

Запорізький національний технічний університет



Радіоелектроніка Інформатика Управління

2(18)'2007

Науковий журнал

Видається двічі на рік

Зареєстрований **29 січня 2003 року**
Державним комітетом інформаційної політики,
телебачення та радіомовлення України.

Свідоцтво – серія **КВ № 6904**

Засновник і видавник – *Запорізький національний технічний університет*

Запоріжжя, ЗНТУ
2007

Передплатний індекс – 22914

Журнал «Радіоелектроніка. Інформатика. Управління» видається з 1999 року.

Журнал входить до переліку наукових фахових видань України, в яких можуть публікуватися результати дисертаційних робіт на здобуття наукових ступенів доктора і кандидата технічних наук та фізико-математичних наук за спеціальністю «Радіофізика». Статті, що публікуються в журналі, реферуються в базах даних та РЖ ВІНТІ (Росія) і «Джерело» (Україна)

РЕДАКЦІЙНА КОЛЕГІЯ

Головний редактор – д.т.н. Потапенко Є. М.

Заст. головного редактора – к.т.н. Дубровін В. І.

Члени редколегії :

д.ф.-м.н. Ахметшин А. М.
д.т.н. Волков О. В.
д.ф.-м.н. Горбань О. М.
д.ф.-м.н. Горр Г. В.
д.т.н. Гостєв В. І.
д.ф.-м.н. Дробахин О. О.
д.ф.-м.н. Ковальов О. М.

д.ф.-м.н. Корніч Г. В.
д.ф.-м.н. Онуфрієнко В. М.
д.т.н. Павлов О. А.
д.т.н. Піза Д. М.
д.ф.-м.н. Погосов В. В.
д.т.н. Труфанов І. Д.
д.ф.-м.н. Чумаченко В. П.

Редакційно-видавнича рада:

Беліков С. Б., Внуков Ю. М., Піза Д. М., Яримбаш С. Т., Потапенко Є. М., Дубровін В. І., Петруша Ю. П., Флора В. Д., Соболь Ю. О., Дедков М. В., Рибіна Ю. А., Табунщик Г. В., М'ясников М. Л., Богданова Л. Ф., Кобзар Н. М., Федоренко Е. С.

Рекомендовано до видання вченою радою Запорізького національного технічного університету, протокол № 4 від 26.11.2007 р.

Рукописи проходять незалежне рецензування з залученням провідних фахівців, за результатами якого редакційна колегія приймає рішення про опублікування

Журнал зверстаний редакційно-видавничим відділом
Запорізького національного технічного університету

Адреса редакції: 69063, м. Запоріжжя, вул. Жуковського, 64, ЗНТУ,
редакція журналу «РІУ»

Тел: (061) 769-82-96 – редакційно-видавничий відділ
Факс: (061) 764-21-41
E-mail: rvv@zntu.edu.ua
TQM@zntu.edu.ua

Web-сайт журналу розташований в Internet за адресою: <http://www.zntu.edu.ua/base/news2005/ric.htm>

© 2007 Запорізький національний технічний університет

РАДИОЭЛЕКТРОНИКА ТА ТЕЛЕКОМУНІКАЦІЇ 5

Н. П. Гапоненко, Е. В. Огренич
ПРОЕКТИРОВАНИЕ ГОФРИРОВАННЫХ
ТЕПЛООТВОДОВ..... 5

С. П. Гулин
АНАЛИЗ СПЕКТРА ОТКЛИКА НЕЛИНЕЙНОСТИ,
ПРЕДСТАВЛЕННОЙ ФУНКЦИЕЙ
ДИНАМИЧЕСКОГО НАСЫЩЕНИЯ,
ПРИ МНОГОЧАСТОТНОМ ВОЗДЕЙСТВИИ 8

*А. А. Захарченко, В. Е. Кутний, И. М. Прохорец,
А. В. Рыбка, М. А. Хажмурадов*
ОПРЕДЕЛЕНИЕ ПАРАМЕТРОВ ПЕРЕНОСА
ЗАРЯДА В СДТЕ ДЕТЕКТОРАХ γ -ИЗЛУЧЕНИЯ
МЕТОДАМИ МАТЕМАТИЧЕСКОГО
МОДЕЛИРОВАНИЯ13

*Д. М. Пиза, Н. П. Чернобородова,
М. П. Чернобородов*
УСОВЕРШЕНСТВОВАНИЕ МЕТЕОКАНАЛА
АЭРОДРОМНОГО РАДИОЛОКАЦИОННОГО
КОМПЛЕКСА «ДНЕПР-А» 16

В. И. Псарев, Л. А. Пархоменко, Ю. В. Пьянкова
О ТЕХНИКЕ ИДЕНТИФИКАЦИИ ПРИ
СИСТЕМНОМ АНАЛИЗЕ ТОНКИХ
ОСТРОВКОВЫХ ПЛЕНОК 20

МАТЕМАТИЧНЕ ТА КОМП'ЮТЕРНЕ МОДЕЛЮВАННЯ.....26

А. Ю. Берко, В. А. Висоцька
ЗАСТОСУВАННЯ МЕРЕЖ ПЕТРИ ДЛЯ
МОДЕЛЮВАННЯ СИСТЕМ ЕЛЕКТРОННОЇ
КОМЕРЦІЇ.....26

С. Н. Герасим, М. А. Козлов
СТОХАСТИЧЕСКАЯ МОДЕЛЬ ИЗМЕНЯЮЩИХСЯ
МЕЖПРОИЗВОДСТВЕННЫХ ПОТОКОВ.....33

В. А. Гороховатский, С. В. Кузьмин
СТАТИСТИЧЕСКИЙ АНАЛИЗ ЗНАЧЕНИЙ
ХАРАКТЕРНЫХ ПРИЗНАКОВ
ПРИ РАСПОЗНАВАНИИ ИЗОБРАЖЕНИЙ.....37

В. И. Долгов, А. В. Нелася
ГЕОМЕТРИЧЕСКИЙ ПОДХОД К СЛОЖЕНИЮ
ДИВИЗОРОВ ГИПЕРЭЛЛИПТИЧЕСКОЙ
КРИВОЙ44

Т. С. Дьячук, Р. К. Кудерметов
РАСПРЕДЕЛЕНИЕ ЗАДАНИЙ В
МНОГОПРОЦЕССОРНОЙ СИСТЕМЕ.....51

А. А. Егошина
ФОРМАЛЬНАЯ МОДЕЛЬ
СЛОВООБРАЗОВАТЕЛЬНОЙ СЕМАНТИКИ54

В. Н. Журавлев, Е. А. Архипова
АНАЛИЗ МЕТОДА РАСЧЕТА ПАРАМЕТРА
ЭФФЕКТИВНОСТИ МАСКИРОВАНИЯ РЕЧИ
В ТЕХНИЧЕСКИХ КАНАЛАХ УТЕЧКИ57

М. Б. Ильяшенко
АЛГОРИТМ ОПТИМАЛЬНОГО
РЕЗЕРВИРОВАНИЯ РАСПРЕДЕЛЕННЫХ
ВЫЧИСЛИТЕЛЬНЫХ РЕСУРСОВ
ПО КРИТЕРИЮ ЭФФЕКТИВНОГО
ИСПОЛЬЗОВАНИЯ ВЫЧИСЛИТЕЛЬНЫХ
УЗЛОВ 64

С. В. Курапов
КОНСТРУКТИВНЫЙ АЛГОРИТМ ДЛЯ
РАСКРАСКИ КУБИЧЕСКИХ ГРАФОВ..... 69

А. И. Липчанский, А. А. Давыдов
КЛАССИФИКАЦИЯ ОСНОВНЫХ ПОКАЗАТЕЛЕЙ
ДЛЯ ОЦЕНКИ НАДЕЖНОСТИ И
ДИАГНОСТИРУЕМОСТИ КОМПЬЮТЕРНЫХ
СИСТЕМ 76

М. А. Новотарский
АЛГЕБРА ПРОЦЕССОВ ДЛЯ ИМИТАЦИОННОГО
МОДЕЛИРОВАНИЯ СЛОЖНЫХ ДИСКРЕТНЫХ
СИСТЕМ 79

*А. Л. Становский, Т. В. Лысенко, Т. И. Носенко,
Е. В. Лысенко*
СИНХРОНИЗИРУЮЩЕЕ УПРАВЛЕНИЕ
ЭКСТРЕМУМАМИ ОТДАЛЕННЫХ
СОСТОЯНИЙ 86

НЕЙРОІНФОРМАТИКА ТА ІНТЕЛЕКТУАЛЬНІ СИСТЕМИ90

Е. В. Бодянский, Н. Е. Кулишова
ИНТЕРПОЛИРУЮЩАЯ НЕЙРО-ФАЗЗИ-СЕТЬ
ДЛЯ МОДЕЛИРОВАНИЯ ЦВЕТОПЕРЕДАЧИ
ПЕЧАТНОЙ СИСТЕМЫ.....90

Т. В. Киррич, В. И. Дубровин
АНАЛИЗ МОДИФИКАЦИЙ
САМООРГАНИЗУЮЩИХСЯ КАРТ КОХОНЕНА
ПО КРИТЕРИЯМ РЕГУЛЯРНОСТИ И ТОЧНОСТИ
АППРОКСИМАЦИИ96

О. Г. Руденко, А. В. Заец, Ю. Э. Ткач
УПРАВЛЕНИЕ НАГРУЗКОЙ В РАСПРЕДЕЛЕННОЙ
КОМПЬЮТЕРНОЙ СИСТЕМЕ ПРИ ПОМОЩИ
НЕЙРОННЫХ СЕТЕЙ 99

А. В. Савельев
ОБЗОР ИЗОБРЕТЕНИЙ В ОБЛАСТИ
НЕЙРОКИБЕРНЕТИКИ
И НЕЙРОМОДЕЛИРОВАНИЯ 101

ПРОГРЕСИВНІ ІНФОРМАЦІЙНІ ТЕХНОЛОГІЇ 112

Ю. С. Афонин, В. И. Дубровин, Т. В. Манило
ПРИМЕНЕНИЕ МЕТОДОВ ВЕЙВЛЕТ-АНАЛИЗА
И ПРЕОБРАЗОВАНИЯ ФУРЬЕ В ЗАДАЧАХ
ДИАГНОСТИКИ ОБЪЕКТОВ
ПО ВИБРОАКУСТИЧЕСКОМУ СИГНАЛУ 112

Г. Г. Киричек
ЕЛЕКТРОННА БІБЛІОТЕКА ЯК ОБ'ЄКТ
ОСВІТНІХ ТЕХНОЛОГІЙ 122

В. И. Дубровин, Л. Ю. Дейнега
ПОДСИСТЕМА ВИЗУАЛИЗАЦИИ БАЗЫ ЗНАНИЙ
ЭКСПЕРТНОЙ СИСТЕМЫ ТЕХНИЧЕСКОЙ
ДИАГНОСТИКИ 117

М. П. Мусиенко, В. И. Томенко
БЕСПРОВОДНЫЕ ИНФОРМАЦИОННО-
КОММУНИКАЦИОННЫЕ
И НАВИГАЦИОННЫЕ ТЕХНОЛОГИИ
НА ТРАНСПОРТНЫХ СРЕДСТВАХ 127

ТЕОРІЯ І МЕТОДИ АВТОМАТИЧНОГО УПРАВЛІННЯ 132

В. И. Гостев, С. Н. Скуртов, И. В. Панченко
ОПРЕДЕЛЕНИЕ УПРАВЛЯЮЩИХ
ВОЗДЕЙСТВИЙ НА ВЫХОДЕ НЕЧЕТКОГО
РЕГУЛЯТОРА ПРИ ИДЕНТИЧНЫХ
ТРЕУГОЛЬНЫХ ФУНКЦИЯХ
ПРИНАДЛЕЖНОСТИ С ОГРАНИЧЕНИЕМ 132

*Е. М. Потапенко, А. Е. Казурова, Е. В. Душинова,
В. И. Левыкина*
АСИМПТОТИЧЕСКОЕ ВЫПОЛНЕНИЕ
МАТЕМАТИЧЕСКИХ ОПЕРАЦИЙ НА
МНОГОФАЗНЫХ СИГНАЛАХ 144

Т. Б. Никитина
СИНТЕЗ ЦИФРОВОГО НЕЛИНЕЙНОГО
РОБАСТНОГО УПРАВЛЕНИЯ
МНОГОКАНАЛЬНЫМИ СИСТЕМАМИ 134

КЕРУВАННЯ У ТЕХНІЧНИХ СИСТЕМАХ 148

Е. М. Потапенко, А. В. Соломаха
СИНТЕЗ И АНАЛИЗ КОМПЕНСАТОРА
ПАССИВНОГО LC-ФИЛЬТРА НА ВЫХОДЕ
ПРЕОБРАЗОВАТЕЛЯ ЧАСТОТЫ 148

В. Е. Цыганиш, В. М. Зверев, Ю. С. Белоуваненко
АНАЛИЗ ЗАДАЧИ ОПТИМАЛЬНОГО
УПРАВЛЕНИЯ ПРОЦЕССОМ ПРЕОБРАЗОВАНИЯ
ЭНЕРГИИ В ТУРБОАГРЕГАТЕ 151

АВТОРЫ НОМЕРА 155

РАДІОЕЛЕКТРОНІКА ТА ТЕЛЕКОМУНІКАЦІЇ

РАДИОЕЛЕКТРОНИКА И ТЕЛЕКОММУНИКАЦИИ

RADIO ELECTRONICS AND TELECOMMUNICATIONS

УДК 621.396.6.017:004.942

Н. П. Гапоненко, Е. В. Огренич

ПРОЕКТИРОВАНИЕ ГОФРИРОВАННЫХ ТЕПЛОТВОДОВ

Исследуются процессы теплоотдачи в гофрированных теплоотводах. Получены зависимости коэффициентов теплоотдачи конвекцией и излучением от угла гофрирования. Показана возможность оптимизации массогабаритных показателей. Предложен алгоритм проектирования.

ВВЕДЕНИЕ

Тепловой режим теплонагруженных элементов в радиоэлектронной аппаратуре в большинстве случаев обеспечивается применением теплоотводящих устройств или вынужденным охлаждением. Оба эти пути влияют на массогабаритные показатели радиоаппаратуры, требуют ресурсных затрат и оптимизации конструкций теплоотводящих устройств.

Одним из наиболее распространенных теплоотводящих устройств является радиатор. Существует большое количество конструкций радиаторов [1, 2], что не позволяет в общем виде решить задачу оптимизации их массогабаритных показателей. Исследованы только простые элементы конструкций [3], для которых показана возможность уменьшения массы в несколько раз. Оптимизация таких элементов при использовании комплексных массогабаритных критериев рассмотрена в [4]. В большинстве случаев это значительно увеличивает

площадь, которую занимают теплоотводящие элементы на печатных платах.

Уменьшение площади может осуществляться переходом к пространственным формам теплоотводящих элементов, например с гофрированными поверхностями. Расчет тепловых процессов в таких теплоотводах может проводиться средствами инженерного анализа [5, 6], которые позволяют учитывать нелинейные свойства теплоотдачи конвекцией и излучением. К сожалению, решение задач теплопереноса этими средствами требует значительных временных ресурсов, что затрудняет их использование в алгоритмах оптимизации.

Целью работы является сокращение времени проектирования гофрированных теплоотводов с помощью средств инженерного анализа. Для решения этой задачи, в первую очередь, необходимо значительно уменьшить время анализа тепловых режимов программными средствами. Такая возможность представляется, если известны аналитические соотношения для коэффициентов теплоотдачи. Эти соотношения могут быть получены обработкой результатов моделирования тепловых процессов компьютерными средствами инженерного анализа, которые позволяют отказаться от трудоемких и дорогостоящих экспериментальных исследований.

1 ИССЛЕДОВАНИЕ КОЭФФИЦИЕНТОВ ТЕПЛОТДАЧИ

При переходе от плоских форм к гофрированным поверхностям можно обеспечить различные формы гофр. Наиболее простыми из них являются поверхности с прямоугольными и треугольными гофрами (рис. 1), где 1 – источник тепла; 2 – поверхность теплоотвода.

Теплоотводы с прямоугольными гофрами (рис. 1, а) могут применяться при значительных расстояниях между теплоотводящей поверхностью и другими элементами в электронных аппаратах, когда отсутствует ограничение на межплатное расстояние. При наличии таких ограничений на высоту радиатора более эффективно использование теплоотводящих поверхностей с треугольным профилем (рис. 1, б).

Исследование влияния параметров теплоотвода на коэффициенты теплоотдачи проводилось программными средствами [5] для гофрированных теплоотводов с треугольным профилем. Основными параметрами таких конструкций являются ширина пластины B и толщина d пластины, высота гофр h , угол гофрирования γ . Эти параметры связаны с шагом гофрирования и длиной развертки соотношениями:

$$\gamma = \arctg \frac{s}{2h}; \quad l = 2h \cos \frac{\gamma}{2},$$

где s – шаг гофрирования; l – длина развертки одного гофра.

В качестве модели использовалась одна секция гофр с толщиной $d = 1$ мм, шириной $B = 40$ мм. Через внешние поверхности теплоотвод отсутствовал, поверхность гофр была изотермической. Угол гофрирования изменялся от $7,5^\circ$ до 180° . Температура поверхности гофр задавалась постоянной 60°C и 80°C , температура окружающей среды составляла $t_{\text{ат}} = 40^\circ\text{C}$. Движение

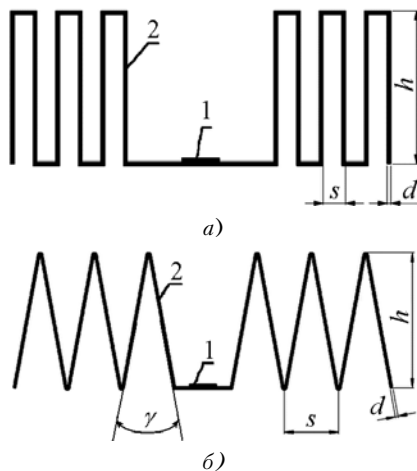


Рисунок 1 – Гофрированные теплоотводы с прямоугольными и треугольными гофрами

теплоносителя осуществлялось в каналах, формируемых гофрами, что соответствует вертикальному расположению теплоотвода.

Зависимости коэффициента теплоотдачи конвекцией при естественном охлаждении и длине развертки $l = 120$ мм имеют вид, показанный на рис. 2, где α_c – коэффициент теплоотдачи конвекцией, t_g – температура гофрированной поверхности.

Полученные зависимости нормировались коэффициентами теплоотдачи теплоотвода при $\gamma = 180^\circ$. В результате аппроксимации получено выражение для коэффициента теплоотдачи конвекцией:

$$\alpha_c = \alpha_{c0} \cdot \left[1 + \frac{1}{l} \cdot \left(1,68 \cdot 10^{-3} + \frac{0,284}{\gamma} \right) \right], \quad (1)$$

где α_{c0} – коэффициент теплоотдачи конвекцией при $\gamma = 180^\circ\text{C}$.

Для расчета α_{c0} использовались известные соотношения [1]. Выражение (1) обеспечивает точность расчета коэффициента теплоотдачи не хуже $0,5\%$.

Зависимость коэффициента теплоотдачи излучением α_r от угла гофр представлена на рис. 3. Полученные зависимости нормировались относительно угла $\gamma = 180^\circ$ и аппроксимировались выражениями:

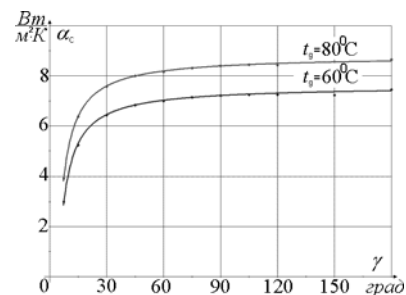


Рисунок 2 – Зависимость коэффициента теплоотдачи конвекцией от угла гофра

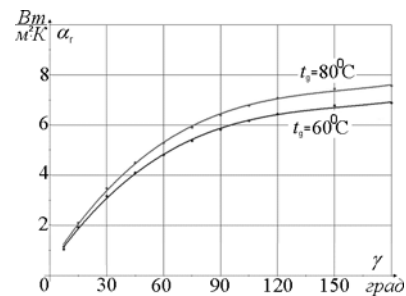


Рисунок 3 – Зависимость коэффициента теплоотдачи излучением от угла гофра

$$\alpha_r = \alpha_{r0} \cdot \left[\frac{1-\varepsilon}{\varepsilon} + \left[\sin^{-1} \frac{\gamma}{2} - \left(1 - \sqrt{1 + \left(\frac{2B}{l} \right)^2} - \frac{2B}{l} \right)^{-1} \right] \cdot \left(1 - \sqrt{1 + \left(\frac{2B}{l} \right)^2} - \frac{2B}{l} \right)^{-1} \right]^{-1}, \quad (2)$$

где α_{r0} – коэффициент теплоотдачи излучением при $\gamma = 180^\circ$; ε – коэффициент черноты поверхности теплоотвода.

Для расчета α_{r0} использовались известные соотношения [1]. Выражение (2) обеспечивает точность расчета α_r не хуже 2 %. При необходимости увеличения точности расчет радиационной составляющей теплоотвода может проводиться инженерными средствами проектирования, в которых эта составляющая рассчитывается значительно быстрее, чем конвективная. Соотношения (1) и (2) позволяют значительно сократить время вычислений и использовать для оптимизации гофрированных теплоотводов обычные компьютерные средства.

2 ПРОЕКТИРОВАНИЕ ГОФРИРОВАННОГО ТЕПЛОТВОДА

Тепловая модель гофрированного теплоотвода может быть представлена в виде ребра с различными коэффициентами теплоотдачи в центре и на краях теплоотвода (рис. 4), где P – мощность источника тепла, α_{ext} , α_{int} – коэффициенты теплоотдачи на краях и в центре теплоотвода, L – длина развертки теплоотвода. Коэффициент α_{int} определяется с помощью соотношений (1) и (2). Коэффициент α_{ext} определяется как сумма коэффициентов теплоотдачи с гофрированной поверхности и поверхности в неограниченном пространстве.

Тепловая модель, изображенная на рис. 4, пересчитывается к однородной модели при усредненном коэффициенте теплоотдачи:

$$\alpha = \frac{2}{n} \alpha_{ext} + \frac{n-2}{n} \alpha_{int}, \quad (3)$$

где n – количество граней радиатора.

В этом случае распределение температуры в теплоотводе имеет вид:

$$\theta(x) = \frac{P}{\lambda b S} \cdot \frac{\text{ch}[b(L-x)]}{\text{sh}(bL)}; \quad b^2 = \frac{\alpha U}{\lambda S}, \quad (4)$$

где $\theta(x)$ – перепад температуры между теплоотводом и окружающей средой; λ – коэффициент теплопроводности материала; $S = dB$ – площадь поперечного сечения теплоотвода; $U = 2(d+B)$ – периметр поперечного сечения теплоотвода.

