮ С МАТРИЦЕЙ ДВУНАПРАВЛЕННЫХ СВЯЗЕЙ, РЕАЛИЗУЕМОЙ МЕТОДОМ ГОЛОГРАФИИ ФУРЬЕ

З.С. Бекашева, В.Н. Васильев, А.А. Востриков, А.В. Павлов

Определены возможные механизмы порождения абдуктивной, индуктивной и творческой гипотез двуслойной нейронной сетью с матрицей связей, реализуемой методом голографии Фурье.

Ключевые слова: голография, нейронные сети.

В искусственном интеллекте актуальны задачи реализации механизмов творческого мышления и правдоподобных выводов (абдуктивного и индуктивного). По мнению ряда исследователей, эти задачи должны решаться в рамках концепции образного мышления. Поэтому интерес вызывает голографическая парадигма, основанная на аналогиях между работой мозга и голограммы. Принцип образности реализуется голографической обработкой изображений как аналогов паттернов внутренней репрезентации (ПВР) воспринимаемой информации, формирующихся в коре головного мозга.

В работах [1, 2] предложен подход к порождению творческих гипотез при реализации хаотической динамики в схеме голографии Фурье с обращением волновых фронтов в плоскостях изображений и корреляций, реализующий модель двуслойной нейронной сети. В работе [3] предложен подход к реализации правдоподобных рассуждений в той же оптической схеме, но при конвергентной динамике. При этом порождается как паттерн абдуктивной или индуктивной гипотезы, так и паттерн творческой гипотезы. Паттерн порождаемой творческой гипотезы описывается моделью линейной регрессии условий задачи по знаниям. Представляет интерес адекватность порождаемой гипотезы как имеющимся знаниям, представленным в памяти нейронной сети (записанным на голограмме), так и реальности.

Известно, что модель линейной регрессии строит наилучшую по критерию минимума среднего квадрата ошибки оценку для стационарных процессов, в то время как реальные процессы существенно нестационарны. Соответственно, нестационарными должны быть и ПВР. Вместе с тем экспозиционные характеристики регистрирующих сред характеризуются существенной нелинейностью и ограниченностью динамического диапазона, что ведет к пространственно-частотной фильтрации. В работе методом численного моделирования исследовано влияние различных видов фильтрации на статистические характеристики порождаемых творческих гипотез. Использованы фильтры, адекватные реальным характеристикам голографических сред – сигмоидальный, гауссообразный, сумма двух функций Гаусса, экспоненциальный (низкочастотная фильтрация) и полосовой (режектирующий). Показано, что низкочастотная фильтрация ведет к увеличению стационарности по дисперсии, сопряженному с уменьшением дисперсии. Высокочастотная фильтрация ведет к увеличению стационарности по математическому ожиданию, превалирующую роль играют низкие пространственные частоты, режекция которых наиболее эффективно повышает оценку стационарности. При этом по мере увеличения информативности обрабатываемой реализации (числа радиусов корреляции в размере реализации) зависимость оценки стационарности от степени ослабления низких частот снижалась, и оценки для разных фильтров стремились к одному значению.

Таким образом, свойство ограниченности динамического диапазона регистрирующих сред имеет следствием стационаризацию порождаемой сетью творческой гипотезы. При этом оценка стационарности по математическому ожиданию возрастает с ростом информативности реализации, а оценка стационарности по дисперсии, напротив, снижается. В результате изменение параметров фильтра позволяет изменять «творческие способности» сети, выбирая требуемый компромисс между адекватностью гипотезы имеющимся знаниям (стационарная гипотеза) и «полетом фантазии», сопряженным с нестационарностью гипотезы по дисперсии, что предполагает работу в области высоких пространственных частот.

Работы выполнены при поддержке РФФИ, грант 09-01-00165-а.

1. Павлов А.В. О возможности применения фурье-голографии в задаче моделирования творческого мышления: требования к передаточным характеристикам реверсивных голографических сред // Оптический журнал. – 2008. – Т. 75. – № 10. – С. 42–49.
2. Павлов А.В. О возможности моделирования творческого мышления методом голографии Фурье // Труды XI Национальной конференции по искусственному интеллекту. Т. 1. – М.: URSS, 2008. – С. 384–392.
3. Павлов А.В. О возможности реализации абдуктивного и индуктивного выводов методом голографии Фурье // Конгресс по интеллектуальным системам и информационным технологиям AIS-IT'09. 3-10 сентября 2009. Дивноморское. Россия. Т.2. – С.130–139.