Итерационный алгоритм проектирования теплоотвода с минимальными массогабаритными показателями (MS-стратегия), учитывающий нелинейную зависимость коэффициентов теплоотдачи, имеет вид:

Шаг 1. Уточняется значение угла γ ;

Шаг 2. Уточняется значение толщины теплоотвода $d^{(i)}$;

Шаг 3. Уточняется L из условия обеспечения теплового режима:

$$L = \frac{1}{b} \text{acth} \left(\theta_m \frac{\lambda b S}{P} \right),$$

где $\theta_m = \theta(0)$;

Шаг 4. Рассчитывается критерий $m^{(i)} \cdot S^{(i)}$;

Шаг 5. Проверяется условие:

$$m^{(i)} \cdot S^{(i)} \rightarrow \min. \quad (5)$$

Если условие (5) не выполняется, то переход на шаг 2;

Шаг 6. Проверяется условие:

$$\left| \frac{d^{(i)} - d^{(i-1)}}{d^{(i-1)}} \right| \leq \varepsilon, \quad (6)$$

где ε – точность вычислений.

Если условие (6) не выполняется, то переход на шаг 1. Иначе – конец алгоритма.

Методика применялась для проектирования алюминиевого теплоотвода при $B = 40$ мм. Мощность источника тепла задавалась 2 Вт. Полученные зависимости приведены на рис. 5.

Из рис. 5 видно, что применение гофрированных поверхностей позволяет уменьшить MS-критерий теплоотводов примерно в 2 раза. При этом длина теплоотвода уменьшается в 6 раз по сравнению с первоначальной. Масса теплоотвода увеличилась всего в 1,7 раза.

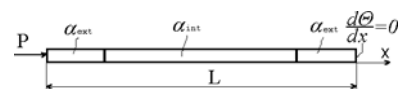


Рисунок 4 – Тепловая модель

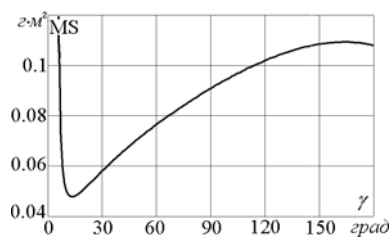


Рисунок 5 – Зависимость MS-критерия от угла гофр γ

ВЫВОДЫ

Анализ процессов теплоотдачи в гофрированных теплоотводах показал, что зависимости коэффициентов теплоотдачи от угла гофрирования легко нормируются относительно соответствующих значений для плоских поверхностей. Результаты аппроксимации представлены простыми выражениями, что позволяет использовать их в алгоритмах оптимизации массогабаритных параметров без применения инженерных средств проектирования.

Применение процедур оптимизации показало, что в гофрированных теплоотводах возможно уменьшение массогабаритных показателей более чем в 2 раза. При этом площадь, занимаемая теплоотводом на плате, уменьшается почти в 6 раз. Изменение массы обратно пропорционально изменению MS-критерия.

ПЕРЕЧЕНЬ ССЫЛОК

1. Дульнев Р. Н. Тепло- и массообмен в радиоэлектронной аппаратуре. – М.: Высш. шк., 1984. – 247 с.

УДК 621.372.011.72

С. П. Гулин

АНАЛИЗ СПЕКТРА ОТКЛИКА НЕЛИНЕЙНОСТИ, ПРЕДСТАВЛЕННОЙ ФУНКЦИЕЙ ДИНАМИЧЕСКОГО НАСЫЩЕНИЯ, ПРИ МНОГОЧАСТОТНОМ ВОЗДЕЙСТВИИ

Предложен метод анализа спектра установившегося отклика нелинейности, представленной аналитической трансцендентной функцией, при многочастотном воздействии на основе гипергеометрической функции Гаусса. Полученные результаты позволяют моделировать поведение широкого класса электронных устройств и компонентов в режимах малого и большого сигналов с произвольным спектром.

ПОСТАНОВКА ЗАДАЧИ

При анализе электронных систем в целом часто возникает необходимость исследовать работу отдельных функциональных узлов, работающих в нелинейных режимах при многочастотном воздействии в широком диапазоне изменения его нормы.

Для моделирования режимов работы указанных объектов исследования применяются математические модели с использованием функций, которые позволяют решать поставленную задачу с различной степенью

2. Роткоп Л. Л., Спокойный Ю. Е. Обеспечение тепловых режимов при конструировании радиоэлектронной аппаратуры. – М.: Советское радио, 1976. – 232 с.
3. Ройзен Л. И., Дулькин И. Н. Теловой расчет оребренных поверхностей. под. ред. В. Г. Фастовского. – М.: Энергия, 1977. – 256 с.
4. Gaponenko N., Ogrenich E. Strategy of flanged radiators design // Proceedings of the International Conference TCSET'2006. – P. 554–556.
5. Алямовский А. А. SolidWorks. Компьютерное моделирование в инженерной практике. – СПб.: БХВ-Петербург, 2005. – 800 с.
6. Каплун А. Б., Морозов Е. М., Олферьева М. А. ANSYS в руках инженера: Практическое руководство. – М.: Едиториал УРСС, 2003. – 272 с.

Надійшла 12.09.07
Після доробки 15.10.07

Досліджуються процеси тепловіддачі в гофрованих тепловідводах. Отримані залежності коефіцієнтів тепловіддачі конвекцією та випромінюванням в залежності від кута гофрування. Показана можливість оптимізації массогабаритних показників. Запропонован алгоритм проектування.

Heat transfer in goffered heat sink are studied. Dependencies of heat transfer coefficients on goffered angle are obtained. The possibility of optimizing mass and size characteristics is shown. A new design algorithm is proposed.

точности, однако не обладают достаточной степенью гибкости для изменения их формы. В последние годы при решении подобных задач наметилось новое направление, связанное с концепцией «управляемого динамического насыщения» (КУДН) [1–5]. В одной из работ указанного направления А. Д. Канном [1] была предложена модель динамического насыщения безынерционного усилителя, которая описывается функцией вида

$$Y(X) = A \cdot \operatorname{sgn}(X) \cdot [1 + (B/|X|)^S]^{-1/S}, \quad (1)$$

где X и Y – входное и выходное напряжения, а B и A – их уровни насыщения; S – параметр, регулирующий кривизну годографа (1).

Необходимо отметить, что модель А. Д. Канны (1) не является пионерской в использовании и развитии принципа «динамического насыщения». Достаточно

© Гулин С. П., 2007

сослаться на работу Л. К. Регенбогена [2], в которой была представлена простая модель безынерционного нелинейного устройства с использованием принципа управляемого насыщения

$$Y(X) = A \cdot [1 + (1/X)^S]^{-1/S}. \quad (2)$$

Однако, как показано в работе С. Л. Лойка [3], модели вида (1–2) имеют существенные недостатки, обусловленные существованием точек сингулярности, которые не позволяют адекватно моделировать спектры отклика и нелинейные эффекты в устройствах при многочастотном воздействии. Ни в одной из известных автору публикаций, включая [1–5], не были исследованы причины проявления отмеченных недостатков и не были предложены пути их преодоления. Эту задачу удалось решить в работах [6] и [7], позволивших предложить функцию

$$Y(X) = Y_0 + A \cdot [1 + (B/X)^p]^{-1/S}, \quad (3)$$

которая в рамках КУДН обобщает модели (1–2) и обеспечивает большую гибкость в изменении формы ее годографа и, самое главное, устраняет недостатки, характерные для моделей А. Д. Канна и Л. Н. Регенбогена.

Автором в [7] на основе функции (3) были также предложены несколько ее модификаций, обеспечивающих моделирование передаточных характеристик широкого класса безынерционных нелинейных элементов (БНЭ) и устройств, не проявляющих свойств конвергентности [8], на основе КУДН, например:

$$Y_1(X) = Y_0 + A + \operatorname{sgn}(X) \cdot [1 + (B/|X|)^p]^{-1/S}, \quad (3a)$$

либо

$$Y_2(X) = Y_0 \cdot |X| \cdot A \cdot \operatorname{sgn}(X) \cdot [1 + (B/|X|)^p]^{-1/S}. \quad (3b)$$

Смысл параметров A и B в (3) тот же, что и в (1); p , S – параметры, регулирующие кривизну годографа функции (3) в пределах квадранта с учетом условий [7], причем, $p \neq S$, $S > 1$, $p > 0$; $Y_0 = Y(0)$. При этом параметры A и Y_0 определяются из ВАХ, амплитудной или передаточной характеристики моделируемого безынерционного нелинейного устройства.

Кроме того, автором в [6] был предложен метод анализа спектра установившегося отклика БНЭ, передаточная характеристика которого описывается функцией динамического насыщения (ФДН) (3) на воздействие вида:

$$\begin{aligned} X_{\text{вх}}(t) &= X_0 + \sum_{i=1}^N X_i \cdot \cos(\omega_i t + \varphi_i) = \\ &= X_N \cdot \left[1 + \sum_{i=0}^N \frac{X_i}{X_N} \cdot \cos(\omega_i t + \varphi_i) \right], \end{aligned} \quad (4)$$

где $X_0 = \text{const}$; X_i , ω_i , φ_i – амплитуда, круговая частота и начальная фаза i -й компоненты (4), соответственно; $X_0 = X_0 - X_N$; X_N – норма, определяемая соотношением $X_N = \sqrt{\sum_{i=0}^N |X_i|^2}$.

При этом предполагается, что частоты ω_i – несоизмеримы, а их любые возможные линейные комбинации с целыми коэффициентами являются числами, не равными друг другу.

Однако алгоритм спектрального анализа отклика нелинейности (3) [6] оказался громоздким, поскольку включал операции обращения степенных рядов. Цель настоящей работы состоит в разработке компактного алгоритма анализа спектра установившегося отклика исследуемой нелинейности, свободного от указанного недостатка.

РЕШЕНИЕ

План решения сформулированной задачи заключается в том, чтобы, используя аналитическую связь (3), с учетом (4), найти компоненты отклика заданной комбинационной частоты $\omega_\Sigma = \sum_{i=1}^N n_i \cdot \omega_i$ для различных порядков нелинейности, суммируя которые, с учетом их фазовых соотношений, получить искомый результат.

С этой целью представим (3) с учетом (4) в следующем виде:

$$Y(t) = Y_0 + A \cdot (1 + a_0^{-1} \cdot \tilde{x}^{-p})^{-1/S}, \quad (5)$$

где

$$a_0 = (X_N/B)^p; \quad (6a)$$

$$\begin{aligned} \tilde{x} &= \left[1 + \sum_{i=0}^N \frac{X_i}{X_N} \cdot \cos(\omega_i \cdot t + \varphi_i) \right] = 1 + z; \\ z &= \sum_{i=0}^N \frac{X_i}{X_N} \cdot \cos(\omega_i \cdot t + \varphi_i). \end{aligned} \quad (6b)$$

Так как поведение функции (3) при вариации нормы X_N уже изучено в [6], рассмотрим выражение в круглых скобках (5):

$$(1 + a_0^{-1} \cdot \tilde{x}^{-p})^{-1/S} = [1 + a_0^{-1} \cdot (1 + z)^{-p}]^{-\frac{1}{S}}. \quad (7)$$

Поскольку параметр p – произвольное действительное число, то результат возведения в степень $(-p)$ бинома $(1 + z)$ может быть представлен выражением [9]

$$(1 + z)^{-p} = {}_2F_1(p; b; b; -z), \quad (8)$$

где ${}_2F_1(p; b; b; -z)$, $= \sum_{k=0}^{\infty} \frac{(p)_k \cdot (-1)^k}{k!} \cdot z^k$ – гипергеометрическая функция Гаусса; $(p)_k = p \cdot \dots \cdot (p - k + 1) =$

$= \frac{\Gamma(p+k)}{\Gamma(p)}$ – факториальная функция, $z > 0$; $(p)_0 = 1$;

$\Gamma(p) = \int_0^{\infty} \exp(-t) \cdot t^{p-1} dt$ – гамма-функция Эйлера ($p > 0$) [10].

Для дальнейших преобразований без уменьшения общности ограничимся рассмотрением ФДН (3) в первом квадранте декартовой системы координат при условии $Y_0 = Y(0) = 0$.

Вводя параметр $\gamma = -1/S$ и используя (8), представим (5) в виде

$$Y(t) = A \cdot (1 + a_0^{-1})^\gamma \cdot \lim_{m \rightarrow \infty} \left[1 + \sum_{k=1}^m \frac{(p)_k \cdot (-1)^k}{(1 + a_0) \cdot k!} \cdot z^k \right]^\gamma. \quad (9)$$

Для возведения в произвольную действительную степень содержимого квадратной скобки (9), согласно [10], применяем формулу полиномиального разложения

$$\left(1 + \sum_{l=1}^L U_l \right)^\gamma = \sum_{k_1=0}^{\infty} \sum_{k_2=0}^{\infty} \dots \sum_{k_L=0}^{\infty} \gamma \cdot (\gamma - 1) \cdot \dots \times \left(\gamma - \sum_{l=1}^L k_l + 1 \right) \cdot \prod_{l=1}^L \frac{U_l^{k_l}}{k_l!}. \quad (10)$$

С учетом (10), соотношение (9) принимает вид

$$Y(t) = A_0 \cdot \lim_{m \rightarrow \infty} \underbrace{\sum_{k_1=0}^{[\theta(m)]} \sum_{k_2=0}^m \dots \sum_{k_m=0}^m \gamma \cdot (\gamma - 1) \cdot \dots \left(\gamma - \sum_{l=1}^m k_l + 1 \right)}_{\sum_{l=1}^m k_l = m; \sum_{l=1}^m l \cdot k_l = m; k_l \geq 0, l = \overline{1, m}} \cdot \prod_{l=1}^m \frac{(p)_{k_l} \cdot (-1)^{k_l} \cdot z^{l \cdot k_l}}{(1 + a_0)^\gamma \cdot k_l!}, \quad (11)$$

где $A_0 = A \cdot (1 + a_0^{-1})^\gamma$; $[\theta(m)]$ – целая часть числа возможных разбиений целого положительного числа m на m целых неотрицательных чисел-частей $k_l, l = \overline{1, m}$.

Решение поставленной задачи включает определение числа всевозможных разбиений $\theta(m)$ целого положительного числа m , равного значению верхнего индекса суммирования в (11), на m целых неотрицательных чисел-частей m_l , которые, согласно [11], удовлетворяют уравнение Диофанта

$$k_1 + 2 \cdot k_2 + 3 \cdot k_3 + \dots + m \cdot k_m = m, \quad (12)$$

при одновременном выполнении дополнительного условия

$$m = \sum_{l=1}^m k_l, \quad m \rightarrow \infty. \quad (13)$$

Харди и Рамануджаном [11] была получена приближенная формула, определяющая число неограниченных разбиений $\theta(m)$ числа m :

$$\theta(m) = (4 \cdot \sqrt{3} \cdot m)^{-1} \cdot \exp(\pi \cdot \sqrt{2m/3}) \times [1 + O(m^{-1/4 + \varepsilon})], \quad (14)$$

где ε удовлетворяет неравенству $0 < \varepsilon < 1/4$.

Поскольку потребности практики при анализе спектра отклика современных узкополосных нелинейных систем и устройств телекоммуникации и связи на многочастотное воздействие пока удовлетворяются значением $m \leq 10 \dots 20$ [3–5], то для такого диапазона значений m целесообразно использовать рекуррентное соотношение одного из следствий пентагональной теоремы Л. Эйлера [12]:

$$\theta(m) - \theta(m-1) + \theta(m-2) - \theta(m-5) - \theta(m-7) + \dots + (-1)^m \cdot \{\theta[m - m \cdot (3m-1)/2] + \theta[m - m \cdot (3m+1)/2]\} + \dots = 0, \quad (15)$$

где $\theta(m) = 0$ для всех отрицательных m , а $\theta(0) = 1$.

Учитывая (6б), производим последующие преобразования (11)

$$Y(t) = A_0 \cdot \lim_{m \rightarrow \infty} \sum_{k=1}^{[\theta(m)]} \underbrace{\sum_{k_1=0}^m \sum_{k_2=0}^m \dots \sum_{k_m=0}^m \gamma \cdot (\gamma-1) \cdot \dots \cdot (\gamma-m+1)}_{\sum_{l=1}^m k_l=m; \sum_{l=1}^m l \cdot k_l=m; k_l \geq 0, l=\overline{1,m}} \cdot \left[\prod_{l=1}^m \frac{(p)_{k_l} \cdot (-1)^{k_l}}{(1+a_0)^{\gamma} \cdot k_l!} \right] \times \sum_{d=0}^m \binom{m}{d} \cdot \left[\sum_{i=1}^N \frac{X_i}{X_0 - X_N} \cdot \cos(\omega_i t + \phi_i) \right]^d$$

Используя формулу Эйлера для косинуса и вводя комплексные, комплексно-сопряженные и нормированные амплитуды для i -й компоненты спектра воздействия (4): $\dot{\bar{X}}_i = \bar{X}_i \cdot e^{j\omega_i t}$, $\ddot{\bar{X}}_i = \bar{X}_i \cdot e^{-j\omega_i t}$, $\bar{X}_i = X_i / (X_0 - X_N)$, $X_{0N} = (X_0 - X_N) / X_N$, приводим (16) к виду

$$Y(t) = A_0 \cdot \lim_{m \rightarrow \infty} \sum_{k=1}^{[\theta(m)]} \underbrace{\sum_{k_1=0}^m \sum_{k_2=0}^m \dots \sum_{k_m=0}^m \gamma \cdot (\gamma-1) \cdot \dots \cdot (\gamma-m+1)}_{\sum_{l=1}^m k_l=m; \sum_{l=1}^m l \cdot k_l=m; k_l \geq 0, l=\overline{1,m}} \cdot \left[\prod_{l=1}^m \frac{(p)_{k_l} \cdot (-1)^{k_l}}{(1+a_0)^{\gamma} \cdot k_l!} \right] \times X_{0N} \cdot \sum_{d=0}^m \sum_{s=0}^d \frac{m! \binom{d}{s} \cdot 2^{-d}}{d! \cdot (m-d)!} \cdot \left[\sum_{i=1}^N \dot{\bar{X}}_i \cdot e^{j\omega_i t} \right]^s \cdot \left[\sum_{i=1}^N \ddot{\bar{X}}_i \cdot e^{-j\omega_i t} \right]^{\gamma-s} \quad (17)$$

Мультиномиальные разложения в квадратных скобках (17), определяются формулами [10]:

$$\left[\sum_{i=1}^N \dot{\bar{X}}_i \cdot e^{j\omega_i t} \right]^s = \sum_{q_1=0}^s \sum_{q_2=0}^s \dots \sum_{q_N=0}^s s! \times \prod_{i=1}^N \frac{\dot{\bar{X}}_i^{q_i}}{q_i!} \cdot e^{j \cdot \left(\sum_{i=1}^N q_i \right) \cdot \omega_i t}, \quad (18)$$

$$\left[\sum_{i=1}^N \ddot{\bar{X}}_i \cdot e^{-j\omega_i t} \right]^{\gamma-s} = \sum_{r_1=0}^{\gamma-s} \sum_{r_2=0}^{\gamma-s} \dots \sum_{r_N=0}^{\gamma-s} (\gamma-s)! \times \prod_{i=1}^N \frac{\ddot{\bar{X}}_i^{r_i}}{r_i!} \cdot e^{-j \cdot \left(\sum_{i=1}^N r_i \right) \cdot \omega_i t}, \quad (19)$$

где $\{q_i\}$ и $\{r_i\} (i = \overline{1, N})$ – множество всевозможных разбиений целых чисел s и $(\gamma-s)$ на N неотрицательных чисел-частей, соответственно.

Если определить g_i как меньшее из чисел q_i и r_i , то большее из них определится суммой $g_i + |n_i|$. Огра-

ничиваясь положительной частотной полуосью, введем следующие обозначения:

$$n = \sum_{i=1}^N |n_i|, \quad \beta = \sum_{i=1}^N g_i. \quad (20)$$

Причем круговым частотам компонент с комплексными амплитудами $\dot{\bar{X}}_i$ (16) соответствует неравенство $n_i \geq 0$, а круговым частотам компонент с комплексно-сопряженными амплитудами $\ddot{\bar{X}}_i$ – неравенство $n_i < 0$.

Учитывая (19), результаты работы [6] и, применяя известные факториальные соотношения [10], получаем

$$(g_i + |n_i|)! = |n_i|! \cdot (|n_i| + 1)_{g_i}, \quad (21)$$

$$\gamma \cdot (\gamma-1) \cdot \dots \cdot (\gamma-n-2 \cdot \beta+1) = (-1)^n \cdot 2^{2\beta} \cdot (-\gamma)_n \cdot \left(\frac{n-\gamma}{2} \right)_{\beta} \cdot \left(\frac{n-\gamma+1}{2} \right)_{\beta}, \quad (22)$$

где для целых k и m $(k)_m = k \cdot (k+1) \cdot \dots \cdot (k+m-1)$ – символ Похгаммера, а $(k)_m = 1$ для $m \leq 0$.