Бекашева Зоя Салаватовна – СПбГУ ИТМО, студ., zoya.bekyasheva@phoi.ifmo.ru; Васильев Владимир Николаевич – СПбГУ ИТМО, д.т.н., проф., ректор; Востриков Алексей Алексеевич – СПбГУ ИТМО, студ., lehav@inbox.ru; Павлов Александр Владимирович – СПбГУ ИТМО, к.т.н., доц., pavlov@phoi.ifmo.ru

О МНОГОЧАСТОТНОЙ ГЕНЕРАЦИИ Er:YLF-ЛАЗЕРА С ДИОДНОЙ НАКАЧКОЙ И ПАССИВНЫМ Fe:ZnSe-ЗАТВОРОМ

В.С. Жилина, Д.Ю. Сачков, В.Ю. Храмов

Теоретически исследуются спектрально-энергетические параметры лазерной генерации в кристалле Er:YLF в режиме пассивной модуляции добротности, реализуемой затворами на основе кристаллов Fe:ZnSe.

Ключевые слова: лазер, эрбиевые кристаллы, пассивная модуляция добротности.

Увеличение импульсной мощности лазеров, излучающих в области 3 мкм, способно существенно улучшить результаты воздействия на твердые и мягкие биологические ткани. Одним из возможных способов увеличения импульсной мощности лазера является использование режима модуляции добротности.

В рамках настоящей работы создана теоретическая модель Er:YLF-лазера с диодной накачкой и пассивной модуляцией добротности Fe:ZnSe-затвором. Модель позволяет описать одновременную генерацию на произвольном числе дискретных спектральных линий трехмикронного перехода иона Er^{3+} с учетом процессов нелинейного межзонного взаимодействия по кросс-релаксационным схемам. В работе учитывалась генерация только на трех длинах волн, соответствующих основным максимумам сечения вынужденного излучения кристалла Er:YLF (2,66 мкм, 2,71 мкм и 2,81 мкм [1]), для которых реализована генерация излучения при проведении экспериментальных исследований. Параметры кристалла Fe:ZnSe для этих длин волн заимствованы из работы [2]. Результаты расчетов находятся в хорошем согласии с предварительными данными, полученными на лабораторном макете Er:YLF-лазера с Fe:ZnSe-затвором.

Как и следовало ожидать исходя из данных о сечениях вынужденного излучения и поглощения [1], генерация начинается на длине волны 2,66 мкм, для которой отношение сечения усиления активной среды к сечению поглощения в пассивном затворе является максимальным. После излучения нескольких импульсов на этой длине волны следующие импульсы генерации происходят на длине волны 2,71 мкм, а затем на длине волны 2,81 мкм. Данный эффект смещения длины волны генерации в неселективном резонаторе, в первую очередь, связан с заселением нижнего лазерного уровня $^4I_{13/2}$ и качественно соответствует основным положениям работы [3].

Теоретическое моделирование показало, что существует нетривиальная спектральная зависимость энергии пиков генерации от величины начального коэффициента пропускания пассивного затвора T_0 . В области высоких значений T_0 энергия генерации пика излучения на длине волны 2,66 мкм превышает энергию пиков генерации на больших длинах волн. При уменьшении величины начального пропускания затвора T_0 характер энергетической зависимости меняется на противоположный.

Результаты работы показывают, что в импульсно-периодическом режиме Er:YLF-лазера возможно управление спектром его генерации за счет изменения длительности импульса накачки, средней мощности излучения накачки и скважности следования импульсов. В частности, в неселективном резонаторе осуществим режим генерации в виде последовательности чередующихся по спектру генерации импульсов излучения на длинах волн 2,66 и 2,81 мкм, причем импульсы генерации на длине волны 2,66 мкм могут завершать последовательность импульсов генерации в пачке. Данный режим может найти применение при решении ряда медицинских задач, для которых важен больший коагуляционный эффект при облучении биоткани коротковолновым излучением при завершении хирургических операций.