Учитывая (14–21), получаем конечную формулу отклика $Y(t)$:

$$Y(t) = A_0 \cdot \lim_{m \rightarrow \infty} \underbrace{\sum_{k=1}^{\lfloor \theta(m) \rfloor} \sum_{k_1=0}^m \dots \sum_{k_m=0}^m (-1)^n \cdot 2^{2\beta} \cdot \left(\frac{1}{S}\right)_n}_{\sum_{l=1}^m k_l=m; \sum_{l=1}^m l \cdot k_l=m; k_l \geq 0, l=1, \overline{m}} \cdot \left(\frac{n + \frac{1}{S}}{2}\right)_\beta \cdot \left(\frac{n + \frac{1}{S} + 1}{2}\right)_\beta \cdot \left[\prod_{l=1}^m \frac{(p)_{k_l} \cdot (-1)^{k_l}}{(1+a_0)^{k_l} \cdot k_l!} \right] \times \\ \times X_{0N}^m \cdot \sum_{d=0}^m \frac{m! \cdot 2^{-j}}{(m-j)!} \cdot \sum_{s=0}^j \sum_{q_1=0}^s \dots \sum_{q_N=0}^s \sum_{r_1=0}^{\gamma-s} \sum_{r_2=0}^{\gamma-s} \dots \sum_{r_N=0}^{\gamma-s} \prod_{i=1}^N \frac{\ddot{X}_i^{g_i} \cdot \dot{X}_i^{g_i+|n_i|}}{g_i! \cdot |n_i|! \cdot (|n_i|+1)_{g_i}} e^{j\omega_0 t}. \quad (23)$$

Используя (23), после соответствующих преобразований получаем формулы, определяющие амплитуду комбинационного колебания V_Σ частотой ω_Σ и величину постоянной составляющей отклика V_0

$$\frac{1}{2} \cdot V_\Sigma = A_0 \cdot \lim_{m \rightarrow \infty} \underbrace{\sum_{k=1}^{\lfloor \theta(m) \rfloor} \sum_{k_1=1}^m \dots \sum_{k_m=0}^m (-1)^n \cdot 2^{2\beta} \cdot \left(\frac{1}{S}\right)_n}_{\sum_{l=1}^m k_l=m; \sum_{l=1}^m l \cdot k_l=m; k_l \geq 0, l=1, \overline{m}} \cdot \left(\frac{n + \frac{1}{S}}{2}\right)_\beta \cdot \left(\frac{n + \frac{1}{S} + 1}{2}\right)_\beta \cdot \left[\prod_{l=1}^m \frac{(p)_{k_l} \cdot (-1)^{k_l}}{(1+a_0)^{k_l} \cdot k_l!} \right] \times \\ \times X_{0N}^m \cdot \sum_{d=0}^m \sum_{s=0}^d \frac{m! \cdot 2^{1-d}}{(m-d)!} \cdot \sum_{q_1=0}^s \dots \sum_{q_N=0}^s \sum_{r_1=0}^{\gamma-s} \sum_{r_2=0}^{\gamma-s} \dots \sum_{r_N=0}^{\gamma-s} \prod_{i=1}^N \frac{\ddot{X}_i^{g_i} \cdot \dot{X}_i^{g_i+|n_i|}}{g_i! \cdot |n_i|! \cdot (|n_i|+1)_{g_i}}. \quad (24)$$

$$V = A \cdot \lim_{m \rightarrow \infty} \underbrace{\sum_{k=1}^{\lfloor \frac{\theta(m)}{2} \rfloor} \sum_{k_1=0}^m \dots \sum_{k_m=0}^m 2^{2\beta} \cdot \frac{1}{S} \cdot \left(\frac{1}{S}\right)_\beta \cdot \left(\frac{1}{S} + 1\right)_\beta}_{\sum_{l=1}^m k_l=m; \sum_{l=1}^m l \cdot k_l=m; k_l \geq 0, l=1, \overline{m}} \cdot \left[\prod_{l=1}^m \frac{(p)_{k_l} \cdot (-1)^{k_l}}{(1+a_0)^{k_l} \cdot k_l!} \right] \times \\ \times X_{0N}^m \cdot \sum_{d=0}^m \sum_{s=0}^d \frac{m! \cdot 2^{-d}}{(m-d)!} \cdot \sum_{q_1=0}^s \dots \sum_{q_N=0}^s \sum_{r_1=0}^{\gamma-s} \sum_{r_2=0}^{\gamma-s} \dots \sum_{r_N=0}^{\gamma-s} \prod_{i=1}^N \frac{|\dot{X}_i|^{2g_i}}{(g_i!)^2}. \quad (25)$$

Результаты практических расчетов по соотношениям (23–25) настоящей работы полностью совпадают с результатами расчетов, использующих аналогичные соотношения работы [6], однако требуют значительно меньших затрат машинного времени.

ВЫВОДЫ

Практическая ценность работы состоит в том, что на основе гипергеометрической функции Гаусса предложен новый, свободный от недостатков метода [6] алгоритм расчета спектра установившегося отклика произвольной аналитической безынерционной нелинейности на многочастотное полиамплитудное воздействие. Предложенный алгоритм позволяет производить корректную оценку уровня нелинейных искажений отклика нелинейных безынерционных устройств, работающих с реальными сигналами, с помощью известных критериев совместно с ранее разработанной методикой

[13] построения их математических моделей на основе КУДН.

Работа выполнена в рамках госбюджетной НИР № ДБ03916 «Дослідження алгоритмів обробки сигналів в умовах інтенсивної протидії з урахуванням неідеальності каналів передачі та приймання».

ПЕРЕЧЕНЬ ССЫЛОК

1. Cann A. J. Nonlinearity model with variable knee sharpness // IEEE Trans. Aerosp. Electron. Syst. – 1980. – Vol. AES-16, № 11. – P. 874–877.
2. Regenbogen L. K. Nonlinearity Model With Variable Knee Sharpness // IEEE Transaction Aerospace Electronic Systems. – 1980. – Vol. AES-16, May. – P. 410–414.
3. Louka S. L. On the Use of Cann's Model for Nonlinear Behavioral-Level Simulation // IEEE Trans. on Vehicular Technology. – 2000. – Vol. 49, № 5, September. – P. 1982–1985.
4. Бобков А. М., Яковлев Н. Н. Аппроксимация характеристики нелинейного безынерционного элемента // Радиотехника. – 1986. – № 5. – С. 25–26.
5. Верлань А. Ф., Горшко И. О., Гушель Т. П. Аппроксимация экспериментальных зависимостей полиномами

- с дробным показателем степени // Электронное моделирование. – 2002. – Т. 24, № 3. – С. 101–106.
6. Гулин С. П. Анализ спектра отклика нелинейности, представленной аналитической трансцендентной функцией, на многочастотное воздействие большой нормы // Радиоэлектроника. Информатика. Управління. – 2004. – № 1. – С. 21–28.
 7. Гулин С. П. Условия применимости модели динамического насыщения в задачах анализа спектра отклика нелинейных устройств // Радиоэлектроника. Информатика. Управління. – Запоріжжя, ЗНТУ, 2005. – № 2(14). – С. 21–28.
 8. Данилов Л. В., Матханов П. Н., Филиппов Е. С. Теория нелинейных электрических цепей. – Л.: Энергоатомиздат. Ленингр. отд-ние, 1990. – 256 с.: ил.
 9. Люк Ю. Специальные математические функции и их аппроксимации. Перевод с англ. – М.: Мир, 1980. – 608 с.
 10. Справочник по специальным функциям. Под ред. М. Абрамовица и И. Стигана. – М.: Наука, 1979. – 832 с.: ил.
 11. Сачков В. Н. Комбинаторные методы дискретной математики. – М.: Наука, 1977. – 320 с.
 12. Эндриус Г. Теория разбиений. Перевод с англ. – М.: Наука, 1982. – 256 с.
 13. Гулин С. П. Определение параметров адаптивной модели нелинейных компонентов, представленной аналитической трансцендентной функцией, на основе

экспериментальных характеристик // Радиоэлектроника. Информатика. Управління. – 2005. – № 2(13). – С. 25–32.

Надійшла 4.04.07

Після доробки 15.05.07

Запропоновано метод аналізу усталеного відгуку нелінійності, що представлена аналітичною трансцендентною функцією, при багаточастотному впливі на основі гіпергеометричної функції Гауса. Отримані результати дозволяють моделювати поведінку широкого класу електронних пристроїв в режимах малого та великого сигналів з довільним спектром.

The method of analysis of spectrum of the set response of nonlinearity is offered, by the represented analytical transcendent function, at multifrequency influence on the basis of hypergeometrical function of Gausse. The got results allow to design the conduct of wide class of electronic devices and components in the modes of small and large signals with an arbitrary spectrum.

УДК 539.1.074

А. А. Захарченко, В. Е. Кутний, И. М. Прохорец,
А. В. Рыбка, М. А. Хажмурадов

ОПРЕДЕЛЕНИЕ ПАРАМЕТРОВ ПЕРЕНОСА ЗАРЯДА В CdTe ДЕТЕКТОРАХ γ -ИЗЛУЧЕНИЯ МЕТОДАМИ МАТЕМАТИЧЕСКОГО МОДЕЛИРОВАНИЯ

Предложен метод быстрого определения параметров переноса заряда с использованием дозиметрических характеристики детектора и с применением математического моделирования.

ВВЕДЕНИЕ

Характеристики полупроводниковых детекторов, используемых для регистрации γ -излучения, существенно зависят от свойств переноса заряда: подвижности μ и времени жизни τ носителей заряда – электронов (e) и дырок (h). Качество детекторов характеризуется уровнем сбора неравновесного заряда, образующегося в полупроводнике под воздействием γ -излучения. Для количественного описания процесса сбора заряда обычно используется произведение $\mu\tau$ [1].

У получивших в последние годы широкое распространение детекторов на базе CdTe $(\mu\tau)_e$ обычно на порядок и более превышает $(\mu\tau)_h$ [1]. Такое различие параметров переноса приводит к неполному сбору за-

ряда, что сильно влияет на спектрометрические и дозиметрические характеристики детекторов [2, 3]. Для расчета параметров коррекции необходимо определить произведение $\mu\tau$.

Прямые измерения параметров переноса заряда [4] в CdTe затруднены из-за высокого сопротивления образцов, что приводит к ошибкам измерения сравнимым по величине со значениями определяемых параметров. Часто используемые методики, основанные на анализе отклика детекторов при облучении α -частицами [1] и низкоэнергетичным γ -излучением [5], применимы только для детекторов спектрометрического качества, у которых $(\mu\tau)_e$ превышает $1 \times 10^{-3} \text{ см}^2/\text{В}$.

В последние годы разработан ряд методик определения $(\mu\tau)_{e,h}$ в CdTe детекторах γ -излучения спектрометрического качества с помощью методов математического моделирования [6, 7]. Для детекторов дозиметрического качества эта задача остается нерешенной.

Цель данной работы заключается в разработке и верификации модели расчета параметров переноса заряда в CdTe детекторах на основе их дозиметрических характеристик (чувствительности). Актуальность рассматриваемой задачи обусловлена необходимостью развития новой элементной базы систем радиационного контроля в атомной энергетике, медицине и экологии.

1 АЛГОРИТМ ОПРЕДЕЛЕНИЯ $\mu\tau$ В CdTe ДЕТЕКТОРАХ

В работе [8] нами обоснована модель расчета амплитудных спектров и чувствительности планарных CdTe (CdZnTe) детекторов γ -излучения с использованием программы EGSnrc.

EGSnrc – это классический код Монте Карло, позволяющий изучение отдельных взаимодействий фотонов в объеме детектора. Разработанный нами программный код, встраиваемый в EGSnrc, позволяет проводить для каждого отдельного взаимодействия модификацию сбора заряда распределением носителей случайным гауссианом, чтобы воспроизводить эффекты генерации и уширения шумом электроники.

Для расчета параметров переноса заряда используется следующий алгоритм:

1. Экспериментально измеряются спектры при облучении CdTe детектора γ -квантами от источников ^{241}Am (60 кэВ) и ^{137}Cs (662 кэВ) с известной мощностью экспозиционной дозы (МЭД) излучения.

2. По спектру ^{241}Am рассчитывается положение фотопика, соответствующее линии 60 кэВ.

3. По спектру ^{137}Cs рассчитывается скорость счета детектора $N_{\text{сч}}$ с порогом дискриминации, соответствующем энергии 60 кэВ.

4. Рассчитывается чувствительность детектора как $N_{\text{сч}}/\text{МЭД}$.

5. С помощью программы моделирования рассчитываются положение фотопика 60 кэВ и чувствительность детектора при энергии излучения 662 кэВ на сетке значений $(\mu\tau)_{e,h}$.

6. В координатах $(\mu\tau)_h - (\mu\tau)_e$ строятся линии равных значений (изолинии) положения фотопика 60 кэВ и чувствительности детектора при энергии излучения 662 кэВ.

7. На построенной диаграмме точка пересечения линий, соответствующих экспериментально полученному положению фотопика 60 кэВ и чувствительности на линии 662 кэВ, определяет величины $(\mu\tau)_e$ и $(\mu\tau)_h$.

Расчеты показывают, что при уровнях эквивалентного шумового заряда (ENC) менее 4,5 кэВ положение фотопика 60 кэВ (^{241}Am) изменяется не более чем на 1 %. Чувствительность детектора на линии 662 кэВ (^{137}Cs) изменяется в диапазоне 1–2 %. Таким образом, влиянием шума до величины ENC ~4,5 кэВ можно пре-

небречь, что существенно уменьшает время моделирования.

2 РАСЧЕТ И СРАВНЕНИЕ С ЭКСПЕРИМЕНТОМ

Верификация алгоритма расчета проведена сопоставлением параметров $(\mu\tau)_{e,h}$, полученных при измерении с помощью α -источника и из результатов моделирования. Исследовался CdTe детектор дозиметрического качества с размерами $4 \times 4 \times 1,5$ мм. Напряжение смещения – 130 В. Измеренное положение фотопика 60 кэВ соответствует энергии 38,5 кэВ, чувствительность детектора на линии 662 кэВ равна 24,3 импульс/мкР. На рис. 1 показана диаграмма расчета $(\mu\tau)_{e,h}$. Сплошные кривые – изолинии положения фотопика 60 кэВ. Штрихпунктирные линии – изолинии чувствительности детекторов при энергии γ -излучения 662 кэВ.

В табл. 1 приведены для сравнения рассчитанные значения $(\mu\tau)_{e,h}$ и полученные при измерении с помощью α -источника. Расхождение между ними согласуется с результатами [7], из которых следует, что значения $(\mu\tau)_{e,h}$, рассчитанные при подгонке α -спектров, обычно занижены от 1,3 до 3 раз по сравнению с данными, получаемыми при моделировании спектров. Причиной этого является то, что при определении эффективности сбора заряда в детекторах с помощью α -источника существенную роль играют поверхностные уровни-ловушки. Использование при измерениях и в расчетах γ -квантов высокой энергии (662 кэВ), равномерно проникающих на всю глубину детектора, позволяет минимизировать влияние поверхности.

Из полученного результата следует, что рассмотренный алгоритм может применяться для определения параметров переноса заряда в CdTe детекторах γ -излучения с хорошей точностью.

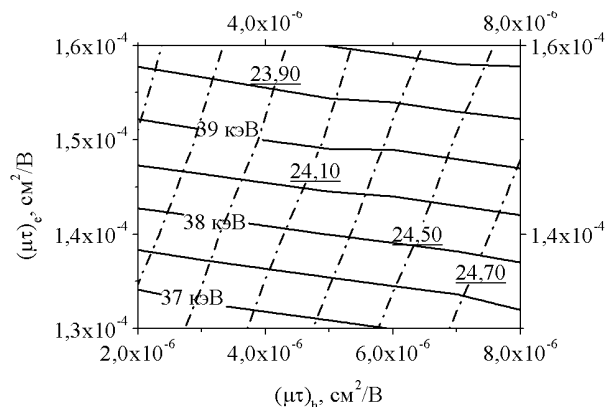


Рисунок 1 – Диаграмма расчета $(\mu\tau)_{e,h}$ для CdTe детектора размерами $4 \times 4 \times 1,5$ мм

Таблица 1 – Значения $(\mu\tau)_{e,h}$ для CdTe детектора размерами $4 \times 4 \times 1,5$ мм, полученные при моделировании и при измерении с помощью α -источника

Образец, размеры, мм	Моделирование		α -источник	
	$(\mu\tau)_e, \text{см}^2/\text{В}$	$(\mu\tau)_h, \text{см}^2/\text{В}$	$(\mu\tau)_e, \text{см}^2/\text{В}$	$(\mu\tau)_h, \text{см}^2/\text{В}$
$4 \times 4 \times 1,5$	$(1,45-1,5) \times 10^{-4}$	$(4,8-6) \times 10^{-6}$	$1,1 \times 10^{-4}$	$4,6 \times 10^{-6}$

Таблица 2 – Результаты расчета $(\mu\tau)_{e,h}$ для группы CdTe детекторов размерами $2 \times 2 \times 2$ мм

№ образца	Положение пика 60 кэВ, кэВ	Чувствительность на линии 662 кэВ, импульс/мкР	$(\mu\tau)_e, \text{см}^2/\text{В}$	$(\mu\tau)_h, \text{см}^2/\text{В}$
6	44,0	7,5	$(7,25-7,4) \times 10^{-4}$	$(4-6) \times 10^{-5}$
7	46,0	7,7	$(7,0-7,25) \times 10^{-4}$	$(5,5-8) \times 10^{-5}$
15	46,7	7,5	$(9,2-9,5) \times 10^{-4}$	$(4,8-6,5) \times 10^{-5}$
16	43,0	7,7	$(6,6-6,8) \times 10^{-4}$	$(5,5-7,8) \times 10^{-5}$

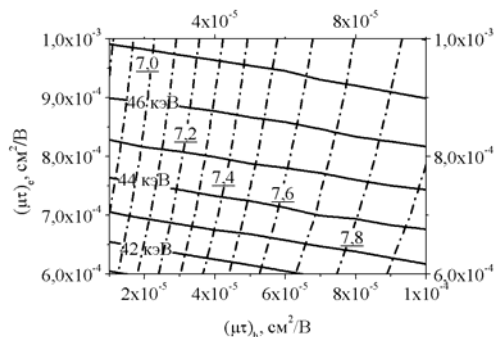


Рисунок 2 – Диаграмма расчета $(\mu\tau)_{e,h}$ для группы CdTe детекторов размерами $2 \times 2 \times 2$ мм

Дальнейшие исследования выполнены для группы CdTe детекторов размерами $2 \times 2 \times 2$ мм, изготовленных из одного слитка. Напряжение смещения – 90 В. На рис. 2 приведена диаграмма расчета $(\mu\tau)_{e,h}$, а в табл. 2 – полученные значения $(\mu\tau)_{e,h}$.

Как видно из таблицы, CdTe детекторы, изготовленные из одного слитка достаточно однородны, что согласуется с экспериментальными данными [1].

ЗАКЛЮЧЕНИЕ

В работе решена задача расчета параметров переноса заряда в планарных CdTe детекторах γ -излучения с использованием их дозиметрических характеристик.

Научная новизна работы заключается в том, что впервые предложен надежный метод расчета параметров переноса заряда в дозиметрических детекторах γ -излучения на основе высокоомных полупроводниковых соединений (на примере CdTe). В дальнейшем рассмотренный метод может быть распространен на детекторы разной геометрии (полусферические, цилиндрические и т. д.).

Практическая ценность работы состоит в том, что разработанный метод позволяет определять параметры

переноса заряда с хорошей точностью, что имеет существенное значение для калибровки детекторов. Метод может использоваться также для контроля параметров блоков детектирования в процессе их эксплуатации.

ПЕРЕЧЕНЬ ССЫЛОК

- Schlesinger T. E., Toney J. E., Yoon H. et al. Cadmium zinc telluride and its use as a nuclear radiation detector material // Materials Science and Engineering. – 2001. – 32. – P. 103–189.
- Eisen Y., Horovitz Y. Correction of incomplete charge collection in CdTe detectors // Nuclear Instruments & Methods in Physics Research A. – 1994. – 353. – P. 60–66.
- Rybka A. V., Davydov L. N., Shlyakhov I. N. et al. Gamma-radiation dosimetry with semiconductor CdTe and CdZnTe detectors // Nuclear Instruments & Methods in Physics Research A. – 2004. – 531. – P. 147–156.
- Павлов Л. П. Методы измерения параметров полупроводниковых материалов. – 2-е изд., перераб. и доп. – М.: Высшая школа. – 1987. – 239 с.
- Ruzin A., Nemirovsky Y. Methodology for evaluation of mobility-lifetime product by spectroscopy measurements in CdZnTe spectrometers // Journal of Applied Physics – 1999. – 82 (9). – P. 4166–4171.
- Miyajima S., Sakuragi H., Matsumoto M. Extraction of mean free path of charge carriers in CdZnTe crystals from measured full-energy peaks // Nuclear Instruments & Methods in Physics Research A. – 2002. – 485. – P. 533–538.
- Sato G., Parsons A., Hullinger D. et al. Development of a spectral model based on charge transport for the Swift/BAT 32K CdZnTe detector array // Nuclear Instruments & Methods in Physics Research A. – 2005. – 541. – P. 372–384.
- Захарченко А. А., Кутний В. Е., Рыбка А. В., Хажмурадов М. А. Моделирование дозиметрических характеристик CdTe (CdZnTe) детекторов γ -излучения // Радиозлектроника и информатика. 2006. – 2. – С. 28–33.

Надійшла 2.03.07

Запропоновано метод швидкого визначення параметрів переносу заряду з використанням дозиметричних характеристик детектора і з використанням математичного моделювання.

A method of fast determination of the detector charge transportation (collection) parameters with the use of the detector dosimetric characteristics and by means of mathematical simulation is proposed.

Д. М. Пиза, Н. П. Чернобородова, М. П. Чернобородов

УСОВЕРШЕНСТВОВАНИЕ МЕТЕОКАНАЛА АЭРОДРОМНОГО РАДИОЛОКАЦИОННОГО КОМПЛЕКСА «ДНЕПР-А»

Проведен анализ технических решений, использованных при проектировании аэродромного радиолокационного комплекса (АРЛК) «Днепр-А», и алгоритмов обработки метеоданных, источником которых является канал цели. Оценена эффективность работы метеоканала опытного образца локатора. Предложены пути дальнейшего усовершенствования метеоканала комплекса «Днепр-А», апробированы новые алгоритмы устранения влияния мешающих отражений от земной поверхности.

ВВЕДЕНИЕ

Непрерывное увеличение количества авиационных полетов вызывает необходимость постоянного совершенствования существующих систем управления воздушным движением (УВД). Перспектива внедрения в Европейском воздушном пространстве новой концепции кооперативного УВД (Cooperative ATM) и режима «Свободного полета» (Free flight) выдвигают более жесткие требования к средствам поддержки необходимого уровня безопасности воздушного движения. Одной из составляющих усовершенствования средств поддержки является повышение качества метеоинформации, используемой для прогнозирования погоды, оперативного предупреждения о надвигающихся опасных явлениях природы (ураганы, шквалы и др.).

Создание специализированных доплеровских метеорологических РЛС [1] определенным образом повысило качество метеорологической информации за счет извлечения из отраженных сигналов сведений о радиальной составляющей скорости метеобъектов. Однако, усовершенствование метеорологических РЛС не решило проблему оперативности поступления метеоинформации для управления воздушным движением. Альтернативой применению метеорологических РЛС является использование метеоканала в структуре аэродромных и трассовых РЛС УВД. Особенно актуально наличие встроенного метеоканала для стран, в которых отсутствует развитая сеть метеостанций. В США, например, развертываемой национальной сетью метеолокаторов нового поколения NEXRAD запланировано охватить не все аэропорты страны. В таких местах по-прежнему продолжают эксплуатировать РЛС УВД со встроенным метеоканалом – диспетчерский радар ASR-9 [2].

ПОСТАНОВКА ЗАДАЧИ

В НПК «Искра» (г. Запорожье) впервые в Украине разработан когерентно-доплеровский аэродромный радиолокационный комплекс (АРЛК) «Днепр-А» – аналог ASR-9 [3] производства США. Поэтому представляет интерес провести анализ использованных технических решений, оценить их эффективность и предложить пути дальнейшего усовершенствования метеоканала АРЛК «Днепр-А».

АНАЛИЗ ТЕХНИЧЕСКИХ РЕШЕНИЙ

В АРЛК «Днепр-А» используется вращающаяся зеркальная антенна с косеканс-квадратной диаграммой направленности. В качестве зондирующих сигналов станцией излучаются одиночные импульсы (для обзора «мертвой» зоны) и фазоманипулированные 13-позиционным кодом Баркера импульсы мощностью 130 кВт, обеспечивающие дальность действия в 170 км. Прием отраженных сигналов на длине волны в 10 см ведется по двум разнесенным на $2,5^\circ$ по углу места лучам. Ширины диаграмм направленности лучей в азимутальной плоскости – $2,5^\circ$.