1. Labb'e C., Doualan J.-L., Girard S., Moncorg'e R., Thuau M. Absolute excited state absorption cross section measurements in $\text{Er}^{3+}:\text{LiYF}_4$ for laser applications around 2,8 μm and 551 nm // J. Phys.: Condens. Matter. – 2000. – V. 12. – P. 6943–6957.
2. Kernal J., Fedorov V.V., Gallian A., Mirov S.B., Badikov V.V. 3,9–4,8 μm gain-switched lasing of Fe:ZnSe at room temperature // Optics Express. – 2005. – V. 13. – I. 26. – P. 10608–10615.
3. Иночкин М.В., Назаров В.В., Сачков Д.Ю., Хлопонин Л.В., Храмов В.Ю. Динамика спектра генерации трехмикронного Er:YLF-лазера при полупроводниковой накачке // Оптик. журнал. – 2009. – Т. 76. – № 11. – С. 62–67.

Жилина Вера Сергеевна – СПбГУ ИТМО, студ., vera-sergeevna@yandex.ru; Сачков Дмитрий Юрьевич – СПбГУ ИТМО, аспирант, dsachkov@gmail.com; Храмов Валерий Юрьевич – СПбГУ ИТМО, д.т.н., проф., зав. кафедрой, vkhramov@gmail.com

OPTICAL AND OPTICAL ELECTRONIC SYSTEMS OPTICAL TECHNOLOGIES

ON THE ISSUE OF WOOD SPECIES IDENTIFICATION WITH THE USE OF SPECTRUM ANALYSIS METHODS

A. Voronin, E. Smirnova, A. Smirnov

The problems and conceptions of computer-based system construction for wood species identification (recognition) on the basis of spectral information analysis are described in this report.

Keywords: identification, spectrum, spectrophotometer, diffraction grating, integrating sphere, wood species.

Andrey Voronin – Saint Petersburg State University of Information Technologies, Mechanics and Optics, Postgraduate, V-electronics@mail.ru

Elena Smirnova – Saint Petersburg State University of Information Technologies, Mechanics and Optics, Postgraduate, smirnova-elen@yandex.ru

Alexander Smirnov – Saint Petersburg State University of Information Technologies, Mechanics and Optics, D.Sc, Professor.

LASER CLEANING OF HISTORICAL MONUMENTS

V. Parfenov, A. Gerashchenko, M. Gerashchenko, I. Grigoryeva

Results of experimental work on laser cleaning of historical copper monuments are presented. Surface cleaning was carried out by means of pulsed Nd:YAG laser (wave length of 1064 nm, pulse duration of 30-110 μ s, energy density 10-25 J/cm²). Efficiency of laser treatment was controlled using X-ray fluorescence spectrometer. Performed investigations demonstrated a high efficiency of laser cleaning and possibility of laser use with specified parameters in restoration of copper monuments.

Keywords: laser cleaning, pulsed Nd:YAG laser, monuments of history and culture, restoration, X-ray fluorescence spectral analysis.

Vadim Parfenov – Saint Petersburg Electrotechnical University "LETI" (ETU), Associate Professor, PhD, Associate Professor, vadim_parfenov@mail.ru

Anastasia Gerashchenko - Saint Petersburg Electrotechnical University "LETI" (ETU), Postgraduate, opinio@inbox.ru

Mikhail Gerashchenko - Saint Petersburg Electrotechnical University "LETI" (ETU), Postgraduate, sir_kofaioh@inbox.ru

Irina Grigoryeva - Saint Petersburg Electrotechnical University "LETI" (ETU), Postgraduate, grigoryevai@rambler.ru

PROCESS OF LAYERING FORMATION RESEARCH AT ACID LEACHING OF ALKALI BOROSILICATE GLASS

I. Meshkovskiy, A. Olekhovich

The model of layering formation at acid leaching of alkali borosilicate glass Na 7/23 [1] is presented in this article. Elements, created by acid leaching, can be used as optical elements.

Keywords: acid leaching, simulation, layering, strata, porosity, alkali borosilicate glass, porous matrix, two-phase glass.

Igor Meshkovskiy – Saint Petersburg State University of Information Technologies, Mechanics and Optics, Department Head, D.Sc., Professor, igorkm@runnet.ru

Artyom Olekhovich - Saint Petersburg State University of Information Technologies, Mechanics and Optics, Postgraduate, romanart1@inbox.ru

PHOTONICS AND OPTICAL INFORMATICS

INTERACTION LAWS OF COUNTER-PROPAGATING LIGHT WAVES CONSISTING OF FEW OSCILLATIONS IN THE NONLINEAR DIELECTRIC MEDIA

E. Buyanovskaya, S. Kozlov

Mechanisms of counter-propagating light waves interaction from low fluctuation count in non-linear dielectric mediums are theoretically considered. The analytical solution of the equation describing the field dynamics of counter-propagating light waves from low fluctuation count, interacting in the dielectric medium with instantaneous cube nonlinearity is obtained. Changes of light pulse time and spectral structure under counter-propagating pulse interaction are given. The efficiency of these phenomena is shown to be determined by oncoming pulse energy and is independent from its duration and spectral structure.

Keywords: pulses consisting of the few oscillations of light field, interaction of counter-propagating waves, nonlinear dielectric media.

Elizaveta Buyanovskaya – Saint Petersburg State University of Information Technologies, Mechanics and Optics, Postgraduate, lee.buyanovskaya@gmail.com

Sergei Kozlov – Saint Petersburg State University of Information Technologies, Mechanics and Optics, Dean, Department Head, D.Sc., Professor

ANALYSIS AND SYNTHESIS OF COMPLEX SYSTEMS UNITS AND APPLIANCES OF ENGINEERING SYSTEMS

ALGEBRAIC CONDITIONS OF TOTAL SYNCHRONISM FOR MULTIVARIATE DYNAMIC PLANTS

N. Dudarenko, M. Polyakova, A. Ushakov

The algebraic conditions problem of total synchronism for multivariate dynamic plants is concerned in this article. The problem is solved with geometric interpretation. Synchronous process is described as a motion in linear casing along determined vector.

Keywords: algebraic organization, multivariate plants, total synchronous process, eigenvector, linear casing.

Natalia Dudarenko – Saint Petersburg State University of Information Technologies, Mechanics and Optics, Associate professor, PhD, Dudarenko@yandex.ru

Maya Polyakova – Saint Petersburg State University of Information Technologies, Mechanics and Optics, Postgraduate, 12noch@mail.ru

Anatoliy Ushakov – Saint Petersburg State University of Information Technologies, Mechanics and Optics, Professor, D.Sc., Professor, Ushakov-AVG@yandex.ru

ALGORITHM OF ACCELERATED RESETTING SOLUTION OF LINEAR ALGEBRAIC EQUATION SYSTEMS AND ITS USE IN MATHEMATICAL MODELING OF ELECTRONIC DEVICES

M. Revin, N. Savyolov

Numerical experiments with the developed program, realizing the variant of a Gaussian elimination method which allows getting considerable advantage in calculation speed when resetting solution of the linear algebraic equations systems are stated in the article. The analysis of a new method in comparison with Sherman-Morrison's known method is given. Application of this updating for acceleration of condition equations formation process has allowed developing the microprocessor system capable to supervise a temperature operating mode of electrotechnical object in real time.

Keywords: method, modeling, experiments, the condition equations, microprocessor system, electrotechnical object.

Mikhail Revin - South - Russian State Technical University (Novocherkassk Polytechnic Institute), Postgraduate, misharevin86@mail.ru

Nikolay Savyolov - South - Russian State Technical University (Novocherkassk Polytechnic Institute), Associate professor, PhD, savelov@mail.ru

COMPUTER SYSTEMS AND INFORMATION TECHNOLOGIES

A SET-THEORETIC APPROACH TO THE LOGICAL INFERENCE IN KNOWLEDGE BASES

I. Bessmertny

The paper concerns to the problem of combinatorial complexity of inference in artificial intelligence applications based on production rules. An approach to speeding up the knowledge acquisition is proposed that specifies indexing the facts and operations with sets of indices to prune unnecessary facts. In contrast to the most known algorithms that provide pre-collecting facts in nodes of the search tree, the indices exist apart from the rule sets, that allows disjointing bases of facts and bases of rules logically and physically and to simplify the knowledge base updating. An opportunity of alternation the matching of rules to set-theoretic operation on tuples of variables is proven.

Keywords: artificial intelligence, facts indexing, inference.