В качестве входной информации для метеоканала станции используются модули спектральных составляющих доплеровских фильтров. В ближней зоне действия АРЛК с целью уменьшения влияний отражений от местных предметов обрабатывается радиолокационная информация, полученная из верхнего луча приемной антенны, а в дальней – информация нижнего луча. Оператор с рабочего места может изменять дальность, на которой необходимо переключать рассчитанную в лучах метеоинформацию для ее отображения.

Метеоканал «Днепр-А» выполнен на одном сигнальном процессоре ADSP-21160, одновременно обрабатывающем данные обоих приемных каналов станции. Результат работы метеоканала выводится на экране рабочего места оператора АРЛК в виде метеокарты. Метеоинформация может отображаться одновременно с информацией об обнаруженных воздушных судах. Размеры элементов метеокарты: $2,4 \text{ км} \times 5^\circ$ в ближней зоне (до $23,4 \text{ км}$ – может задаваться оператором АРЛК) и $2,4 \text{ км} \times 2,5^\circ$ в дальней. Для каждого элемента метео-

карты рассчитывается амплитуда отраженного от метеоявления сигнала. Эта амплитуда пропорциональна радиолокационной отражаемости метеообразования Z и позволяет провести обнаружение и классификацию метеообъектов, а также оценить интенсивность осадков. Рассчитанная амплитуда усредняется в «скользящем окне» за 3 оборота антенны, сравнивается с заранее рассчитанными значениями амплитуд сигналов от дождей различной интенсивности и каждому элементу метеокарты присваивается группа в соответствии с интенсивностью осадков. Проградуированная таким образом метеокарта отображается на индикаторе. Градации отличаются цветом.

Для вычисления амплитуды сигналов A_m , отраженных от метеоявлений, в отсутствие отражения от земной поверхности (преимущественно, в дальней зоне) модули амплитуд с выходов фильтров дискретного преобразования Фурье (ДПФ) суммируются по правилу сложения мощностей [1]:

$$A_m = \sqrt{A_0^2 + A_1^2 + \dots + A_{N-1}^2}, \quad (1)$$

где $A_0 \dots A_{N-1}$ – амплитуды сигналов с выходов фильтров преобразования Фурье, N – количество фильтров.

Поскольку источником обрабатываемых данных является канал цели, то, для уменьшения влияния мешающих отражений от земной поверхности, применяемые алгоритмы обнаружения метеоявлений основаны на отличиях спектральных характеристик последних. Эти отличия достигнуты благодаря высокой стабильности передатчика, применению весового окна при вычислении преобразования Фурье с малыми уровнями боковых лепестков, а также уменьшению межфильтрового «просачивания» (последнее обеспечивается формированием нулей АЧХ, расположенных на центральных частотах главных лепестков остальных фильтров). Для понимания сути реализованных в метеоканале АРЛК «Днепр-А» алгоритмов обработки рассмотрим примеры радиолокационных спектров.

На рис. 1, а представлены амплитуды модулей спектральных составляющих отраженных сигналов от местных предметов, находящихся на расстоянии 4,5...5,1 км от АРЛК, нескольких подряд излученных частотных пачек в азимутальном секторе 165°19'...167°49'. Нумерация фильтров в пределах каждой частотной пачки – слева направо. Из анализа рис. 1, а следует, что:

- спектральные составляющие местных предметов сосредоточены в 0, 1 и $N-1$ фильтрах с равноамплитудными значениями в первом и $N-1$ фильтрах;
- уровень амплитуд в остальных фильтрах (2, ..., $N-2$) не зависит от наличия неподвижных местных предметов.

На рис. 1, б представлены модули амплитуд спектральных составляющих отраженных сигналов от дождя средней интенсивности. Спектр отраженных сигналов сосредоточен в трех-четыре фильтрах с одним-двумя локальными максимумами, причем амплитуды сигналов в них отличаются незначительно. Это дает возможность различать их и сигналы от местных предметов.

Если в одном элементе разрешения по дальности одновременно присутствуют отклики от местного предмета и метеообразования, и их спектры перекрываются, то симметрия сигналов в первом и $N-1$ фильтрах нарушается. На рис. 1, в представлены амплитуды сигналов на выходах фильтров ДПФ, отраженных от местного предмета и метеоявления, где видно нарушение симметрии спектральных составляющих в первом и $N-1$ доплеровских фильтрах.

Для устранения влияния отражений от земной поверхности на рассчитываемую метеоканалом величину A_m , в ближней зоне АРЛК реализуется довольно громоздкий алгоритм уменьшения влияния этих отражений (в дальней зоне предполагается их полное отсутствие) путем сглаживания на интервале дальности 2,4 км (по размеру элемента метеокарты) амплитуд одноименных фильтров с учетом асимметрии в первом и $N-1$ фильтрах. Основу операции сглаживания составляет усреднение амплитуд с отбрасыванием нескольких

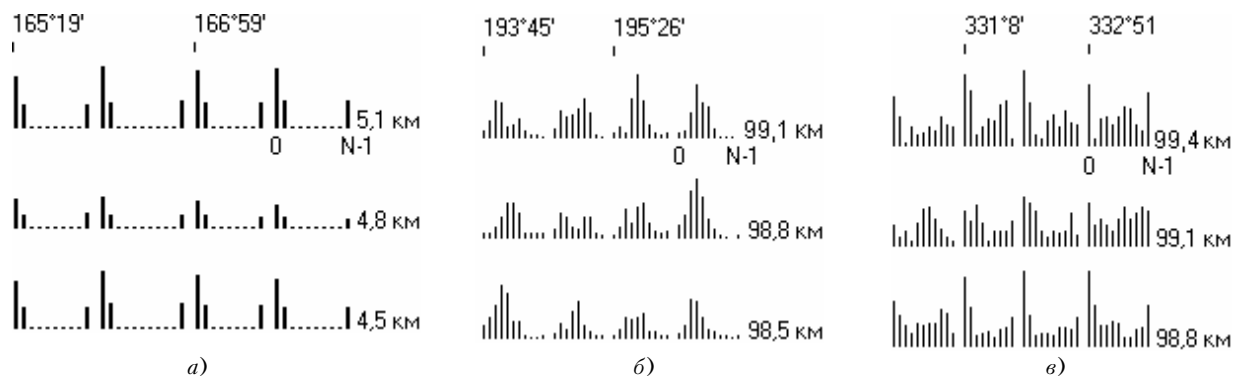


Рисунок 1 – Модули спектральных составляющих местного предмета (а), дождя средней интенсивности (б), местного предмета и облака (в)

максимальных и минимальных значений по информации нескольких частотных пачек. В дальней зоне величина A_m определяется согласно (1) по этим усредненным значениям. В ближней зоне после усреднения выполняется поиск большего из локальных максимумов $A_{\max i}$. Если он не относится к нулевому фильтру, то в качестве A_m принимается найденное значение:

$$A_m = \max(A_{\max i}). \quad (2)$$

В противном случае, в отсутствие асимметрии, A_m вычисляется по информации первого и $N - 1$ фильтров:

$$A_m = f_1(A_1, A_{N-1}), \quad (3)$$

а при наличии асимметрии – нулевого и меньшего из первого и $N - 1$ фильтров:

$$A_m = f_2[A_0, \min(A_1, A_{N-1})], \quad (4)$$

где f_1 и f_2 – обозначение функциональной зависимости; A_0, A_1, A_{N-1} – амплитуды нулевого, первого и $N - 1$ фильтров, соответственно.

РЕЗУЛЬТАТЫ ЭКСПЕРИМЕНТАЛЬНЫХ ИССЛЕДОВАНИЙ МЕТЕОКАНАЛА И ПОВЫШЕНИЕ ЕГО ЭФФЕКТИВНОСТИ

Недостаточное качество представления метеоданных наряду с выявленными (в ходе наладки опытного образца радиолокатора в г. Запорожье) низкой вероятностью правильного обнаружения метеоявлений и высокой вероятностью их ложных обнаружений вызвало необходимость проведения дальнейших исследований.

Для оценки результата работы метеоканала, на рис. 2 приведен внешний вид экрана оператора АРЛК «Днепр-А» в режиме отображения только метеоинформации. Показанная на рис. 2 большая изрезанность внешних контуров, отсутствие внутренних элементов изображения облаков, а также большое количество разрозненных метеоячеек различных уровней интенсивности обусловлено, во-первых, большими размерами элемента разрешения метеокарты. Это значительно затрудняет возможности своевременного обнаружения процесса зарождения стремительно развивающихся кучево-дождевых облаков (Сb), наиболее опасных для полетов воздушных транспортных средств. Поэтому, предложено уменьшить размеры элемента метеокарты до $600 \text{ м} \times 1^\circ$ (согласно размерам кучевых облаков Сu [4]).

Во-вторых, это обусловлено также и неверным предположением об отсутствии мешающих отражений от земной поверхности в дальней зоне: на рис. 1, в продемонстрировано присутствие таковых на дальностях 98,8...99,4 км, что вчетверо превышает границу ближ-



Рисунок 2 – Внешний вид экрана оператора АРЛК «Днепр-А» в режиме отображения только метеоинформации

ней зоны в 23,4 км (даже в условиях равнинной местности). Кроме того, при одном и том же местоположении РЛС может иметь место существенная изрезанность ландшафта местности как по дальностной, так и по азимутальной координатам. Вследствие этого предложено отказаться от разделения обрабатываемой дальности на ближнюю и дальнюю зоны и пересмотреть алгоритмы устранения влияния отражений от земной поверхности.

В-третьих, это обусловлено характером примененных алгоритмов расчета величины A_m . Поскольку метеобразования имеют широкий спектр [1], то оценка A_m по значению амплитуд одного-двух доплеровских фильтров (согласно (2)–(4)) будет существенно меньшей, чем согласно (1). Как показано на рис. 1, б, в, основная часть спектра отраженных сигналов от метеоявлений сосредоточена как минимум в трех-четырех фильтрах. Следовательно, для повышения точности расчета величины A_m необходимо обрабатывать возможно большее число фильтров.

Частое чередование ячеек карты с разными уровнями градаций (см. рис. 2) и существенные изменения отображаемой метеоинформации от оборота к обороту происходило из-за неправильного выделения метеобразований на фоне отражений от земной поверхности. Поэтому, если в спектре принятого сигнала присутствуют отражения от местного предмета, то при наличии максимума в нулевом фильтре и равноамплитудных значениях в первом и $N - 1$ фильтрах величину A_m предложено рассчитывать согласно (1), но без учета амплитуд упомянутых фильтров по формуле:

$$A_m = \sqrt{A_2^2 + A_3^2 + \dots + A_{N-2}^2}. \quad (5)$$

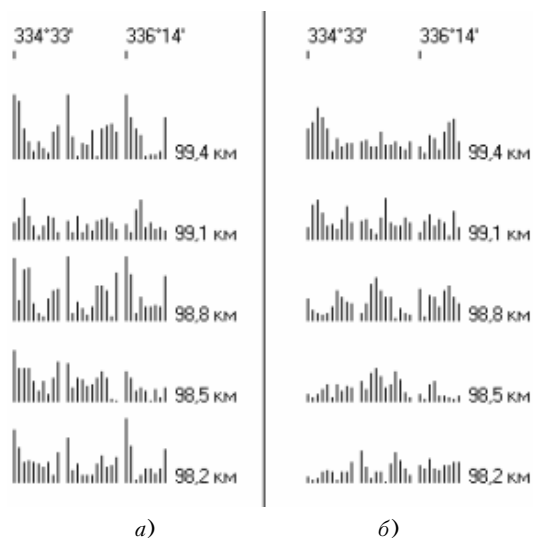


Рисунок 3 – Спектр сигналов, отраженных от метеообразований, в нижнем луче (а) и верхнего луча (б) на фоне мешающих отражений от земной поверхности

Если в одном элементе разрешения по дальности одновременно присутствуют отклики от местного предмета и метеообразования, и их спектры перекрываются, то симметрия сигналов в первом и $N - 1$ фильтрах нарушается. В этом случае амплитуду сигнала от метеобъекта предложено рассчитывать по формуле, в которой исключена амплитуда сигнала нулевого фильтра и учтена разность амплитуд первого и $N - 1$ фильтров (эта разность несет соответствующую часть информации о метеообразовании):

$$A_m = \sqrt{|A_1^2 - A_{N-1}^2| + A_2^2 + A_3^2 + \dots + A_{N-2}^2}. \quad (6)$$

Дополнительные возможности классификации метеобъектов получены путем перехода к межканальной обработке, основанной на различиях в спектральных портретах одного и того же участка пространства по высоте. Если в нижнем луче в одном элементе разрешения по дальности присутствуют отражения от метеообразования и мешающих отражений от земной поверхности, то в спектре отраженного сигнала от этого же элемента в верхнем луче такая помеха может отсутствовать. На рис. 3 продемонстрированы такие различия: спектра сигналов нижнего луча (см. рис. 3, а) – сформирован из отражений от земной поверхности и метеобъектов спектра сигналов верхнего луча (см. рис. 3, б) – наблюдаются отражения только от метеоявлений. Поэтому, в таких случаях для повышения точности выделения метеобъектов необходимо брать для оценки амплитуды метеообразований A_m информацию верхнего луча. Значительное превышение амплитуды метеосигнала в нижнем луче над верхним может



Рисунок 4 – Результат работы усовершенствованного алгоритма обработки и отображения метеоинформации

свидетельствовать о выпадении осадков или, например, наличии горизонтальных сдвигов ветра.

На рис. 4 приведен пример обработки исходных данных для рис. 2 с учетом предложенных усовершенствований метеоканала. Как показано на рис. 4, изменения в алгоритмах обработки метеоданных позволили получить не только качественное улучшение отображения метеобстановки, но и повысить вероятность правильного обнаружения метеоявлений (например, на рис. 2 отсутствуют облака в левой части рисунка).

ВЫВОДЫ

Проведен анализ технических решений, использованных при проектировании метеоканала аэродромного радиолокационного комплекса «Днепр-А», и алгоритмов обработки метеоданных, источником которых является канал цели. Оценена эффективность работы метеоканала опытного образца локатора. Предложены пути дальнейшего усовершенствования метеоканала комплекса «Днепр-А», апробированы новые алгоритмы устранения влияния мешающих отражений от земной поверхности.

Дальнейшие работы по совершенствованию метеоканала следует вести по таким направлениям:

- повышение точности оценки амплитуды метеосигнала с учетом ожидаемого коэффициента заполнения облучаемого импульсного объема метеобъектом;
- улучшение распознавания метеобъектов путем сравнения данных метеоканала, метеолокатора и метеостанции.

ПЕРЕЧЕНЬ ССЫЛОК

1. Довиак Р. Дж. Метеорологические доплеровские РАС / Р. Дж. Довиак, Д. С. Зрнич, Д. С. Сирманс // ТИИЭР. – М., 1979. – Т. 67, № 11. – С. 63–102.
2. Mark E. Weber Advances in Operational Weather Radar Technology // Lincoln Laboratory Journal. – Lexington, 2006. – Vol. 16, № 1. – P. 9–22.
3. Тейлор-мл. Дж. Новая диспетчерская радиолокационная станция ASR-9 / Дж. Тейлор-мл., Г. Брунинс // ТИИЭР. – М., 1985. – Т. 73, № 2. – С. 128–135.
4. Яновський Ф. Й. Метеонавігаційні радіолокаційні системи повітряних суден: Навчальний посібник. – К.: НАУ, 2003. – 304 с.

Надійшла 16.11.07

Проведено аналіз технічних рішень, використаних при проектуванні аеродромного радіолокаційного комплексу (АРЛК) «Дніпро-А», і алгоритмів обробки метеоданих,

джерелом яких є канал цілі. Оцінено ефективність роботи метеоканалу дослідного зразка локатору. Запропоновано шляхи подальшого удосконалення метеоканалу комплексу «Дніпро-А», випробувані нові алгоритми зменшення впливу відбиттів від земної поверхні.

The analysis of technical solutions have been performed, when making a design of airfield radar complex (ARC) «Dnepr-A» weather channel, as well as of algorithms of processing weather data a source of which is target channel. There has been estimated efficiency of weather channel operation within the ARC development model. The ways of further enhancement of complex «Dnepr-A» weather channel were offered, new algorithms of eliminating an effect of underlying surface clutter have been approved.

УДК 681.51.012: 539.23

В. И. Псарев, Л. А. Пархоменко, Ю. В. Пьянкова

О ТЕХНИКЕ ИДЕНТИФИКАЦИИ ПРИ СИСТЕМНОМ АНАЛИЗЕ ТОНКИХ ОСТРОВКОВЫХ ПЛЕНОК

Получены аналитические формулы для проведения системного анализа структурного огрубления островковых пленок на подложке. Для этой цели предложен метод установления сходства и различия между характеристиками теоретического и экспериментального распределения островков по размерам. Предложенный метод позволяет получать информацию о протекании процессов в системе микроостровков по мере их приближения к состоянию равновесия.

ВВЕДЕНИЕ

Применение тонких пленок в технике стало возможным после освоения методов их получения с заданными физико-химическими свойствами. Это открыло широкие возможности использования тонких пленок в оптике, космической и атомной промышленности и СВЧ-технике; в качестве элементов микросхем и тензодатчиков; в криогенной технике и других областях. Качество и надежность работоспособного состояния в процессе их функционирования в устройствах и приборах, в особенности, под влиянием внешних факторов, существенно зависят от структурной стабильности пленочного материала, достоверная информация о которой должна быть заранее получена методами системного физико-химического анализа. Рассмотрим в этой связи пленки островкового типа, подверженные огрублению из-за остwaldовской коагуляции микроостровков, осложненной рядом сопутствующих процессов. Их существенной характеристикой является функция плотности распределения островков по размерам.

© Псарев В. И., Пархоменко Л. А., Пьянкова Ю. В., 2007

Вызываемое внутрисистемными процессами огрубление островков изменяет характер их распределения по размерам, что оказывает существенное влияние на формирование микроструктурного состояния и свойства пленочного материала. Познание внутрисистемных процессов по признакам вызываемой ими трансформации экспериментальных распределений – задача, решение которой может быть распространено на широкий класс островковых пленок. Подразумевается возможность получения полезной информации путем выявления сходства и различия между экспериментальным распределением – образом и теоретической функцией распределения микроостровков по размерам – подобием, полученной с учетом определенных физических представлений.

Изучению кинетических особенностей огрубления островковых пленок на подложке посвящены многие работы [1–3]. Однако, в них не учитывалось, что огрубление микроостровков – процесс многофакторный: определяется влиянием структурного, диффузионного, межфазного и других факторов. При этом важное значение приобретает возможность установления характера влияния каждого из факторов в отдельности и в своей их совокупности на кинетические особенности изменения дисперсности микроостровков в процессе их огрубления.

В настоящей работе приведено описание предлагаемой методики системного анализа тонких островко-

вых пленок с привлечением традиционных методов статистического анализа и средств ЭВМ, необходимых для проведения сложных расчетов.

1 ТЕОРЕТИЧЕСКИЕ ПОЛОЖЕНИЯ

Функцию плотности распределения островков по их относительным размерам, отображающую их дисперсность в самых разных случаях, согласно приведенным в работе [4] формулам (1) и (2), можно записать в таком виде:

$$\varphi(u) = \frac{3v_k \gamma(u)}{v_k u \gamma(u) - u + 1} \exp\left[-3v_k \int \frac{\gamma(u) du}{v_k u \gamma(u) - u + 1}\right], \quad (1)$$

где $v_k = v_k(t)$ – приведенная скорость движения r_k в пространстве размеров островков, t – время; $\gamma = \gamma(u)$ – фактор, определяющий механизм процесса; $u = \frac{r}{r_k}$, r – эффективный радиус островка, r_k – критический радиус.

Каждый островок радиуса r_0 на поверхности подложки можно дополнить до сферы радиуса $r = r_0 / \sin \theta$, где θ – краевой угол, и таким образом подойти к эквивалентному распределению сфер в объеме [3]. Переход от числа куполообразных островков на единице площади подложки $n = f(r_0)$ к распределению соответствующих им эффективных сфер в единице объема $N_r = f(r)$ производится с помощью формулы пересчета от двумерного распределения островков к трехмерному [5]. Обратный переход можно осуществить с помощью обращенной формулы той же работы [5].

В размерных переменных эквивалентная функция плотности распределения $f(r, t) = C(r_k) r_k^{-4} \varphi(u)$ существенно зависит от явного вида $\gamma(u)$. В случае диффузионно-контролируемого механизма [6, 7] укрупнения островков в системе $\gamma(u) = u^s$, где s – параметр массопереноса от расплывающихся на поверхности подложки островков к растущим на ней. В случае металлических островковых пленок [3] $\gamma(u) = \varepsilon u^3 + u^{1-\alpha}$, где $\varepsilon = \frac{K r_k^2}{D_s}$, K – скорость прохождения адатомов (адсорбированных атомов) через межфазную границу, D_s – коэффициент поверхностной диффузии адатомов.

При задании фактора $\gamma(u)$ в явном виде с помощью формулы (1) можно определить функцию $\varphi(u)$. Пусть, например, $\gamma(u) = u^s$, тогда формулу (1) можно представить в таком виде

$$\varphi(u) = \frac{3v_k u^s}{v_k u^{1+s} - u + 1} \exp\left[-3v_k \int \frac{u^s du}{v_k u^{1+s} - u + 1}\right]. \quad (2)$$

Учитывая, что параметр $0 \leq s < \infty$, знаменатель в формуле (2) можно разложить на множители: $v_k u^{1+s} - u + 1 = v_k (u_g - u)^2 \psi(u)$. В случае целочисленных значений s многочлен $\psi(u) = \sum_i^s i u^{s-i} \left(\frac{1+s}{s}\right)^{i-1}$; $i = 1, 2, 3, \dots, s$; $u_g = \frac{r_g}{r_k}$, r_g – наибольший размер микроостровков (размах функции $f(r, t)$). Для разных значений s можно получить известные виды теоретических распределений: при $s = 1$ множитель $\psi(u) = 1$ и после интегрирования получим распределение Вагнера [7]; при $s = 2$ и $\psi(u) = u + 3$ получим распределение Лифшица – Слезова [6]; при $s = 3$ и $\psi(u) = u^2 + 2\left(\frac{4}{3}\right)u + 3\left(\frac{4}{3}\right)^2$ распределение будет соответствовать иной системе островков и т. д.

При сопоставлении экспериментальной гистограммы с теоретическим распределением возникает необходимость в определении численных значений параметров, определяющих фактор $\gamma(u)$ и обеспечивающих максимальное сходство между образом и подобием. Такая задача решалась для частного случая [8]. Учитывая выражение (1), путем математических преобразований, аналогичных применяемым в работе [8], общее решение может быть записано в таком виде:

$$v_k(n-3)M_1 + 2v_k M_2 - nM_3 - 2M_4 = 0, \quad (3)$$

где $M_1 = \int_0^{u_g} \gamma^2 u^n \varphi(u) du$; $M_2 = \int_0^{u_g} \gamma \gamma' u^{n+1} \varphi(u) du$; $M_3 = \int_0^{u_g} \gamma u^{n-1} (u-1) \varphi(u) du$; $M_4 = \int_0^{u_g} \gamma' u^n (u-1) \varphi(u) du$; γ' – первая производная по u от $\gamma = \gamma(u)$; n – целые и дробные положительные числа.