Igor Bessmertny - Saint Petersburg State University of Information Technologies, Mechanics and Optics, Associate Professor, PhD, Associate Professor, igor_bessmertny@hotmail.com

COMPLEX SIGNAL PREDICTION BASED ON DYNAMIC SYSTEM ALLOCATION LIMITS

I. Sidorkina, A. Egoshin, D. Shumkov, P. Kudrin

The paper proposes two methods for predicting complex signal, based on multilayer perceptron and SVM, implemented in the software complex. The training set for the predictor is based on the allocation of border implementations of dynamic systems in the signal.

Keywords: forecasting, change-point detection, fractal, dynamic systems, neural networks, SVM.

Irina Sidorkina - Mari State Technical University, Dean, D.Sc., Professor, dekan_fivt@mail.ru

Alexey Egoshin - Mari State Technical University, Postgraduate, g12@bk.ru

Dmitry Shumkov - Mari State Technical University, Postgraduate, dimavova@mail.ru

Pavel Kudrin - Mari State Technical University, Postgraduate, pavelkudrin@gmail.com

NON-STATIONARY IMAGES CLASSIFICATION AND DEVELOPMENT OF SOURCE CODING ALGORITHMS ESTIMATION METHOD

I. Zubakin, Sh. Fahmi

The algorithm of an efficiency estimation of various image compression methods (methods of source coding) on the basis of the non-stationary measure is offered. The technique of efficiency estimation, using compact representative sample of images, ranged according to non-stationary and broadbandness is generated

Keywords: non-stationary measure, broadbandness, classification of images, compact representative sample of images, statistics calculation algorithm.

Igor Zubakin - Saint Petersburg Electrotechnical University "LETI" (ETU), Postgraduate, zubakin_@mail.ru

Shakeeb Fahmi - Saint Petersburg Electrotechnical University "LETI" (ETU), Associate Professor, PhD, shakeebf@mail.ru

STANDARD PROCEDURE OF NATURAL ENVIRONMENT OBJECTS MEASURING BY THE GDV/EPC METHOD

K. Korotkov, D. Orlov, E. Velichko, Yu. Gatchina

This work is aimed on development of procedure of natural environment objects measurements on gas discharge visualization or electro-photon capturing (GDV/EPC) equipment. Main destabilizing factors were revealed during this research and ways of overcoming them were developed. As a result a procedure of measurements was designed, that provides stable and reproducible results.

Keywords: GDV/EPC, procedure of measurements, natural environment objects.

Konstantin Korotkov – Saint Petersburg State University of Information Technologies, Mechanics and Optics, Professor, D.Sc., Professor, korotkov2000@gmail.com

Dmitry Orlov – Saint Petersburg State University of Information Technologies, Mechanics and Optics, Postgraduate, mityaorlov@gmail.com, mitya-9@mail.ru

Elena Velichko – Saint Petersburg State University of Information Technologies, Mechanics and Optics, Postgraduate, heleonores@mail.ru

Yulia Gatchina - Saint Petersburg State University of Information Technologies, Mechanics and Optics, economist, Gatchina@mail.ifmo.ru

PERFORMANCE EVALUATION MODEL FOR DISTRIBUTED SYSTEMS AND TASKS WITH VARIABLE SIZE

A. Khritankov

We develop a performance model of solving computational tasks with variable size on distributed systems. For this purpose we extend a performance model of systems with schedule to allow for stochastic nature of global optimization problem. In addition, we consider an experimental method to estimate task size and calculate a confidence interval for efficiency. Finally, we show that a real distributed system solving a cluster conformation problem follows the model within estimated accuracy.

Keywords: distributed computing, performance evaluation, global optimization.

Anton Khritankov - Moscow Institute of Physics and Technology (State University), Postgraduate, anton.khritankov@gmail.com

ACTOR EXTENSION OF JAVA LANGUAGE IN MPS ENVIRONMENT

A. Zhukova, M. Mazin

The actor extension described in this article adds parallel functional constructions to imperative Java language. Such an extension is useful as it makes parallel programming in Java much easier. Actor model implemented by the extension is well used in functional programming language Erlang. The choice of Meta Programming System (MPS) as the actor extension creation tool made it possible to create an integrated development environment for the extension automatically, moreover, the extension is compatible with other language extensions created in MPS.