При задании $\gamma(u)$ уравнение моментов (3) преобразуется и приобретает конкретное содержание, необходимое и достаточное для определения численных значений параметров островковой системы. Их испытание на достоверность является следующим шагом системного анализа.

По мере движения системы микроостровков к состоянию равновесия происходит непрерывная трансформация функции плотности распределения. Изменяются численные значения ее характеристик: размах, мода и модальное значение размера островков, асимметрия и другие. Сохраняет инвариантность только соотношение между смешанными моментами функции плотности распределения. Воспользуемся для его вывода уравнением моментов распределений, приведенным в рабо-

те [9] (формула (7)). Решая его с учетом (1), получим взаимосвязь между моментами распределения такого вида:

$$(3-n)v_k M_{nm} = nL_n + mL_m, \quad (4)$$

где $M_{nm} = \int_0^{u_g} M du$ – смешанный момент порядка n и m ; n и m – целые и дробные положительные числа, $M = u^n(u_g - u)^m \varphi(u)$; в терминах моментов:

$$\begin{aligned} L_n &= M_{n-1-\gamma, m} - M_{n-\gamma, m}; \\ L_m &= M_{n+1-\gamma, m-1} - v_k M_{n+1, m-1} - M_{n-\gamma, m-1}; \\ M_{n-1-\gamma, m} &= \int_0^{u_g} M[u\gamma(u)]^{-1} du; \\ M_{n-\gamma, m-1} &= \int_0^{u_g} M[(u_g-u)\gamma(u)]^{-1} du; \\ M_{n-\gamma, m} &= \int_0^{u_g} M[\gamma(u)]^{-1} du; \\ M_{n+1, m-1} &= \int_0^{u_g} uM(u_g-u)^{-1} du; \\ M_{n+1-\gamma, m-1} &= \int_0^{u_g} uM[(u_g-u)\gamma(u)]^{-1} du. \end{aligned}$$

На отдельных частных примерах проиллюстрируем применение полученных формул и уравнений.

2 ЧАСТНЫЕ РАСПРЕДЕЛЕНИЯ

Огрубление микроостровков при диффузионно-контролируемом механизме характеризуется фактором $\gamma(u) = u^s$ и скоростью $v_k(t)$. При непрерывном изменении параметра s из формул (3)–(6) [4] следует система уравнений

$$\begin{aligned} u_g &= (1+s)s^{-1}; \quad v_k = (1+s)^{-1}u_g^{-s} = u_g^{-(1-s)}s^{-1}, \\ 4v_k u^{1+s} + u(s-1) - s \Big|_{u_m} &= 0, \end{aligned} \quad (5)$$

$$\begin{aligned} 20v_k^2 u^{2+2s} - 13v_k u^{1+s} z + 2z^2 - \\ - [s(s-1)(u-1) + 2sz](v_k u^{1+s} - u + 1) \Big|_{u_p} &= 0, \end{aligned}$$

где $z = s + u(1-s)$; $u_m = \frac{r_m}{r_k}$, r_m – модальный радиус микроостровков; $u_p = \frac{r_p}{r_k}$, r_p – значение радиуса в точке перегиба на кривой плотности распределения.

При том же условии в соотношении моментов (4) имеем

$$\begin{aligned} L_n &= M_{n-1-s, m} - M_{n-s, m} \\ &\text{и} \\ L_m &= M_{n+1-s, m-1} - v_k M_{n+1, m-1} - M_{n-s, m-1}. \end{aligned}$$

В частности, положив $n = 3$, получим $3|L_3| = m|L_m|$, а при $n = 3$ и $m = 0$ имеет место равенство $M_{2-s} = M_{3-s}$.

Уравнение (3) при $\gamma(u) = u^s$ можно преобразовать к виду

$$v_k(2s+n-3)M_{n+2s} + (n+2s)M_{n+s} - (n+2s)M_{n+s-1} = 0. \quad (6)$$

Или же в размерных переменных, после сокращения на r_k^{n+s-2} , получим

$$r_k^{1+s} - \frac{M'_{n+s}}{M'_{n+s-1}} r_k^s + \frac{v_k(2s+n-3)}{n+2s} \frac{M'_{n+2s}}{M'_{n+s-1}} = 0, \quad (7)$$

где

$$\begin{aligned} M'_{n+s} &= \int_0^{r_g} r^{n+s} f(r, t) dr; \\ M'_{n+s-1} &= \int_0^{r_g} r^{n+s-1} f(r, t) dr; \\ M'_{n+2s} &= \int_0^{r_g} r^{n+2s} f(r, t) dr. \end{aligned}$$

Задавая значение n , можно сформировать систему уравнений и, решая ее, определить s и r_k . Если $n = 3 - 2s$, то из уравнения (6) следует: $M_{3-s} = M_{2-s}$, а из уравнения (7): $r_k = \frac{M'_{3-s}}{M'_{2-s}} = u_s^{-1} r_s$, где u_s – среднее значение u в распределении $\varphi(u)$; r_s – среднее значение r в распределении $f(r, t)$. Численное значение $k_s = u_s^{-1}$ используется при переходе от экспериментального среднего радиуса r_s к критическому r_k .

В табл. 1 приведены массивы данных четырех примерных распределений. Первое из них характеризуется отрицательной асимметрией, достигает максимума при $u_m = 1$, а верхняя граница $u_g \approx 2$. Предполагая, что это распределение типа Вагнера [7], т. е. $s = 1$, после нормирования находим моменты M_1 и M_2 . Их отношение $\frac{M_2}{M_1} = 0,99988 \approx 1$, значение $u_s = 0,89106$. Следовательно, $k_s = u_s^{-1} = 1,1223 \approx 9/8$ [7]. Предположение подтверждается. Согласно формулам и урав-

Таблица 1 – Данные массивов и характеристики примерных распределений при разных значениях фактора $\gamma(u) = u^s$

№ массива	Значения варианты u									
	0	0,2	0,4	0,6	0,8	1,0	1,2	1,4	1,6	1,8
1	0	3,776	8,971	15,355	21,663	24,288	20,255	8,174	0,478	0,001
2	0	5,776	10,971	17,355	23,663	24,288	20,255	8,174	0,478	0,001
3	0	9,776	14,971	21,355	27,663	24,288	20,255	8,174	0,478	0,001
4	0	3,776	8,971	15,355	21,663	24,288	22,255	10,174	2,478	0,01
	s	$q_s \times 10^4$	u_g	v_k	u_m	u_{p_1}	u_{p_2}	u_{p_3}	u_s	k_s
1	1	1,2	2	0,25	1	–	0,6084	1,3458	0,8911	1,1223
2	0,82	3,0	2,2195	0,2858	0,922	0,1053	0,4129	1,3328	0,8630	1,1587
3	0,546	0,6	2,8315	0,3664	0,7136	–	–	1,2390	0,8167	1,2244
4	1,34	0,2	1,7452	0,2022	1,082	–	0,8078	1,3361	0,9186	1,0886

Примечание. $q_s = \left| 1 - \frac{M_{3-s}}{M_{2-s}} \right|$ – параметр идентификации распределений.

нениям (5) находим $u_g = 2$; $v_k = 0,25$; $u_m = 1$; $u_{p_1} = 0,6084$; $u_{p_2} = 1,3458$, что соответствует распределению Вагнера. Его вид следует из выражения (2) при $s = 1$: $\varphi(u) = Cu(2-u)^{-5} \exp\left(-\frac{6}{2-u}\right)$.

Аналогичным образом проведены расчеты и для других массивов примерных распределений при условии $M_{3-s} = M_{2-s}$. Характерно, при $s < 1$ коэффициент $k_s > 9/8$; при $s > 1$ величина $k_s > 9/8$.

При задании фактора $\gamma(u) = \varepsilon u^3 + u^{1-\alpha}$ значения характеристик u_g , v_k , u_m и u_p также можно определить из формул (3)–(6) [4]

$$3\varepsilon u^{3+\alpha} - 4\varepsilon u^{2+\alpha} + (1-\alpha)u - 2 + \alpha \Big|_{u_g} = 0,$$

$$v_k = (3u - 4)u^{\alpha-2}(2 + \alpha)^{-1} \Big|_{u_g}, \quad (8)$$

$$u(\alpha + 3\varepsilon u^{1+\alpha} - 2\varepsilon u^{2+\alpha}) + 1 - \alpha = 4v_k u^{2-\alpha} (1 + \varepsilon u^{2+\alpha})^2 \Big|_{u_m},$$

$$20v_k^2 \gamma^4 - 13v_k \gamma^2 z + 2z^2 - [\gamma \gamma''(u-1) + 2\gamma' z](v_k u \gamma - u + 1) \Big|_{u_p} = 0,$$

где $z = \gamma - \gamma'(u-1)$; γ' и γ'' – первая и вторая производные по u от $\gamma = \gamma(u)$. Полученные иным путем уравнения системы (8) содержатся в работах [1–3].

Определение численных значений параметров α и ε можно произвести с помощью уравнения моментов, которое вытекает из выражения (3)

$$\frac{v_k \varepsilon^2 (n+3+\alpha)}{r_k^{(n+5+\alpha)}} M'_{n+6+\alpha} + \frac{2\varepsilon v_k (n+1)}{r_k^{(n+3)}} M'_{n+4} - \frac{\varepsilon (n+6+\alpha)}{r_k^{(n+2+\alpha)}} M'_{n+3+\alpha} + \frac{\varepsilon (n+6+\alpha)}{r_k^{(n+1+\alpha)}} M'_{n+2+\alpha} + \frac{v_k (n-1-\alpha)}{r_k^{(n+1-\alpha)}} M'_{n+2-\alpha} - \frac{n+2-\alpha}{r_k^n} M'_{n+1} + \frac{n+2-\alpha}{r_k^{(n-1)}} M'_n = 0, \quad (9)$$

где $M'_\chi = \int_0^{r_g} r^\chi f(r, t) dr$ – моменты распределения порядка $\chi = n, n+1, n+2+\alpha, \dots$ при целочисленном и дробном значении n .

3 СОПОСТАВЛЕНИЕ С ЭКСПЕРИМЕНТОМ

Воспользуемся данными распределения платиновых островков на $\gamma\text{-Al}_2\text{O}_3$ -подложке по их эффективным радиусам (взяты из работ [10, 1]). В табл. 2 приведены количественные характеристики распределений островков по размерам в пленке, нагревание которой производилось на воздухе при температуре 700 °С. Идентификация экспериментальных гистограмм (нормированных на единицу) производилась в предположении действия диффузионно-контролируемого механизма (1) ($\gamma(u) = u^s$) и механизма (2), дополнительно учитывающего межфазный фактор ($\gamma(u) = \varepsilon u^3 + u^{1-\alpha}$).

Характеристики распределений островков по размерам (u_g, v_k, u_m, u_p, r_k и др.) после одного часа нагрева при 700 °С, рассчитанные в предположении действия одного и другого механизмов, близки по своим численным значениям. Это указывает на то, что основ-

Таблиця 2 – Количественные характеристики гистограмм островковой пленки Pt на $\gamma\text{-Al}_2\text{O}_3$ -подложке после нагрева при 700 °C на воздухе [10]

r , нм	$n_s \cdot 10^{-7}$, мм ⁻²	$N_r \cdot 10^{-12}$, мм ⁻³	Механизм (1)	Механизм (2)
1 час; $n_0 = 6,972 \cdot 10^8$ мм ⁻² ; $N_0 = 31,146 \cdot 10^{12}$ мм ⁻³ ; $r_s = 8,0$ нм 1 ($s = 0,754$; $q_s = 8,4 \cdot 10^{-5}$); 2 ($\alpha = 0,50$; $\varepsilon = 0,01$); $g = 26,25$ %				
2,14	8,68	2,0023	$u_g = 2,3263$	$u_g = 2,4080$
5,0	15,12	3,2531	$v_k = 0,3017$	$v_k = 0,3450$
7,86	31,36	17,7142	$u_m = 0,8837$	$u_m = 0,7313$
10,72	13,44	7,6514	$u_p = 1,3214$	$u_p = 1,2816$
13,58	0,84	0,3902	$r_k = 8,85$ нм	$r_k = 9,26$ нм
16,44	0,28	0,1359	$r_m = 7,82$ нм	$r_m = 6,77$ нм
19,30	0	0	$r_g = 20,6$ нм	$r_g = 22,3$ нм
16 часов; $n_0 = 4,635 \cdot 10^8$ мм ⁻² ; $N_0 = 20,51 \cdot 10^{12}$ мм ⁻³ ; $r_s = 9,05$ нм 1 ($s = 0,546$; $q_s = 6,5 \cdot 10^{-5}$); 2 ($\alpha = 0,257$; $\varepsilon = 0,01$); $g = 41,74$ %				
1,69	1,70	0,011	$u_g = 2,8315$	$u_g = 2,1572$
3,93	6,50	1,610	$v_k = 0,3654$	$v_k = 0,2867$
6,17	11,15	5,465	$u_m = 0,7136$	$u_m = 0,9085$
8,41	10,35	4,874	$r_k = 10,63$ нм	$r_k = 10,43$ нм
10,65	9,0	4,731	$r_m = 7,56$ нм	$r_m = 9,48$ нм
12,89	4,5	2,134	$r_g = 30,10$ нм	$r_g = 22,5$ нм
15,13	2,7	1,455	$u_p = 1,239$	$u_{p_1} = 0,1986$
17,37	0,45	0,243	$k_s = 1,1747$	$u_{p_2} = 0,3106$
19,61	0	0	–	$u_{p_3} = 1,3377$

ным лимитирующим фактором их огрубления в пленке на подложке является поверхностная диффузия адатомов от расплывающихся ($r < r_k$) к растущим ($r > r_k$) микроостровкам. Теоретические функции плотности распределения, соответствующие одному и другому механизму, практически с одинаковой степенью точности описывают экспериментальное распределение. В этом случае из-за ограниченности экспериментального материала трудно выявить степень влияния каждого из факторов на процесс огрубления островков в пленке.

Результаты расчетов экспериментальных данных после нагрева в течение 16 часов оказались противоречивыми. При действии механизма (1) кривая плотности распределения имеет одну точку перегиба; в случае механизма (2) – три точки перегиба. По мере приближения островковой системы к состоянию равновесия параметр s при диффузионном механизме должен увеличиваться. Он же уменьшается от $s = 0,754$ (1 час) до $s = 0,546$ (16 часов). Это обстоятельство приводит к необоснованному увеличению со временем характеристик u_g и v_k и уменьшению значения u_m .

В то же самое время при действии возможного механизма (2) идет закономерная оствальдовская коагуляция островков на $\gamma\text{-Al}_2\text{O}_3$ -подложке. Уменьшаются со временем значения u_g и v_k при одновременном увеличении величины u_m . Экспериментальное распределение после 16 часов нагрева практически совпадает с теоретическим: $M_{3-\gamma(u)} = 0,8419$ и $M_{2-\gamma(u)} = 0,8437$. В системе увеличивается доля растущих островков от $g = 26,28$ % до $g = 41,74$ % после 16 часов нагрева пленки.

Полученный результат по своему содержанию вскрывает проявление эффектов термической деградации и так называемого «отравления» керамических катализаторов (применяются в нефтеперерабатывающей промышленности). Для их изготовления используются чешуйки $\gamma\text{-Al}_2\text{O}_3$ с осажденными на них островками металлов [10]. Воздействие внешнего фактора на межфазную границу (согласно [10] – образование слоя PtO_2 на поверхности островков при нагревании пленок на воздухе) оказывает существенное влияние на диффузионные потоки адатомов в системе островков, что ис-

ключает возможность закономерной их коагуляции в соответствии с механизмом (1).

ЗАКЛЮЧЕНИЕ

Показано, что изменение характерных структурных свойств островковых пленок можно выявить путем системного анализа распределений микроостровков по размерам, их собственных моментов и соотношений между ними с привлечением средств ЭВМ. Дано описание способа проведения такого анализа в системе островков на подложке. Для этой цели устанавливается сходство и различие экспериментального распределения – образа с теоретическим – подобием, учитывающим возможный внутрисистемный механизм изменения дисперсности микроостровков при их огрублении.

Получены аналитические формулы и уравнения, с помощью которых можно оценить качество и достоверность результатов идентификации распределений. Приведены примеры их применения с учетом действующих ограничений.

Предлагаемая методика системного анализа островковых пленок вполне может обеспечить достоверную информацию об их структурной стабильности, необходимой, например, при изготовлении пленочных микросхем в радиоэлектронной отрасли.

ПЕРЕЧЕНЬ ССЫЛОК

1. Псарев В. И., Пархоменко Л. А., Куликов А. Ф. Структурная устойчивость тонких островковых пленок. Расчеты и компьютерный анализ // Радиоэлектроника. Информатика. Управление. – 2003. – № 1. – С. 16–20.
2. Псарев В. И., Пархоменко Л. А., Куликов А. Ф. Анализ устойчивости островковых тонких пленок. // Складні системи і процеси. – 2003. – № 2. – С. 25–32.

3. Псарев В. И. Компьютерный анализ металлических островковых пленок // Металлы. – 1999. – № 6. – С. 105–110.
4. Psarev V. I., Parkhomenko L. A. A new method for the structural analysis of thin island films // Technical Physics Letters. – 2005. – V. 31. – No. 10. – P. 888–890.
5. Псарев В. И., Куликов А. Ф., Пшенцов С. И. Влияние межфазной поверхностной энергии на процесс коагуляции микрочастиц при нагревании металлических сплавов // Поверхность. Физика, химия, механика. – 1985. – № 12. – С. 22–27.
6. Лифшиц И. М., Слезов В. В. О кинетике диффузионного распада пересыщенных твердых растворов // ЖЭТФ. – 1958. – Т. 35. – № 2/8. – С. 179–192.
7. Wagner C. Theorie der Alterung von Niederschlagen durch Umlösen (Ostwald – Reifung) // Zeitschr. f. Electrochem. – 1961. – V. 65. – No. 7/8. – P. 581–591.
8. Псарев В. И. Установление сходства между образом и подобием распределения микрочастиц по размерам в дисперсной системе // Изв. вузов. Физика. – 1991. – Т. 34. – № 12. – С. 80–84.
9. Псарев В. И. Проблема моментов распределений в статистической физике // Изв. вузов. Физика. – 1997. – Т. 40. – № 4. – С. 92–97.
10. Wynblatt P., Gjostein. A model study of catalyst particles coarsening // Scripta metal. – 1973. – V. 7. – No. 9. – P. 969–976.

Надійшла 16.11.07
Після доробки 5.12.07

Одержані аналітичні формули для проведення системного аналізу структурного огрублення островкових плівок на підкладці. Для цієї мети запропоновано метод встановлення подібності і відмінності між характеристиками теоретичного і експериментального розподілів островків за розмірами. Запропонований метод дозволяє одержувати інформацію про плин процесів в системі мікроостровків при її наближенні до рівноважного стану.

Analytical formulas for a systemic analysis of the structural coarsening of an island films on a solid substrate are obtained based on a comparison of the characteristics of theoretical and experimental distributions of the islands size. The proposed method provides useful information on the occurrence of various processes in a microislands system as it approaches the equilibrium state.

МАТЕМАТИЧНЕ ТА КОМП'ЮТЕРНЕ МОДЕЛЮВАННЯ

МАТЕМАТИЧЕСКОЕ И КОМПЬЮТЕРНОЕ МОДЕЛИРОВАНИЕ

MATHEMATICAL AND COMPUTER MODELLING

УДК 681.518:681.327.8

А. Ю. Берко, В. А. Висоцька

ЗАСТОСУВАННЯ МЕРЕЖ ПЕТРІ ДЛЯ МОДЕЛЮВАННЯ СИСТЕМ ЕЛЕКТРОННОЇ КОМЕРЦІЇ

У статті проаналізовано основні проблеми електронної комерції та запропоновано методи вирішення цих проблем.

ВСТУП. ЗАГАЛЬНА ПОСТАНОВКА ПРОБЛЕМИ

Двадцять перше століття кинуло людству виклик у формі всесвітньої «павутини» Internet та появи інтерактивного бізнесу чи віртуальної економіки. Інтерактивний бізнес – це бізнес, побудований на спільних діях бізнес-процесу в особі бізнесмена і комп'ютера або іншого засобу зв'язку обміну інформацією. Віртуальна економіка – це економіка, заснована на інтерактивному бізнесі та на головному законі людини – економії часу. Вже сьогодні у розвинутих державах існує багато випадків реалізації даного закону, а саме: використовуючи засоби зв'язку, не виходячи з дому, управляти технологічними лініями на виробництві або фінансово-комерційною діяльністю; вести бухгалтерський облік; здійснювати дистанційне навчання, читання

книг і періодики; купувати товари; виконувати банківські, біржові й інші фінансові операції. Нині Internet впливає як на зовнішні відносини між компаніями та їх партнерами чи клієнтами, так і на внутрішню структуру самих компаній. Електронна комерція заснована на структурі традиційної комерції, а використання електронних мереж додає їй гнучкості.

ЗВ'ЯЗОК ВИСВІТЛЕНОЇ ПРОБЛЕМИ ІЗ ВАЖЛИВИМИ НАУКОВИМИ ТА ПРАКТИЧНИМИ ЗАВДАННЯМИ

Електронна комерція (*electronic commerce*) – це придбання/продаж товару/послуги за допомогою електронних носіїв, чи через мережу, подібну до Internet. Дане поняття може включати в себе замовлення, оплату та доставку товарів/послуг. Складові електронної комерції (ЕК) представлені на рис. 1 [1]. Процеси, які становлять цикл ЕК [2]: доступ до інформації, оформлення замовлення, оплата, виконання замовлення,

© Берко А. Ю., Висоцька В. А., 2007



Рисунок 1 – Складові електронної комерції

післяпродажне обслуговування і підтримка. Типи ЕК [7]: торгівля інформацією, товарообіг, надання послуг.

Категорії ЕК [1]: бізнес – бізнес (наприклад, компанія, що використовує мережу для замовлень постачальникам, отримання рахунків і оплати); бізнес – споживач (електронна роздрібна торгівля); бізнес – адміністрація (операції, що укладаються між компаніями та урядовими організаціями); споживач – адміністрація (ще не існує, але із зростанням попередніх двох категорій уряди можуть розширити електронну взаємодію в таких сферах, як наприклад, соціальні виплати); споживач – споживач (взаємодія користувачів для обміну комерційною інформацією, досвідом, аукціонною торгівлею між фізичними особами тощо)

Інструменти, якими фірма користуватиметься для реалізації можливостей ЕК, можна умовно розділити на наступні групи [1]: бізнес-додатки; електронні магазини; шлюз в EDI-систему (electronic data interchange – технології електронного обміну даними); зв'язок із фінансовими організаціями через різні платіжні системи.