Keywords: actor model, problem-oriented extension, parallel programming, language-oriented programming.

Anna Zhukova – Saint Petersburg State University, Student, anna.zhukova@math.spbu.ru

Maxim Mazin - Saint Petersburg State University of Information Technologies, Mechanics and Optics, Postgraduate, mazine@rain.ifmo.ru

MATERIALS TECHNOLOGY AND NANOTECHNOLOGIES

NANOWHISKER GROWTH UNDER ELECTRON BEAM IMPACT: MATHEMATICAL MODEL

S. Chivilikhin, A. Golubok, I. Mukhin

We present mathematical model of nanowhisker growth at the conductor surface in the vacuum chamber of scanning electron microscope under focused electron beam impact. Carbon ions with positive charge come out from chamber surface and are captured by electron beam field. Ions move along electron beam and subside at

nanowhisker surface. This process provides nanowhisker growth. In that model electrical field and beam potential are approximated by electrical field and potential of uniformly charged wire. For that simplification Vlasov equation is written, that gives equation of nanowhisker surface movement. By numerical methods we simulated and visualized the process of nanowhisker growth. Computing program builds dynamic grid, which simulates growing whisker – dynamic growth in real time. Obtained results go well with experimental data.

Keywords: nanowhisker, mathematical model of growth, equation of surface movement.

Sergey Chivilikhin - Saint Petersburg State University of Information Technologies, Mechanics and Optics, Associate Professor, PhD, leading scientist, sergey.chivilikhin@gmail.com

Alexander Golubok – Analytical Instrument Engineering Institute RAS, Deputy director for scientific work, D.Sc., aogolubok@mail.ru

Ivan Mukhin - Academic Physics and Technology University (APTU) RAS, Postgraduate, imukhin@yandex.ru

THERMAL PHYSICS AND THEORETICAL HEAT ENGINEERING

APPLICATION OF SUPERCOMPUTER TECHNOLOGIES FOR THE THERMAL ANALYSIS OF RADIO-ELECTRONIC DEVICES

I. Shibut

The main stages of computer thermal analysis on the basis of supercomputer technologies are considered. Theoretical and methodical approaches to preparation and carrying out the thermal analysis of radio-electronic devices are stated. The example of the thermal analysis of the TV “Horizont” using the multiprocessor version of package LS-DYNA on the basis of supercomputer “SKIF k - 1000” is considered.

Keywords: thermal analysis, supercomputer, radio-electronic devices.

Irina Shibut - United Institute of Informatics Problems of the National Academy of Sciences of Belarus (UIIP NAS of Belarus), Postgraduate, irina_shibut@mail.ru

COMPUTER-AIDED DESIGN SYSTEMS

HARDWARE-SOFTWARE DESIGN OF COMPLEX FUNCTIONAL UNITS ON “SYSTEM ON CHIP”

I. Zubakin, Sh. Fahmi, S. Shagarov

With the increase of possibilities of VLSI design technology, «system on chip» becomes more and more attractive and necessary means of integrated EDA creation with programmable structure. Certain design stages (hardware and program parts of the circuit project) and circuit verification as a whole on the example of the multiplier design with hardware-software EDA means on the basis of «system on chip» technology are considered in this article.

Keywords: system on chip, level of register transfers, complex functional block, programmable logic integrated circuit.

Igor Zubakin - Saint Petersburg Electrotechnical University "LETI" (ETU), Postgraduate, zubakin_@mail.ru

Shakeeb Fahmi - Saint Petersburg Electrotechnical University "LETI" (ETU), Associate Professor, PhD, shakeebf@mail.ru

Sergei Shagarov - Saint Petersburg Electrotechnical University "LETI" (ETU), student, shagarovss@mail.ru

OPTIMIZATION OF GRAPH ISOMORPHISM IDENTIFICATION ALGORITHMS

S. Moskalenko, Yu. Gatchin

A row of optimization moments for modern algorithms of graph isomorphism identification is proposed in this article. Integration of these moments into GI algorithms will increase the speed of their work with calculation accuracy preservation.

Keywords: graph theory, image recognition, graph isomorphism detection.