Сфери поширення ЕК [12]: маркетинг, продаж і сприяння продажу; перепродаж, попередні домовленості, поставки; фінансування й страхування; комерційні операції: замовлення, отримання, оплата; обслуговування та підтримка продукту; спільна розробка продукту; розподільне спільне виробництво; використання загальних і приватних послуг; адміністрування бізнесу (концесії, дозвіл, податки, митниця тощо); транспорт, техніка перевезень і постачання; загальні закупівлі; автоматична торгівля електронними товарами; бухгалтерський облік; дозвіл спірних моментів.

АНАЛІЗ СУЧАСНИХ ДОСЛІДЖЕНЬ І ПУБЛІКАЦІЙ

Особливості електронного ринку Internet полягають у наступному [12]: *відкритість* – доступ як для компаній будь-яких розмірів, так і для споживачів; *гло-*

бальність, тобто доступ до нього можливий з будь-якої точки земної кулі; *відвертість* ринку – характеризується досить низьким бар'єром для входу на нього фірм. Internet надає можливість скорочення каналів поширення товарів і ліквідації проміжних ланок, таких як дистриб'ютор і оптовий продавець, на зміну яким приходять прямий зв'язок виробник – покупець. Причиною такого скорочення є спроможність фірм взяти на себе функції, що традиційно виконуються фахівцями проміжних ланок, оскільки Internet володіє більш ефективною можливістю взаємодії зі споживачами й одночасно дозволяє відстежувати інформацію про споживачів. З технічного боку це зумовлено розвитком технологій побудови і ведення баз даних та автоматичного опрацювання запитів, які надходять.

Особливості Internet як засобу роздрібною торгівлі [7]: споживач взаємодіє з Internet і здійснює запити про купівлю безпосередньо через середовище Internet; споживач контролює інформацію про продукцію й послуги та одержує її в будь-який час доби; організація роздрібною торгівлі в Internet вимагає менших вкладень порівняно з організацією торгівлі традиційними способами; порівняно з торгівлею по телебаченню або торгівлею по каталогах споживачі можуть запитувати додаткову інформацію, достатню для здійснення купівлі, в тому числі в інтерактивному режимі.

Переваги ринку Internet [1–3]: активна позиція споживача; порівняно менший об'єм коштів, що вкладаються для виходу на ринок з боку фірм; глобальний характер ринку; цілодобова доступність ринку; можливість одержання необхідного обсягу інформації.

Для розвитку середовища Internet як ринку необхідне здійснення трьох основних умов [1]: завоювання фірмами, що використовують Internet як канал розподілу продукції, довіри з боку споживачів; забезпечення достовірності учасників та їх операцій; створення безпеки передачі і подальшого зберігання даних у середовищі Internet.

Методи завоювання довіри [1–3, 7]:

1. Збільшення гарантів, що переміщують ризик купівлі з покупця на продавця;
2. Значні знижки для перших клієнтів. Перша операція значно поліпшує довірчі відносини між покупцем і фірмою, що раніше майже були відсутніми;
3. Докладна інформація про фірму, включаючи її історію, філософію бізнесу і біографію, знаки схвалення від інших фірм, що засвідчують продавця, і т. п.

Не будучи єдиною технологією, електронна комерція характеризується різнобічністю. Вона об'єднує широкий спектр бізнес-операцій, включаючи [3, 5–7, 10–12]: встановлення контакту, наприклад, між потенційними замовником і постачальником; обмін інформацією; до- і післяпродажну підтримку (докладну інформацію про продукти і послуги, інструкції з використання продукту, відповіді на питання замов-

ників); продаж; електронну оплату (з використанням електронного переказу грошей, кредитних карток, електронних чеків, електронних грошей); управління доставкою та її відстеження для фізичних продуктів, безпосередню доставку продуктів, які можуть розповсюджуватися електронним шляхом; віртуальні підприємства – групи незалежних компаній, об'єднуючих свої зусилля для одержання можливостей надання продуктів і послуг, недоступних для окремих компаній; бізнес-процеси, що розділяються, спільно керовані компанією та її торговими партнерами.

Можливості і переваги електронної комерції [1, 3] вказано у табл. 1.

Таблиця 1

Можливості постачальників	Можливості замовників
Глобальна присутність	Глобальний вибір
Підвищення конкурентоспроможності	Якість послуг
Задоволення потреб замовника	Персоналізація товарів і послуг
Скорочення шляху товару до замовника	Швидка реакція на попит
Економія витрат	Зниження цін
Нові можливості ведення бізнесу	Нові продукти і послуги

Приклади конкретної комерційної вигоди від електронної комерції наступні [1, 7]:

1. Зменшення витрат на рекламу;
2. Зниження витрат на доставку, переважно для товарів, які можуть бути отримані електронним способом;
3. Скорочення витрат на дизайн та стратегічне планування;
4. Великі можливості для маркетингового дослідження ніш на ринку;
5. Однаковий доступ до ринку (для великих корпорацій і невеликих фірм);
6. Доступ до нових ринків збуту;
7. Залучення замовників у розробку і впровадження нових продуктів і послуг.

У числі найважливіших юридичних питань, що вимагають невідкладного рішення за участю світового співтовариства, варто назвати [1, 7, 12]: порядок оподаткування угод в електронній формі; тарифи; вимоги до форми угод і відповідальність; регулювання криптографії; правила аутентифікації; захист інформації; охорону прав споживачів.

Електронна комерція нині впливає на економіку і права громадян. Вимагаються такі правові норми, які сприяли б розвитку глобального та відкритого ринку шляхом уніфікації законодавства і спрощення правил і процедур, застосовуваних у різних країнах. У вироб-

лені подібних норм необхідне поглиблене співробітництво між бізнесом і державною владою, що забезпечило б істотну вигоду виробникам і споживачам в усьому світі. Сьогодні ж нерозвиненість у цілому або фрагментарність правових норм, що відносяться до сфери електронної комерції, а також значні суперечності між законодавствами різних країн є перешкодами для успішного функціонування *on-line economy*.

ВИДІЛЕННЯ ПРОБЛЕМ

Компанії досить часто обмінюються комерційною інформацією та проводять розрахунки між собою і клієнтами в електронній формі. Традиційно для цього використовуються системи на основі EDI (Electronic Data Interchange) та EFT (Electronic Financial Transfer) – стандартів, що накладають жорсткі умови на форми передаваної інформації. Застосовуються такі системи переважно великими компаніями, корпораціями, що мають приватні VAN-мережі [1–3]. На сьогоднішній день як фірми, так і приватні особи мають можливість здійснювати ділові операції через Internet без використання EDI.

Комерційну інформацію, що пересилається через Internet, можна розділити на дві категорії: інформаційні транзакції; фінансові транзакції.

Забезпечення інформацією – основний і дорогий елемент електронної комерції. Інформація в комерції може мати кілька форм:

- статистичні дані (цифри, графіки, аналізи);
- корпоративна інформація (телефонні номери, адреси, структура організації);
- інформація про продукцію або послуги;
- платна інформація (новини, періодичні видання, доступ до баз даних тощо).

Існує три основні класи фінансових транзакцій:

1. Компанія – компанія (платежі між банками, іншими фінансовими установами, перерахування коштів з рахунку однієї фірми на рахунок іншої здійснюються зазвичай електронним переказом, або за допомогою чеків);
2. Компанія – клієнт (основний спосіб платежу – готівка, чеки, дебетові та кредитні картки);
3. Клієнт – клієнт (основний спосіб платежу – готівка та чеки).

Використання Internet для виконання цих типів транзакцій дозволяє замінити представлення або показ готівки, чеків, кредитних карток їх електронними еквівалентами.

На сьогоднішній день існують спеціально розроблені електронні версії платіжних систем, за допомогою яких можна здійснити: замовлення та оплату покупки; банківські розрахунки; інвестиції (в цінні папери) тощо.

Електронна платіжна система – це авторизована інформаційна система, призначена для проведення роз-

рахунків в Internet між фінансовими, комерційними, виробничими, урядовими організаціями, а також окремими користувачами.

Серед основних вимог до платіжних систем можна назвати [1]:

1. Дотримання конфіденційності;
2. Збереження цілісності інформації;
3. Надання засобів оплати;
4. Забезпечення аутентифікації;
5. Проведення авторизації;
6. Мінімізація плати за транзакцію.

Платіжні систем можна прокласифікувати:

1. За типом платежу:
 - а) системи на основі кредитних карток;
 - б) системи на основі Internet-банкінгу;
 - в) системи на основі електронних чеків;
 - г) системи з використанням електронних грошей;
 - д) системи на основі smart-карт;
2. За схемою здійснення платежів:
 - а) кредитні;
 - б) дебетові;
3. За розповсюдженістю:
 - а) міжнародні;
 - б) національні.

Кожне віртуальне підприємство p_i самостійно обирає вид платіжної системи t_i , яка для нього є кращим варіантом проведення фінансових операцій.

Однак, не існує єдиної системи S проведення фінансових операцій між віртуальними підприємствами з різними платіжними системами у віртуальному світі електронного бізнесу, та не існує єдиної стандартної системи S контролю проведення фінансових транзакцій між учасниками віртуального світу електронного бізнесу.

ФОРМУВАННЯ ЦІЛЕЙ (ПОСТАНОВКА ЗАДАЧІ)

Систему фінансових взаємовідносин за допомогою платіжних систем (процесів) між віртуальними підприємствами (учасниками) у Internet-світі електронної комерції можна представити за допомогою мереж Петрі. Така система S буде складатися з чотирьох елементів [4–7, 8, 9]:

1. Скінченої множини учасників $P = \{p_1, p_2, \dots, p_n\}$ системи електронного ринку;
2. Множини процесів взаємовідносин $T = \{t_1, t_2, \dots, t_m\}$;
3. Вхідної функції $I: T \rightarrow P$, тобто направлених зв'язків між процесами та учасниками;
4. Вихідної функції $O: T \rightarrow P$ (направлені зв'язки між учасниками та процесами) (див. рис. 2).

Система проведення та контролю фінансових операцій між віртуальними підприємствами має наступний вигляд:

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

Потужність множини віртуальних підприємств $P - |P| = n$. Потужність множини платіжних систем $T - |T| = m$.

Учасник (віртуальне підприємство електронної комерції) $p_i \in$ вхідною позицією процесу t_i , якщо $p_i \in I(t_i)$ [4]. Аналогічно, $p_i \in$ вихідною позицією процесу t_i , якщо $p_i \in O(t_i)$. Входи і виходи процесів є комплектами позицій. Комплект є узагальненням множини. В комплект можуть включатися багаторазово повторені однакові елементи. Прикладом комплекту є $O(t_1) = \{p_2, p_3, p_4, p_4\}$.

Кратність вхідної позиції p_i для процесу t_j це число появ позиції у вхідному комплекті процесу $\#(p_i, I(t_j))$ [4–5]. Аналогічно, кратність вихідної позиції p_i для процесу t_j це кількість появ позиції у вихідному комплекті процесу $\#(p_i, O(t_j))$.

Розширена вхідна функція $I(p_i)$ визначається таким чином [8–9]:

$$\#(t_j, I(p_i)) = \#(p_i, O(t_j)). \quad (2)$$

Розширена вихідна функція $O(p_i)$ визначається таким чином [8–9]:

$$\#(t_j, O(p_i)) = \#(p_i, I(t_j)). \quad (3)$$

Графічним представленням системи фінансових взаємовідносин у віртуальному світі електронної комерції є двохдольний орієнтований мультиграф. Структура системи складається з сукупності учасників і процесів. Відповідно, граф системи складається з двох типів вузлів. Кругок O позначає учасника, а планка $|$ позначає процес. Отже, оскільки вершини графа можна розділити на дві множини – учасників і процеси, то граф є двохдольним.

Орієнтовані дуги (стрілки) з'єднують учасників і процеси, при цьому деякі дуги направлені від учасників до процесів, а деякі від процесів до учасників. Такі дуги позначають вхідні і вихідні комплекти. Кратні виходи представлені кратними дугами. Оскільки можуть існувати кратні дуги між вершинами, то маємо мультиграф. Дуги є направлені, тому такий мультиграф є орієнтований.

Таким чином, граф G системи фінансових взаємовідносин у віртуальному світі електронного бізнесу це двохдольний орієнтований мультиграф, $G = (V, A)$, де $V = \{v_1, v_2, \dots, v_s\}$ – множина вершин, $A = \{a_1, a_2, \dots, a_r\}$ – комплект направлених дуг, $a_i = (v_i, v_k)$, де $v_i, v_k \in V$. Множину V можна розбити на дві підмножини що не перетинаються – P і T , так що $V = P \cup T$, $P \cap T = \emptyset$ і для будь-якої дуги $a_i(v_i, v_k) \in V$ виконується одна з двох умов: або $v_i \in T, v_k \in P$, або $v_i \in P, v_k \in T$.

Комплект направлених дуг A , що відповідають даній системі визначається наступним чином [4]:

$$\#((p_i, t_j), A) = \#(p_i, I(t_j)); \quad (4)$$

$$\#((t_j, p_i), A) = \#(p_i, O(t_j)). \quad (5)$$

Двоїстим графом системи $C = (P, T, I, O)$ є граф системи $\bar{C} = (T, P, I, O)$, тобто граф, в якого учасники і процеси поміняні місцями. Такі графи не представляють практичного інтересу, тому ми їх не розглядаємо.

Маркування – присвоєння певної кількості фішок (грошових одиниць) учасникам системи взаємовідносин.

Маркування – це ресурси користувача, тобто ресурси віртуального підприємства електронної комерції, які він може використати для проведення операцій в системі фінансових взаємовідносин учасників електронної комерції.

Маркування μ системи $C = (P, T, I, O)$ є функцією, яка відтворює множину учасників P в множину невід'ємних цілих чисел N . Маркування можна також подати у вигляді n -мірного вектора $\mu = (\mu_1, \mu_2, \dots, \mu_n)$, де $n = |P|$, $\mu_i \in N$. Вектор μ визначає для кожного учасника p_i кількість фішок в цій позиції μ_i . Частіше застосовується позначення маркування в вигляді функції $\mu(p_i) = \mu_i$.

Маркована система $M = (C, \mu)$ – це сукупність структури системи $C = (P, T, I, O)$ і маркування μ . На графі фішки позначаються у вигляді точки всередині кружка, що позначає дану позицію учасника.

Виконання системи C – це перерозподіл фішок за допомогою запуску процесів. При запуску процесу фішки видаляються з його вхідних позицій учасників і додаються до його вихідних позицій учасників. Процес може запускатися лише в тому випадку, коли він дозволений. Перехід називається дозволим, якщо кожна з його вхідних позицій має число фішок не менше ніж число дуг з позиції в процесі. Фішки у вхідній позиції, що дозволяють процес називаються *дозволяючими*.

Процес $t_j \in T$ в маркованій мережі Петрі $C = (P, T, I, O)$ з маркуванням μ дозволений, якщо для всіх $p_i \in P$ виконується умова [4]

$$\mu(p_i) \geq \#(p_i, I(t_j)). \quad (6)$$

Процес запускається видаленням всіх дозволяючих фішок з його вхідних позицій і наступним додаванням фішок у кожен з його вихідних позицій по одній фішці для кожної дуги. Процес t_j в маркованій системі з маркуванням μ може бути запущений кожний раз, коли він дозволений. В результаті запуску дозволеного

переходу t_j утворюється нове маркування μ' , яке визначається таким співвідношенням [8]:

$$\mu'(p_i) = \mu(p_i) - \#(p_i, I(t_j)) + \#(p_i, O(t_j)). \quad (7)$$

Стан системи C визначається її маркуванням. Запуск будь-якого з процесів змінює стан системи шляхом зміни її маркування. Простір станів системи, що має n учасників це множина всіх маркувань, тобто N_n . Зміна стану системи викликана запуском процесу і визначається функцією δ , яка називається *функцією наступного стану*. Функція наступного стану δ для системи $C = (P, T, I, O)$ з маркуванням μ і процесом $t_j \in P$ визначена тоді і лише тоді, коли $\mu(p_i) \geq \#(p_i, I(t_j))$ для всіх $p_i \in P$. Якщо $\mu(p_i)$ визначена, то $\delta(\mu, t_j) = \mu'$, де $\mu'(p_i) = \mu(p_i) - \#(p_i, I(t_j)) + \#(p_i, O(t_j))$ для всіх $p_i \in P$ [9].

Нехай дана система $C = (P, T, I, O)$ з початковим маркуванням μ^0 . Ця мережа може бути виконана послідовними запусками процесів. Запуск дозволеного процесу t_j призведе до утворення нового маркування $\mu^1 = \delta(\mu^0, t_j)$. В отриманому маркованому графі можна запустити будь-який інший дозволений процес, наприклад t_k . Отримуємо маркування $\mu^2 = \delta(\mu^1, t_k)$. Цей процес можна продовжувати до того часу, доки дозволений хоча б один процес.

При виконанні системи отримуємо дві послідовності. Послідовність маркувань $(\mu^0, \mu^1, \mu^2, \dots)$ і послідовність процесів які запускаються $(t_{j1}, t_{j2}, t_{j3}, \dots)$. Ці послідовності зв'язані таким співвідношенням [4, 8]:

$$\delta(\mu_k, t_{jk}) = \mu_{k+1}, \quad k = 0, 1, 2, \dots \quad (8)$$

За допомогою цього співвідношення, маючи структуру системи $C = (P, T, I, O)$, початкове маркування μ^0 та послідовність станів $(t_{j1}, t_{j2}, t_{j3}, \dots)$ можемо отримати послідовність $(\mu^0, \mu^1, \mu^2, \dots)$. Зворотній процес від послідовності маркувань до послідовності процесів можливий у більшості випадків, за виключенням деяких вироджених випадків.

Для системи $C = (P, T, I, O)$ з маркуванням μ маркування μ' називається *безпосередньо досяжним* з μ , якщо існує процес $t_j \in T$, такий, що $\delta(\mu, t_j) = \mu'$, тобто μ' можна отримати з μ запуском одного з дозволених процесів.

Множина досяжності $R(C, \mu)$ для системи $C = (P, T, I, O)$ з маркуванням μ – це найменша множина маркувань, визначених таким чином:

1. $\mu \in R$;

2. Якщо $\mu' \in R$, то для всіх $\mu'' = \delta(\mu', t_j)$, $t_j \in T$ виконується $\mu'' \in R$, тобто якщо якийсь маркування належить множині досяжності, то всі досяжні з нього маркування також належать їй.

Часто користуються поняттям розширеної функції наступного стану. Розширена функція вхідного стану

визначається для маркування μ та послідовності процесів $\sigma = (t_{j_1}, t_{j_2}, \dots, t_{j_n})$ таким правилом [4, 12]:

$$\mu' = \delta(\mu, \sigma) = \delta(\dots\delta(\delta(\mu, t_{j_1}), t_{j_2})\dots, t_{j_n}), \quad (9)$$

тобто спочатку запускаємо процес t_{j_1} , потім t_{j_2} і так далі, до t_{j_n} .

Дерево досяжності системи S є ілюстрацією множини досяжності $R(S, \mu)$. Оскільки в багатьох випадках множина досяжності є нескінченною, то існують певні правила, які дозволяють відобразити її скінченим деревом досяжності.

Для приведення дерева досяжності до скінченного представлення використовуються такі правила [8–9].

1. Пасивні маркування – тобто маркування в яких немає дозволених процесів є *термінальними вершинами* дерева досяжності (не мають власних піддерев досяжності).

2. Якщо отримане маркування вже зустрічалось в дереві досяжності раніше, то такі маркування називають *дублюючими вершинами*. Розгляд дублюючих вершин зайвий, оскільки отримане піддерево буде аналогічне.

3. Якщо в дереві досяжності отримуємо маркування, яке є більшим за одне з попередніх (на шляху до кореня) маркувань, то у всіх позиціях проставляємо символ ω .

За допомогою цих трьох правил, множину досяжності будь-якої системи можна подати у вигляді скінченого дерева.

АНАЛІЗ ОТРИМАНИХ НАУКОВИХ РЕЗУЛЬТАТІВ

Для прикладу побудуємо систему фінансових взаємовідносин між наступними учасниками: користувач-клієнт (p_1), Інтернет-аукціон (p_2), банк клієнта та Інтернет-аукціону (p_3), Інтернет-магазин (p_4), банк Інтернет-магазину (p_5).

Аналогічно, маємо наступні процеси: оплата за товар/послугу з боку клієнта (t_1), фінансовий процес взаємодії між Інтернет-аукціоном, Інтернет-магазином та банком клієнта (t_2), фінансові операції між Інтернет-магазином та його банком (t_3), переказ грошей на рівні банків та Інтернет-магазину (t_4, t_5).

Для наведеного вище прикладу системи розширена вхідна і вихідна функції будуть виглядати таким чином:

$$\begin{aligned} I(p_1) &= \{\}; & O(p_1) &= \{t_1\}; \\ I(p_2) &= \{t_1, t_2\}; & O(p_2) &= \{t_2\}; \\ I(p_3) &= \{t_1, t_4\}; & O(p_3) &= \{t_2, t_5\}; \\ I(p_4) &= \{t_1, t_1, t_4\}; & O(p_4) &= \{t_2, t_3, t_3\}; \\ I(p_5) &= \{t_3, t_5\}; & O(p_5) &= \{t_4\}. \end{aligned}$$

На рис. 3. подано граф системи з прикладу на рис. 2 з маркуванням $\mu = (1, 2, 4, 8, 15)$. На рис. 3. фішка в позиції p_1 є дозволяючою для процесу t_1 . Приклади запуску процесів системи S подано на рис. 4. На рис. 4. процес t_4 не може бути запущений.

Для прикладу розглянемо систему, зображену на рис. 3. Її початкове маркування $\mu^0 = (1, 1, 3, 2, 1)$ буде

$C = (P, T, I, O);$	$I(t_1) = \{p_1\};$	$O(t_1) = \{p_2, p_3, p_4, p_4\};$
$P = \{p_1, p_2, p_3, p_4, p_5\};$	$I(t_2) = \{p_2, p_3, p_4\};$	$O(t_2) = \{p_2\};$
$T = \{t_1, t_2, t_3, t_4, t_5\};$	$I(t_3) = \{p_4, p_4\};$	$O(t_3) = \{p_5\};$
	$I(t_4) = \{p_5\};$	$O(t_4) = \{p_3, p_4\};$
	$I(t_5) = \{p_3\};$	$O(t_5) = \{p_3\};$

Рисунок 2 – Приклад системи S у віртуальному світі електронної комерції

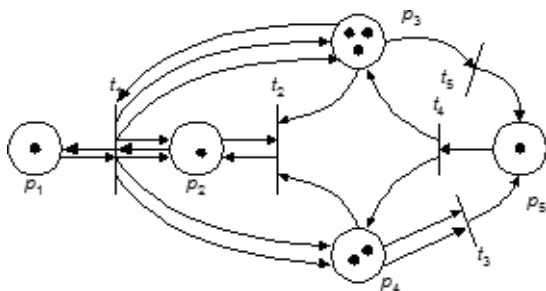


Рисунок 3 – Приклад графу маркованої системи S , представлена мережею Петрі

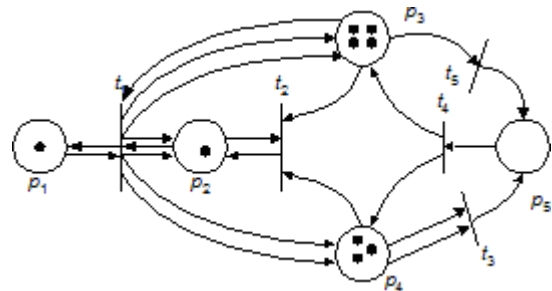


Рисунок 4 – Маркування отримане в результаті запуску переходу t_4 графу зображеного на рис. 3

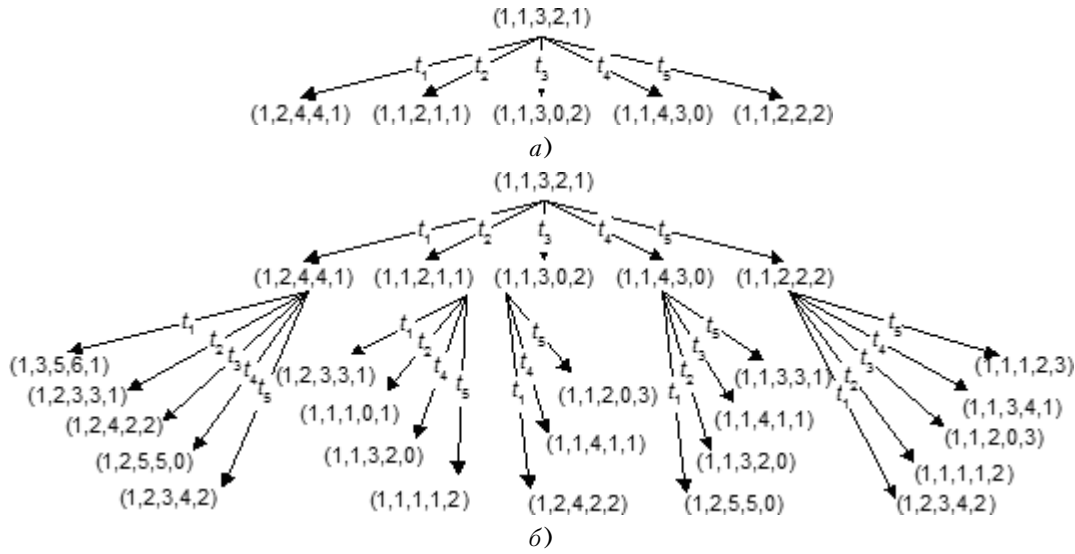


Рисунок 5 – Побудова дерева досяжності:
а – перший крок; б – другий крок

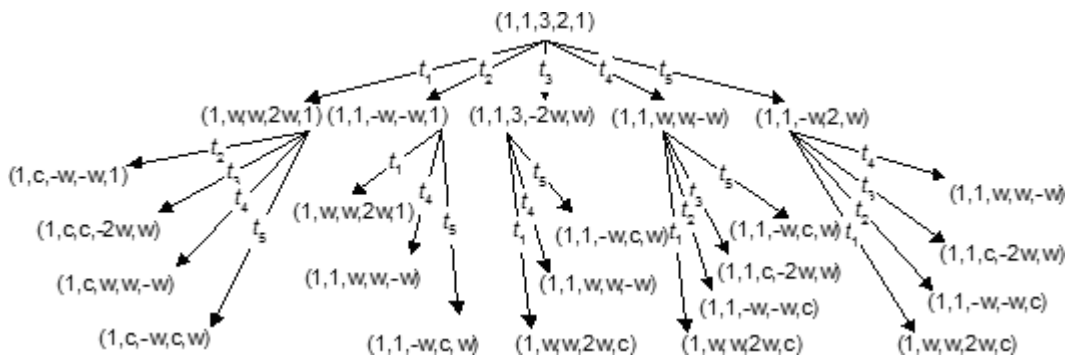


Рисунок 6 – Приведене дерево досяжності

початковою вершиною (коренем) дерева досяжності. В μ^0 дозволені два процеси – $t_1 - t_5$. Нові вершини дерева досяжності будуються для всіх безпосередньо досяжних з μ^0 маркувань і до них з початкової вершини ведуть дуги помічені процесом, що запускається. В результаті запуску t_1 отримаємо $\mu_1^1 = (1, 2, 4, 4, 1)$, а в результаті запуску t_2 отримаємо $\mu_2^1 = (1, 1, 2, 1, 1)$. Таким чином, перший крок побудови дерева досяжності буде мати вигляд, зображений на рис. 5. Аналогічно, здійснюються наступні кроки.

Розглянемо послідовність запусків $t_1 t_1 t_1$. Кожен раз, отримане маркування відрізняється від попереднього лише кількістю фішок в p_2-p_3 (зростає на одиницю) та в p_4 (зростає на дві одиниці). Запускаючи процес t_1 багато раз можна отримати яке завгодно велике число фішок в ньому. Так як запуск процесу збільшує кількість фішок в одній (декількох) позиції залишаючи незмінною кількість фішок в інших, то маркування позиції (позицій), в якій кількість фішок

збільшилось, прийнято позначати символом ω . Символом c позначимо стали кількість фішок в позиції маркування, яка не змінюється на n кроці побудови дерева досяжності. Можна розширити це поняття на послідовність процесів σ , яка збільшує кількість фішок в якійсь позиції, не змінюючи інші позиції. Таким чином, якщо в дереві досяжності отримуємо маркування, яке є більшим за одне з попередніх (на шляху до кореня) маркувань, то у всіх позиціях, що є більшими, проставляємо символ ω (рис. 6).

ВИСНОВКИ І ПЕРСПЕКТИВИ ПОДАЛЬШИХ НАУКОВИХ РОЗВІДОК

Пов'язані з електронною комерцією правові питання спроможні сильно вплинути на її розвиток, проте вони ще чекають свого вирішення. Вони стосуються конфіденційності, оподаткування електронної комерції і кон-

тролю за экспортом криптографічних технологій та правових фінансових взаємовідносин між учасниками віртуального світу електронної комерції. Незалежно від того, як працює компанія, електронна трансформація вже не є предметом вибору. Це вже є необхідністю для успішного розвитку бізнесу.

ПЕРЕЛІК ПОСИЛАНЬ

1. Береза А. М. Електронна комерція. – Київ, 2002.
2. Галіцин В. К., Левченко Ф. А. Обчислювальні системи та мережі. – К.: КНЕУ, 1997.
3. Джерк Н. Разработка приложений для электронной коммерции. – СПб: Питер, 2001.
4. Катренко А. В. Системний аналіз об'єктів та процесів комп'ютеризації. – Львів: «Новий світ – 2000», 2003. – Стор. 286–322.
5. Катренко А. В. Системні аспекти розвитку архітектури підприємства // Інформаційні системи та мережі. – 2002. – № 464. – С. 123–131.

6. Козье Девид. Электронная коммерция. – Москва: Русская Редакция, 1999.
7. Крупник А. Бизнес в интернет. – Москва: Микроарт, 2002.
8. Питерсон Дж. Теория сетей Петри и моделирование систем. – М.: Мир, 1984.
9. Советов Б. Я. Яковлев С. А. Моделирование систем. – М.: ВШ, 1985.
10. Успенский И. Энциклопедия Интернет-бизнеса. – СПб.: Питер, 2001.
11. Холмогоров В. Интернетмаркетинг. – СПб.: Питер, 2001.
12. Эймор Дэниел. Электронный бизнес. – Москва: Вильямс, 1999.

Надійшла 9.02.07

В статье проанализированы основные проблемы электронной коммерции и предложены методы решения этих проблем.

In the given article main problems of electronically commercial are analyzed. New methods for solution of discussed problems are proposed.

УДК 519.2:368.01

С. Н. Герасин, М. А. Козлов

СТОХАСТИЧЕСКАЯ МОДЕЛЬ ИЗМЕНЯЮЩИХСЯ МЕЖПРОИЗВОДСТВЕННЫХ ПОТОКОВ

Предложена стохастическая интерпретация модели динамического распределения ресурсов. Формализм модели использует язык неоднородных марковских цепей. Показана возможность использования предложенной модели в задачах экономического прогнозирования.

ВВЕДЕНИЕ

Предлагаемая модель динамического распределения межпроизводственных потоков (в денежном выражении) отражает взаимосвязь между структурой производства и потребления (на уровне региона или в целом по народному хозяйству) и изменением дохода отдельных предприятий (отраслей) и суммарного дохода по региону в целом (валового национального дохода). В модели используется известный принцип представления структуры производства-потребления в виде матрицы коэффициентов затрат [1] замкнутой системы, а также метод отражения зависимости между объемом дохода и структурой производства-потребления с помощью матрицы коэффициентов прибыльности вложений [2].

Модель позволяет:

1) осуществить выбор оптимальной, с точки зрения роста суммарного дохода, структуры производства-потребления;

2) осуществить выбор оптимальной структуры производства для того или иного предприятия (той или иной отрасли), с точки зрения роста его (ее) доходов, при заданной структуре производства других предприятий региона (других отраслей народного хозяйства);

3) осуществить прогноз изменения суммарного дохода и (или) дохода той или иной отрасли при данной структуре производства-потребления.

ОПИСАНИЕ МОДЕЛИ

Введем некоторые обозначения:

$V(t)$ – суммарный доход (валовой национальный продукт) на момент t ;

$p(t) = (p_1(t), p_2(t), \dots, p_n(t))$ – вектор коэффициентов распределения суммарного дохода (валового национального продукта) по предприятиям региона (отраслям народного хозяйства):

$V_j(t) = v(t)p_j(t)$ – доход j -го предприятия (j -й отрасли) на момент t ;

$P(t) = \|p_{ij}(t)\|_{i,j=1}^n$ – матрица коэффициентов затрат, отражающая структуру производства-потребления:

$p_{ij}(t)$ – часть дохода i -го предприятия (i -й отрасли), направляемая на приобретение продукта j -го предприятия (j -й отрасли),

$$p(t+1) = p(t)P(t);$$

$R_{ij}(t) = \|r_{ij}(t)\|_{i,j=1}^n$ – матрица коэффициентов прибыльности вложений на момент t :

$r_{ij}(t)$ – величина прибыли, получаемой i -м предприятием (отраслью) на единицу затрат по приобретению продукции j -го предприятия (отрасли), на момент t ;

$k_i(t) = \sum_{j=1}^n r_{ij}(t)p_{ij}(t)$ – совокупная прибыль, получаемая i -м предприятием (отраслью) на единицу затрат, произведенных на момент t ;

$k(t) = \sum_{i=1}^n p_i(t)k_i(t) = \sum_{i=1}^n \sum_{j=1}^n p_i(t)p_{ij}(t)r_{ij}(t)$ – величина относительного изменения (увеличения при $k(t) > 0$ и уменьшения при $k(t) < 0$) суммарного дохода по региону (валового национального продукта) на момент t при данной структуре производства:

$$V(t+1) = (1 + k(t))V(t),$$

$k(t) \geq -1$ (получаемый доход не может быть отрицательным).

Как отмечалось выше, предлагаемая модель является моделью замкнутой системы. На практике, однако, ни одна экономическая система не является замкнутой: всегда присутствует товарообмен предприятий данного региона или страны с внешними контрагентами. Для преодоления этого противоречия модель предусматривает введение дополнительного фиктивного предприятия, представляющего в агрегированной форме всех внешних контрагентов. Поскольку в реальной практике одним из контрагентов рынка всегда является государство или отдельные его органы, то в модель в качестве еще одного фиктивного предприятия введен элемент, представляющий государство в структуре производства-потребления. Наконец, поскольку источником всякой стоимости является живой труд, а конечным потребителем продукции всякого предприятия (отрасли) является население, то в модели также использовано агрегированное представление последнего как еще одного фиктивного предприятия (отрасли). На основании вышеизложенного для отдельных элементов матрицы коэффициентов затрат $P(t) = \|p_{ij}(t)\|_{i,j=1}^n$ можно предложить следующие интерпретации.

$p_{i1}(t)$ – часть дохода i -го предприятия, полученного на момент t , расходуемая на выплату заработной платы.

$p_{i2}(t)$ – часть дохода i -го предприятия, полученного на момент t , расходуемая на обязательные отчисления

в государственный и местный бюджеты, а также в различные фонды.

$p_{in}(t)$ – часть дохода i -го предприятия, полученного на момент t , расходуемая на закупку продукции внешних контрагентов.

$p_{1i}(t)$ – часть доходов населения, полученных на момент t , расходуемая на приобретение продукции i -го предприятия.

$p_{2i}(t)$ – часть государственных доходов, полученных на момент t , расходуемая на

- а) закупку продукции i -го предприятия;
- б) государственные инвестиции в i -е предприятие.

$p_{21}(t)$ – часть государственных доходов, полученных на момент t , расходуемая на выплаты населению (зарплаты в бюджетной сфере, пенсии, стипендии, пособия, материальная помощь населению).

$p_{12}(t)$ – часть доходов населения, полученных на момент t , расходуемая на уплату налогов (подоходного, на недвижимое имущество, на транспортные средства).

$p_{2n}(t)$ – часть государственных доходов, полученных на момент t , расходуемая на осуществление платежей внешним контрагентам (закупка продукции внешних производителей, платежи по кредитам из внешних источников, кредитование внешних агентов, безвозмездная помощь внешним субъектам).

$p_{n2}(t)$ – часть доходов, полученных на момент t полученных от внешних контрагентов, получаемая государством (таможенные пошлины, кредиты государству из внешних источников, платежи по кредитам государства внешним агентам).

$p_{1n}(t)$ – часть доходов населения, полученных на момент t , расходуемая на выплаты внешним контрагентам (денежные переводы за пределы региона, расходы на туристические поездки, приобретение продукции непосредственно у внешних производителей).

$p_{n1}(t)$ – часть доходов, полученных на момент t полученных от внешних контрагентов, получаемая населением (денежные переводы из-за пределов региона, заработная плата, полученная за работу на внешних контрагентов и т. п.).

При определении коэффициентов прибыльности вложений для каждой отрасли следует исходить из следующих принципов:

1) поскольку источником всякой стоимости является живой труд, то расходы на заработную плату всегда прибыльны: $r_{i1}(t) > 0, 2 < i < n, \forall t > 0$;

2) отдача от живого труда обычно пропорциональна энерго- и технической вооруженности производства, поэтому для j , соответствующих расходам на энергообеспечение и закупку оборудования также следует положить $r_{ij}(t) > 0, 2 < i < n, \forall t > 0$;

3) затраты на приобретение сырья обычно сами по себе не приводят ни к увеличению, ни к уменьшению

дохода (при нормальных пропорциях относительно других статей расходов);

4) расходы по уплате государственных сборов и налогов обычно непосредственно не продуктивны, поэтому для всех i следует положить $r_{i2}(t) \leq 0$.

При определении конкретных значений этих величин следует учитывать, что они в большой степени зависят от пропорций расходов (то есть реально имеем $r_{ij}(t) = r_{ij}(p_{i1}(t), p_{i2}(t), \dots, p_{in}(t))$), но в любом случае для любого i и j и всех $t > 0$ должно выполняться неравенство $r_{ij}(t) \geq -1$ (то есть даже при самом неразумном вложении средств нельзя потерять больше, чем их имеется в наличии).

СТОХАСТИЧЕСКАЯ ИНТЕРПРЕТАЦИЯ МОДЕЛИ

В классической модели коэффициентов затрат [1] предполагалось, что каждое предприятие (отрасль) на каждом шаге полностью расходует весь полученный на этот момент доход – такое предположение было оправданным на том основании, что при представлении взаимосвязей между отдельными производствами посредством статической модели (охватывающей большие промежутки времени порядка года) такое допущение действительно приближенно соответствовало реальному положению вещей. В предлагаемой модели, так как она, во-первых, является динамической, а во-вторых, поскольку ее временной масштаб имеет порядок недель или месяцев, более оправданным является предположение, что в матрицах коэффициентов затрат для различных периодов не все диагональные элементы будут равны нулю. Следует также отметить, что данное предположение, также как и предположение о неоднородности цепи Маркова с матрицами переходных вероятностей за единицу времени, совпадающими с матрицами коэффициентов затрат для отдельных периодов в рассматриваемой модели, обеспечивает отражение в этой модели различной длительности производственных циклов для различных производств. Кроме того, модель можно модифицировать таким образом, чтобы в ней учитывалось изменение числа предприятий с течением времени: для этого можно использовать рассматривавшуюся в [3] модель цепей Маркова с переменным числом состояний.

В соответствии с принятыми в данной модели предположениями, для величины дохода $V(k, T)$, полученного с момента s в течении промежутка времени T имеем

$$\begin{aligned} V(s, T) &= \sum_{t=0}^T \beta^t \sum_{i=0}^n p_i(s+t-1) \sum_{j=1}^n p_{ij}(s+t) r_{ij}(s+t) = \\ &= \sum_{t=0}^T \beta^t \sum_{i=1}^n p_i(s+t-1) k_i(s+t), \end{aligned} \quad (1)$$

где β – коэффициент дисконтирования (величина, обратная норме банковского процента).

При $T \rightarrow \infty$ из (1) получаем

$$V(s, \infty) = \sum_{t=0}^{\infty} \beta^t \sum_{i=1}^n p_i(s+t-1) k_i(s+t). \quad (2)$$

В случае, когда неоднородная цепь Маркова с матрицами переходных вероятностей за единицу времени, совпадающими с соответствующими матрицами коэффициентов затрат, является эргодической (в сильном смысле), для $V(s, \infty)$ можно получить оценку $\tilde{V}(s, \infty)$, которая включает в себя суммирование только конечного числа слагаемых. Рассмотрим вначале величину

$$\begin{aligned} V'(s, \infty) &= \sum_{t=0}^N \beta^t \sum_{i=1}^n p_i(s, t-1) k_i(s+t) + \\ &+ \sum_{t=N+1}^{\infty} \beta^t \sum_{i=1}^n p_i^* q_i(s+t) = \sum_{t=0}^N \beta^t \sum_{i=1}^n p_i(s, t-1) k_i(s+t) + \\ &+ \sum_{i=1}^n p_i^* \sum_{t=N+1}^{\infty} \beta^t q_i(s+t), \end{aligned}$$

где p_i^* , $1 \leq i \leq n$ – компоненты вектора предельного распределения p^* соответствующей неоднородной цепи Маркова; N – число переходов, достаточное либо для сходимости распределения указанной цепи к предельному, так называемый процесс фокусировки (условия, которые для этого должны выполняться, приведены в [4]), либо для того, чтобы распределение цепи достаточно мало отличалось от своего предела (σ -фокусировка [5]).

В общем случае, использование величины $V'(s, \infty)$ вместо $V(s, \infty)$ не приводит к существенному упрощению расчетов, однако в некоторых частных случаях на основе этой величины можно получить оценки величины $V(s, \infty)$, достаточно приемлемые в смысле точности, получение которых гораздо менее трудоемко с вычислительной точки зрения.

Пусть $r_{ij}(t) \equiv r_{ij} = \text{const}$. В этом случае имеем

$$\begin{aligned} \left| \sum_{t=N+1}^{\infty} \beta^t \sum_{j=1}^n p_{ij}(s+t) r_{ij} \right| &\leq \sum_{t=N+1}^{\infty} \beta^t \sum_{j=1}^n p_{ij}(s+t) |r_{ij}| = \\ &= \sum_{j=1}^n |r_{ij}| \sum_{t=N+1}^{\infty} p_{ij}(s+t) \beta^t \leq R_i \sum_{t=N+1}^{\infty} \beta^t = \\ &= R_i \frac{\beta^{N+1}}{1-\beta}, \end{aligned} \quad (3)$$

где $R_i = \sum_{j=1}^n |r_{ij}| < \infty$, $1 \leq i \leq n$.

Для всякого $\varepsilon > 0$ при достаточно больших N будет справедливо неравенство

$$\frac{\beta^{N+1}}{1-\beta} < \frac{\varepsilon}{R},$$

где $R = \max_{1 \leq i \leq n} R_i$. Тогда для выражения в правой части (3) имеем

$$R_j \frac{\beta^{N+1}}{1-\beta} < \varepsilon.$$

То есть при подходящих значениях β в качестве оценки $\tilde{V}(s, \infty)$ величины $V(s, \infty)$ можно принять

$$\tilde{V}(s, \infty) = \sum_{t=0}^N \beta^t \sum_{i=1}^n p_i(s, t-1) k_i(s+t). \quad (4)$$

Нетрудно заметить, что оценка (4) является удовлетворительной только при специальных значениях параметра β , что значительно сужает область ее применимости. Гораздо больший интерес представляет поэтому случай, когда отдача от вложений зависит только от направления расходов, но не от их источника (такое предположение будет оправдано в том случае, когда модель охватывает предприятия, которые делятся на две группы в зависимости от профиля – производителей и дистрибьютеров однотипной продукции): $r_{ij}(t) = r_j(t)$. В этом случае, подставляя в (1) выражение для $r_{ij}(t)$, получим

$$\begin{aligned} V(s, T) &= \sum_{t=0}^T \beta^t \sum_{i=1}^n p_i(s+t-1) \sum_{j=1}^n p_{ij}(s+t) r_{ij}(s+t) = \\ &= \sum_{t=0}^T \beta^t \sum_{i=1}^n p_i(s+t-1) \sum_{j=1}^n p_{ij}(s+t) r_j(s+t) = \\ &= \sum_{t=0}^T \beta^t \sum_{i=1}^n p_i(s+t) r_i(s+t). \end{aligned}$$

Таким образом, при достаточно больших s можно использовать оценку

$$\tilde{V}(s, T) = \sum_{t=0}^T \beta^t \sum_{i=1}^n p_i^* r_i(s+t). \quad (5)$$

Хотя такая оценка также требует выполнения весьма специфических условий, ее все же можно использовать и в ситуациях, когда предположение, на котором она основана само выполняется лишь приближенно. Примером такого «расширения» модели является модель, для которой имеют место соотношения

$$r_{ij}(t) = r_j(t) + \alpha_t p_{ij}(t),$$

$$\lim_{t \rightarrow \infty} \alpha_t = 0, \lim_{t \rightarrow \infty} \frac{\alpha_t}{\beta^{mt}} = 0 \forall m > 0 (m < \infty).$$

Очевидно, что при выполнении этих условий и при достаточно больших s , для таких моделей применима оценка (5).

При программной реализации в качестве исходных данных используются:

- 1) сведения о числе предприятий региона (или о среднем числе предприятий за достаточно продолжительный период, если выбрана модель с меняющимся числом предприятий);
- 2) сведения из балансовых отчетов предприятий о доходах и расходах за отчетный период;
- 3) сведения из бюджетной сметы о доходах и расходах бюджета за отчетный период;
- 4) статистические сводки о величине совокупного дохода по региону;
- 5) материалы социологических исследований о структуре доходов и расходов населения.