Stanislav Moskalenko - Saint Petersburg State University of Information Technologies, Mechanics and Optics, Postgraduate, stan.moskalenko@gmail.com

Yuri Gatchin - Saint Petersburg State University of Information Technologies, Mechanics and Optics, Department Head, D.Sc., Professor, gatchin@mail.ifmo.ru

INFORMATION SECURITY METHODS AND SYSTEMS

THREAT EVALUATION MODEL OF INFORMATION SAFETY ON THE BASIS OF HACKER CONFERENCE MATERIALS

E. Khusainova

The problem of threat model creation is actual and not enough developed. Threat models should become starting points for designing of the future computer and information protection systems. Construction of threat model within the hacker conferences will allow improving reliability and security of information systems and software products.

Keywords: hacker conferences, threat model, information security, forecasting of attacks.

Elvira Khusainova - Saint Petersburg State University of Information Technologies, Mechanics and Optics, Postgraduate, elechka.de@gmail.com

THE EFFICIENCY ESTIMATION OF THE INTRUDER ATTACKS IN THE PROCESS OF HIS MODEL CREATION

A. Spivak

The design of the security information system always has the stage of the intruder model creation. A well known approach is the usage of violator characteristics as the basis for estimation of his influence on the security system. The changing of the protected object behavior can be used for the estimation of the security violator actions.

Keywords: model, security, intruder, estimation, information.

Anton Spivak - Saint Petersburg State University of Information Technologies, Mechanics and Optics, Postgraduate, anton.spivak@gmail.com

ECONOMICS AND FINANCES. MANAGEMENT

STRATEGY OF FORMATION AND DEVELOPMENT OF THE INDUSTRIAL ENTERPRISE INNOVATIVE POTENTIAL

O. Vasyukhin, E. Pavlova

The authors represent the system of strategic decisions of the innovative-active enterprise. The main types of the corporate strategy directed on innovative development of the enterprise are defined in this article. Readers can get information about the distinctive features of strategy for the creation of the competitive advantages that allow defining of the approaches for substantiation of enterprise innovative potential structure. The authors tried to represent the meaning of the innovative sphere strategies which are defined on the scale of innovative activity. The offered methodological approach will allow developing of the potential purposefully according to individual specificity of enterprise activity.

Keywords: strategies of the enterprise development, innovations, innovative potential.

Oleg Vasyukhin - Saint Petersburg State University of Information Technologies, Mechanics and Optics, Head of Department, D. Sc., Professor, prof-vov52@mail.ru

Elena Pavlova – Saint Petersburg State University of Information Technologies, Mechanics and Optics, Associate professor, PhD, Associate professor, ea_pavlova@mail.ru

BRIEF REPORTS

INTELLIGENT TECHNOLOGIES OF USERS' SUPPORT IN THE FIELD OF NUMERICAL QUANTUM-CHEMICAL COMPUTING OF NANOSTRUCTURE AND NANOMATERIALS MODELING

A. Gus'kov, A. Dunaev, F. Podtelkin, A. Bukhanovsky

Technical aspects of expert system development for intelligent decision support in the field of computer simulation of nanoscale systems and nanomaterials are under consideration.

Keywords: nanotechnology, expert systems, knowledge bases.

Alexander Gus'kov - Research Institute of Saint Petersburg State University of Information Technologies, Mechanics and Optics, applicant, i@garfa.spb.ru

Anton Dunaev - Research Institute of Saint Petersburg State University of Information Technologies, Mechanics and Optics, senior staff scientist, PhD, anton.dunaev@gmail.com

Fyodor Podtelkin – Saint Petersburg State University of Information Technologies, Mechanics and Optics, student, podtelkin@rain.ifmo.ru

Alexander Bukhanovsky – Saint Petersburg State University of Information Technologies, Mechanics and Optics, Professor, D.Sc., Professor, avb_mail@mail.ru

STUDY OF MECHANISMS RESPONSIBLE FOR THE MOLECULAR LAYERS COMPOSITION FORMATION

E. Kaliteevskaya, V. Krutyakova, T. Razumova, A. Starovoytov

The equilibrium composition of the molecular layers on glass was studied depending on the end groups structure and conjugation chain length of polymethine dye.

Keywords: thin films, polymethine dyes, stereoisomer.

Elena Kaliteevskaya – Saint Petersburg State University of Information Technologies, Mechanics and Optics, senior staff scientist, PhD, lenakalit@yandex.ru

Valentina Krutyakova - Saint Petersburg State University of Information Technologies, Mechanics and Optics, senior staff scientist, PhD, krutyakova@mail.ru

Tatyana Razumova - Saint Petersburg State University of Information Technologies, Mechanics and Optics, leading staff scientist, D.Sc., senior scientist, tatiana.razumova@gmail.com

Anton Starovoytov - Saint Petersburg State University of Information Technologies, Mechanics and Optics, junior staff scientist, anton.starovoytov@gmail.com

THE BIG CORE PHOTONIC CRYSTAL OPTICAL FIBER DEVELOPMENT

A. Korobeynikov, K. Dukel'skiy, E. Ter-Nersesyants, Yu. Gatchin

The technology of the big core photonic crystal optical fiber generation is offered.

Keywords: microstructured optical fiber, optical fiber.

Anatoliy Korobeynikov – Saint Petersburg State University of Information Technologies, Mechanics and Optics, Professor, D.Sc., Professor, Korobeynikov_A_G@mail.ru

Konstantin Dukel'skiy - Saint Petersburg State University of Information Technologies, Mechanics and Optics, Associate professor, PhD, KDukel@GOI.ru

Egishe Ter-Nersesyants - Saint Petersburg State University of Information Technologies, Mechanics and Optics, Postgraduate, volokno@goi.ru

Yuri Gatchin - Saint Petersburg State University of Information Technologies, Mechanics and Optics, Department Head, D.Sc., Professor, gatchin@mail.ifmo.ru

STATISTICAL PROPERTIES OF HYPOTHESES CAUSED BY TWO-LAYER NEURAL NETWORK WITH BIDIRECTIONAL INTERCONNECTIONS, IMPLEMENTED BY FOURIER-HOLOGRAPHY TECHNIQUE

Z. Bekyasheva, V. Vasil'ev, A. Vostrikov, A. Pavlov

Possible mechanisms for creative, inductive and abductive hypotheses are put forward by two layered neural network with bidirectional interconnections, implemented by Fourier-holography technique.

Keywords: holography, neural networks.

Zoya Bekyasheva – Saint Petersburg State University of Information Technologies, Mechanics and Optics, Student, zoya.bekyasheva@phoi.ifmo.ru

Vladimir Vasil'ev – Saint Petersburg State University of Information Technologies, Mechanics and Optics, Rector, D.Sc., Professor, vasiliev@mail.ifmo.ru

Alexey Vostrikov – Saint Petersburg State University of Information Technologies, Mechanics and Optics, Student, lehav@inbox.ru

Alexander Pavlov – Saint Petersburg State University of Information Technologies, Mechanics and Optics, Associate professor, PhD, senior staff scientist, pavlov@phoi.ifmo.ru

ABOUT MULTIWAVE GENERATION OF Er:YLF LASER WITH DIODE PUMPING AND PASSIVE Fe:ZnSe Q-SWITCH

V. Gilina, D. Sachkov, V. Khramov

The spectral and energy properties of Er:YLF laser with passive Fe:ZnSe Q-switch are investigated and reported.

Keywords: erbium crystal, laser, passive Q-switch.

Vera Zhilina – Saint Petersburg State University of Information Technologies, Mechanics and Optics, Student, vera-sergeevna@yandex.ru

Dmitry Sachkov - Saint Petersburg State University of Information Technologies, Mechanics and Optics, Postgraduate, dsachkov@gmail.com

Valery Khramov – Saint Petersburg State University of Information Technologies, Mechanics and Optics, Department Head, PhD, Professor, vkhramov@gmail.com

Подписано к печати 15.02.2010 Тираж 350 экз. Заказ № 14(66)

Отпечатано в учреждении «Университетские телекоммуникации»
Адрес: 197101, Санкт-Петербург, Кронверкский пр., 49

Журнал зарегистрирован Федеральной службой по надзору в сфере связи
и массовых коммуникаций, свидетельство ПИ № ФС77-33466 от 10.10.2008 г.

Подписной индекс 47 197 по каталогу агентства Роспечать
(Газеты. Журналы)

© Санкт-Петербургский государственный университет
информационных технологий, механики и оптики, 2010