На основании этих данных строятся матрицы коэффициентов затрат и коэффициентов прибыльности вложений.

При этом по желанию пользователя, можно получить:

- 1) прогноз изменения совокупного дохода по региону с заданным упреждением;
- 2) прогноз изменения дохода того или иного предприятия при заданном упреждении;
- 3) оценку долгосрочных тенденций в экономике региона.

На основании этой информации можно:

- 1) оценить влияние изменений в технологии на доход предприятия;
- 2) осуществить выбор оптимальной технологии из ряда альтернатив;
- 3) осуществить выбор оптимального решения при инвестировании в экономику региона.

ВЫВОДЫ

В заключении следует отметить, что указанная модель представляет собой модель марковского типа не только по форме представления данных – ее действительно можно интерпретировать либо как цепь Маркова (с постоянным или переменным числом состояний), либо как полумарковский процесс, описывающий эволюцию некоторой абстрактной «неделимой» денежной единицы, переходящей от одного предприятия к другому.

1. Аллен Р. Математическая экономия. – М.: Иниздат, 1963. – 648 с.
2. Валтер Я. Стохастические модели в экономике. – М.: Статистика, 1976. – 232 с.
3. Герасин С. Н. Стабилизация распределений марковских цепей с переменным числом состояний за конечное время // Доповіді НАН України. – 2004. – № 12. – С. 59–63.
4. Герасин С. Н., Дикарев В. А., Числин Н. И. Сходимость к предельному распределению в конечных неоднородных марковских процессах за конечное время // Доповіді НАН України. – 1998. – № 7. – С. 15–19.
5. Герасин С. Н. Стабилизация решений в задачах динамического распределения ресурсов // Доповіді НАН України. – 2001. – № 10. – С. 73–78.

Запропонована стохастична інтерпретація моделі динамічного розподілу ресурсів. Формалізм моделі використовує мову неоднорідних марківських ланцюгів. Показана можливість використання запропонованої моделі в розв'язанні задач економічного прогнозування.

Stochastic interpretation of model of dynamic allocation of resources is offered. Model formalism is utilized by the language of inhomogeneous Markov's chains. Possibility of the use of the offered model is solution in the tasks of economic prognostication.

УДК 004.932.2

В. А. Гороховатский, С. В. Кузьмин

СТАТИСТИЧЕСКИЙ АНАЛИЗ ЗНАЧЕНИЙ ХАРАКТЕРНЫХ ПРИЗНАКОВ ПРИ РАСПОЗНАВАНИИ ИЗОБРАЖЕНИЙ

Приведены результаты исследований по применению статистического подхода при формировании, оценке значений, анализе свойств и установлении эквивалентности характерных признаков изображений в целях структурного распознавания объектов. Экспериментальные оценки подтверждают возможность построения оптимальных локальных решений.

ВВЕДЕНИЕ

Структурные методы распознавания изображений, основанные на построении характерных признаков фрагментов, имеют несомненные преимущества перед традиционными интегральными подходами в плане гибкости, универсальности, а также устойчивости к фоновым и локальным помехам [1, 2]. Характерными признаками (ХП) (другие названия: точки интереса, ключевые точки) считают значимые отклики локальных фильтров изображений [2–4]. ХП относятся к структурным признакам низкого уровня, по которым либо непосредственно осуществляется распознавание, либо на их основе формируют признаки более высокого уровня. Как правило, ХП получают в результате двухэтапной процедуры, где вначале вычисляется значение признака, а затем принимается решение о значимости и об участии в распознавании. Оптимальные отклики локальных фильтров, на основе которых формируются ХП, описывают конечный набор свойств сигнала изображения в фиксированной локальной окрестности (фрагменте).

Возможность снижать влияние локальных и фоновых помех заложена непосредственно в структурном

подходе к анализу изображения. В то же время действие случайных помех, вызванных условиями формирования изображения, в частности, влиянием дискретизации, геометрическими преобразованиями объектов и цифровыми методами вычисления инвариантных значений ХП, требует дополнительного исследования. Эти помехи носят характер шума и непосредственно сказываются на точности измерения значений признаков и, в конечном итоге, на результатах распознавания.

Цель работы – применение статистического анализа для оценки параметров моделей формирования, оценки значимости и установления эквивалентности значений ХП. Задачами исследования есть изучение распределений, свойств ХП, а также построение достоверных оценок порогов при принятии решения о значении или эквивалентности признаков.

1 СТАТИСТИЧЕСКИЕ МОДЕЛИ ПОСТРОЕНИЯ И ОПИСАНИЯ ХП

Образ объекта представляет собой упорядоченную совокупность признаков. Класс – совокупность образов, обладающих некоторыми общими свойствами [5–8]. Одна из форм описания ХП – это n -мерные числовые векторы $v = (v_1, \dots, v_n)$, $v_i \in R$, $i = \overline{1, n}$ (R – пространство действительных чисел). При решении практических задач часто считают, что значения векторов имеют нормальное распределение с известными параметрами: математическим ожиданием $\mu = (\mu_1, \dots, \mu_n)$ и матрицей

ковариации признаков $\Sigma = \{\sigma_{ij}\}$ (или среднеквадратическим отклонением $\sigma = (\sigma_1, \dots, \sigma_n)$). Информация о типе и пространственном расположении ХП в структурном представлении объекта содержится в векторе $\beta = \{x, y, t\}$, где x, y – координаты, t – тип ХП.

В конкретной реализации имеем множества $V = \{v^k\}$, $\{\beta^k\} = \{x_k, y_k, t_k\}$, $k = \overline{1, s}$, s – сформированное количество ХП. Заметим, что величина t и компоненты вектора $v = (v_1, \dots, v_n)$ должны быть инвариантами относительно допустимых геометрических преобразований фрагмента.

Наиболее простым в вычислительном плане есть принятие решения по одному признаку, т. е. вариант $n = 1$. Здесь признак имеет нормальное распределение $N(\mu, \sigma^2)$ и его среднее μ должно различаться для разных классов при относительно небольшой дисперсии. Примером может быть величина среднего значения $v_L = \frac{1}{L} \sum_{x, y \in D} B(x, y)$, где $B(x, y)$ – функция яркости, L – количество точек во фрагменте D . Признак v_L компактно представляет фрагмент в структурном методе частных корреляций вместо значений функции яркости [7].

Формирование набора ХП, а также отбор среди них множества значимых ХП, – это предмет отдельных процедур. Первая из них Ω_1 осуществляет отображение $\Omega_1: M \rightarrow V$ из множества функций яркости $M = \{B(x, y)\}$ во множество значений признаков $V \subset R^n$. Второе $\Omega: V \rightarrow O$ – формирует значимый набор из s описаний $O = \{o_1, \dots, o_s\}$, $o_i = \{v^i, \beta^i\}$, $i = \overline{1, s}$.

Основными способами построения отображений Ω_1, Ω_2 являются: 1) сопоставление фрагмента с шаблонами (корреляция); 2) свертка с маской, подчеркивающей наличие признака во фрагменте [4]; 3) анализ градиентов функции яркости.

Удобным и универсальным инструментом анализа низкоуровневых признаков считается структурный тензор, который отображает фрагмент в значение некоторой случайной величины [2, 4]. Построение тензора базируется на дифференциальных свойствах изображения путем определения компонент вектора градиента $[B_x, B_y]$ как производных функции яркости $B(x, y)$ по направлениям x, y . Эти компоненты в целях сглаживания помех традиционно вычисляются через свертку с гауссианом (параметры: среднее – 0, дисперсия – σ^2), т. е. $B_x = q_{x, \sigma} * B$, $B_y = q_{y, \sigma} * B$, где $*$ – символ свертки, $q_{x, \sigma}, q_{y, \sigma}$ – пространственные производные двумерного гауссиана

$$q_{\sigma}(x, y) = \frac{1}{2\pi\sigma^2} e^{-(x^2 + y^2)/2\sigma^2}.$$

Декартово произведение транспонированного вектора градиента $[B_x \ B_y]^T$ с самим собой в каждой точке изображения дает матрицу градиентного тензора

$$C = \begin{bmatrix} c_{11} & c_{12} \\ c_{21} & c_{22} \end{bmatrix} = \begin{bmatrix} B_x^2 & B_x B_y \\ B_x B_y & B_y^2 \end{bmatrix},$$

а пространственное сглаживание элементов этого тензора вторым фильтром гауссиана (с параметром $\tilde{\sigma}^2$) приводит к структурному тензору:

$$\Lambda = \begin{bmatrix} \lambda_{11} & \lambda_{12} \\ \lambda_{21} & \lambda_{22} \end{bmatrix},$$

где $\lambda_{ij} = q_{\tilde{\sigma}} * c_{ij}$, ($i, j \in \{1, 2\}$). Вторая свертка усиливает двумерные свойства сигнала. Собственные значения матрицы Λ отражают кривизну автокорреляционной функции изображения в анализируемой точке.

В качестве детекторов локальных признаков как функций структурного тензора используются величины [3]

$$\begin{aligned} a_1 &= \det(\Lambda), \quad a_2 = \det(\Lambda)/\text{tr}(\Lambda), \\ a_3 &= \det(\Lambda) - 0,04(\text{tr}(\Lambda))^2 \end{aligned} \quad (1)$$

(здесь \det – определитель, tr – след матрицы), отражающие вклад ориентации и величины градиента в точке изображения. Значение a_3 соответствует детектору Харриса, который является базовым для многих современных структурных методов распознавания.

Большинство локальных операций, используемых при формировании ХП, связано с цифровыми свертками и носит линейный характер. Однако отдельные из них, например, «немаксимальное подавление», применяемое с целью сокращения количества анализируемых ХП [2, 7], реализует ранговую обработку и нелинейно влияет на оценивание значений признаков.

Результат распознавания в значительной степени зависит от базы моделей (набор эталонов). В большинстве практических задач компьютерного зрения трудно предположить, что вероятностное распределение интегральных признаков изображений имеет существенные различия для разных элементов базы моделей [5]. Как правило, речь идет лишь о принципиальной разнице в значениях отдельных параметров распределений. Сказанное относится и к множеству значений ХП, так как оно обычно имеет значительно большее разнообразие, чем система интегральных признаков. Более того, значения ХП часто оказываются достаточно близкими для разных эталонов.

Одним из возможных вариантов статистической модели есть такой, когда ХП для всех эталонов являются значением одной и той же случайной величины. Здесь тип ХП одинаков для всех классов. Распознавание на основе такого признака осуществить невозможно, однако, эту модель с успехом применяют на предварительном этапе обработки.

В целях реализации стратегии по сокращению объема вычислений формирование множества ХП часто реализуется в два этапа. На первом путем применения «быстрого» решения на основе одного из ключевых признаков фиксируется множество координат ХП, а далее в этих точках вычисляется более полное (и более точное) представление изображения в виде вектора. При этом требование инвариантности, устойчивости к помехам и сохранения различий классов относится к обоим этапам обработки. Примерами процедур первого этапа есть вычисление близких к своему глобальному максимуму значений детектора Харриса, второго – определение в выделенных точках значений вектора локальных потоков [6, 7].

Суть локального потока состоит в представлении функции яркости фрагмента в виде множества векторов, отражающих степень изменчивости яркости путем разложения изображения в ряд Тейлора в точке, соответствующей центру фрагмента. Потоки второго порядка позволяют оценивать ориентацию линий, потоки третьего порядка – кривизну. На основе значений потока можно построить инварианты [6]. Структурный тензор и инварианты на базе локальных потоков относятся к универсальным дифференциальным признакам, которые обеспечивают инвариантность к группе движений (смещение и поворот) и достаточно высокую устойчивость к помехам [7]. Отличительным моментом их построения, что в целом значительно уменьшает требуемый объем вычислений, является одна и та же вычислительная база – компоненты вектора градиента.

Другим вариантом статистического анализа есть модель независимых признаков – каждый ХП каждого из эталонов представляет собой отдельную случайную величину, которая «голосует» за свой эталон [7]. Третий вариант предполагает наличие корреляционных связей и даже статистическую зависимость между признаками как одного, так и разных эталонов [2, 5]. В целом задача создания универсального подхода при формировании ХП еще ждет своего решения.

2 КЛАССИФИКАЦИЯ И СХОДСТВО В ПРОСТРАНСТВЕ ХП

Эффективность распознавания в значительной мере зависит от различия значений ХП для разных эталонов, а также от меры, используемой для их сопоставления. Сходство между ХП определяется в рамках принятой статистической модели [5].

В структурных методах распознавания глобальное сходство объектов строится на частичном сходстве $\rho(v^i, v^j)$ в пространстве ХП с учетом эквивалентности их типа ($t_i = t_j$). Для определения значения $\rho(v^i, v^j)$ может быть использована стандартная мера сходства (различия) в векторном пространстве R^n . Примерами наиболее часто используемых в компьютерном зрении метрик для произвольных двух векторов v^1, v^2 являются

$$\begin{aligned} \rho_1(v^1, v^2) &= \sum_{i=1}^n |v_i^1 - v_i^2|, \\ \rho_2(v^1, v^2) &= 1/\exp\left[\sum_{i=1}^n |v_i^1 - v_i^2|\right]^2, \\ \rho_3(v^1, v^2) &= \sum_{i=1}^n |(v_i^1 - v_i^2)/\sigma_i|, \end{aligned} \quad (2)$$

где $|\dots|$ – знак модуля действительного числа.

Величины (2) непосредственно связаны с вероятностным описанием и имеют статистическую интерпретацию. Например, принятие решения по минимуму ρ_1 для бинарных признаков дает минимальную вероятность ошибки [5].

При известных статистических описаниях базы моделей эффективным может оказаться применение мер, основанных на значениях параметров моделей классов. Если известны статистические характеристики (математическое ожидание и дисперсия) μ_1, σ_1^2 и μ_2, σ_2^2 двух случайных величин, то степень эквивалентности (соответствия) этих величин можно оценить на основе статистики [4]

$$r = \frac{[(\sigma_1^2 + \sigma_2^2)/2 + ((\mu_1 - \mu_2)/2)^2]^2}{\sigma_1^2 \sigma_2^2}. \quad (3)$$

В качестве сравниваемых величин в (3) могут быть, например, характеристики яркости фрагментов, а также значения ХП. Для известных параметров можно рассматривать r как положительную детерминированную величину. При совпадении значений исходных параметров, т. е. при $\mu_1 = \mu_2, \sigma_1^2 = \sigma_2^2$, величина $r = 1$. С ростом различия между парами $\mu_1, \mu_2; \sigma_1^2, \sigma_2^2$ значение r увеличивается. В условиях шума величина r приобретает статистический вид и определяется законами распределений сравниваемых случайных величин.

В случае независимых компонентов вектора $v = (v_1, \dots, v_n)$ можно применить схемы, включающие решения отдельных компонент, и получение общего решения путем их интеграции (например, по значению максимального соответствия среди всех значений компонент, по числу соответствий, превышающих задан-

ний порог, путем суммирования величин соответствий и т. п.). Аналогичные схемы применяются также на более высоких уровнях иерархии анализа множеств ХП при сопоставлении с эталонными моделями.

Для сопоставления ХП можно использовать также функции подобия [5], например, нормированное скалярное произведение, определяющее косинус угла между векторами

$$\rho(v^1, v^2) = \frac{(v^1)^T v^2}{\|v^1\| \|v^2\|},$$

где $\|v\|$ – норма вектора, T – символ транспонирования. Для бинарных признаков функция подобия отражает меру обладания общими атрибутами (значениями признаков).

Если же параметры распределений классов имеют незначительные отличия, для надежного распознавания применяют нормированное евклидово расстояние, а также расстояние Махаланобиса, которое, кроме внутренней нормировки, учитывает корреляционные связи между компонентами вектора признаков

$$d(v^1, v^2) = \sqrt{(v^1 - v^2)^T \Sigma^{-1} (v^1 - v^2)}, \quad (4)$$

где Σ^{-1} – матрица, обратная к матрице ковариации Σ , имеющая размер вектора признаков и вычисленная на подготовительном этапе для множества ХП базы моделей. В случае статистически независимых признаков $\Sigma = \text{diag}(\sigma_1^2, \dots, \sigma_n^2)$, и

$$d^2(v^1, v^2) = \sum_{i=1}^n \left(\frac{v_i^1 - v_i^2}{\sigma_i} \right)^2. \quad (5)$$

В матрице Σ содержится важная для распознавания информация, которая более полно отражает свойства окрестности, для которой формируются значения ХП. Из различных вариантов её использования в структурных методах, где значимым считается каждый ХП, наиболее целесообразным представляется подход, когда отдельный ХП базы моделей имеет индивидуальную матрицу ковариации. Именно этот вариант не ограничивает выбор эталонов и делает анализ универсальным. Другие варианты, например, когда матрица Σ считается одинаковой для множества ХП отдельного класса или одинаковой для всех ХП базы моделей, упрощают решения, но в то же время приводят к ограничениям на базу моделей.

3 РАЗДЕЛЯЮЩИЕ ФУНКЦИИ И УРОВЕНЬ ОШИБОК

В соответствии с байесовским классификатором [5] минимальный уровень ошибки при отнесении ХП к одному из классов осуществляется посредством построения и оптимизации разделяющих функций вида

$$g_j(v) = -\frac{1}{2}(v - \mu^j)^T \Sigma_j^{-1} (v - \mu^j) - \frac{1}{2} \ln |\Sigma_j| + \ln P_j, \quad (6)$$

где J – количество классов изображений, $j = \overline{1, J}$, P_j – априорная вероятность класса j , μ^j – вектор математического ожидания признаков для j -го класса.

При одинаковых ковариационных матрицах ХП имеем упрощение $\Sigma_j = \Sigma$, и можно пренебречь вторым слагаемым, а слагаемое $\ln P_j$ можно отбросить в случае, если априорные вероятности классов одинаковы. В результате оптимальное решающее правило реализуется через определение квадратичного махаланобисова расстояния $d^2 = (v - \mu^j)^T \Sigma^{-1} (v - \mu^j)$ от анализируемого вектора v до каждого из J векторов μ^j средних значений ХП и отнесение вектора v к классу, соответствующему ближайшему среднему значению.

Для случая статистически независимых признаков с многомерным нормальным законом распределения при равных априорных вероятностях для двух классов байесовский уровень ошибки определяется выражением

$$E = \frac{1}{\sqrt{2\pi}} \int_{d/2}^{\infty} \exp\left[-\frac{1}{2}u^2\right] du, \quad (7)$$

где d^2 – квадратичное махаланобисово расстояние между средними векторами классов

$$d^2(\mu^1, \mu^2) = (\mu^1 - \mu^2)^T \Sigma^{-1} (\mu^1 - \mu^2).$$

Вероятность ошибки (7) убывает с ростом d . В случае независимых признаков махаланобисово расстояние имеет вид (5), откуда видно, что каждый из признаков влияет на уменьшение вероятности ошибки. Наилучшими в этом смысле являются признаки, у которых разность средних значений велика по сравнению со стандартными отклонениями. Вместе с тем из (5) видно, что ни один из признаков не бесполезен, если его средние значения для разных классов различны [5].

Если же каждый ХП базы моделей представить независимым случайным вектором с математическим ожиданием μ^{kj} и ковариационной матрицей Σ_{kj} , где j – номер класса, k – номер ХП в базе моделей, то разделяющая функция (6) будет представлена в виде

$$g_{kj}(v) = -\frac{1}{2}(v - \mu^{kj})^T \Sigma_{kj}^{-1} (v - \mu^{kj}) - \frac{1}{2} \ln |\Sigma_{kj}| + \ln P_{kj}, \quad (8)$$

причем оптимизацию (8) нужно производить по всему множеству эталонных значений ХП всех классов. Распознавание в соответствии с (8) более трудоемко в вычислительном плане, однако для случая структурного анализа на основе ХП оно более приемлемо с точки зрения универсальности базы моделей и адекватности представления изображения. Дело в том, что оно учитывает часто возникающую ситуацию, когда значения векторов ХП в силу особенностей локального представления могут быть достаточно близки между собой как в пределах одного эталона, так и для разных эталонов [1].

4 ОЦЕНКА ПАРАМЕТРОВ МОДЕЛЕЙ

Определение параметров распределений и качественная оценка процесса распознавания проводилась для полутоновых изображений автомобилей размером 64×64 (яркость в пределах 0–255), показанных на рис. 1. Можно заметить, что визуально распределения яркостей у эталонов рис. 1, б и рис. 1, в близки между собой. Гистограммы значений детектора Харриса для них оказываются также похожими. На рис. 2 приведено анализируемое изображение (в виде белых крестиков показаны отклики детектора Харриса), которое представляет собой эталон рис. 1, б на сложном фоне. На рис. 3 приведены гистограммы значений детектора Харриса для одного из эталонов (рис. 1, б) и анализируемого изображения (рис. 2).

Можно заметить характерные для всех трех эталонов скопления точек их гистограмм в интервалах значений 1–3, 5, 7, 12, что, в принципе, является дополнительным фактором при выделении ХП, а также может использоваться при распознавании. Диапазон значений детектора Харриса для эталонов составил $3,6 \dots 1,8e10$, а после пороговой обработки, в процессе которой отбирались точки со значениями в пределах 1 % от максимума, этот диапазон сузился до $1,8e8 \dots 1,8e10$ с небольшими отклонениями для разных эталонов (запись eN традиционно понимается как 10^N). При этом порог для точек интереса определялся независимо для каждого из эталонов. Диапазон значений детектора Харриса для изображения рис. 2 составил $1,98e5 \dots 1,98e7$, что несколько отличается от эталонного диапазона и в целом говорит о зависимости отклика от яркости фона.

Параметры распределения значений ХП разных классов в предположении нормального закона оценивались экспериментально следующим образом. Для каждого из эталонов определялись множества ХП в соответствии с детектором Харриса. Затем для сформированных множеств координат вычислялись значения инвариантных локальных потоков $v = (v_1, \dots, v_4)$, основанных на производных до 2-го порядка включительно с учетом сглаживания гауссианом [7]. Учитывая,



Рисунок 1 – Примеры эталонов



Рисунок 2 – Изображение с откликами детектора Харриса

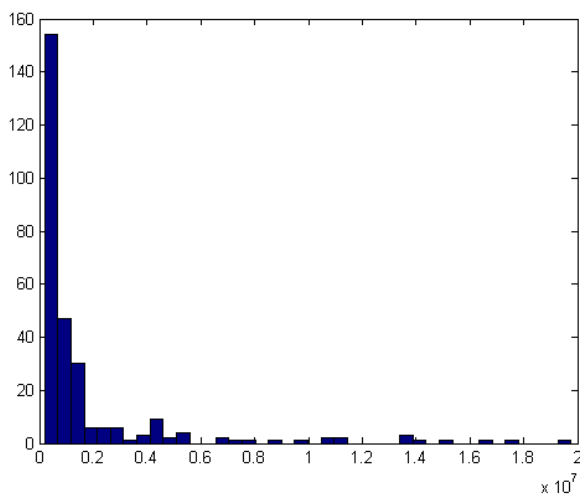
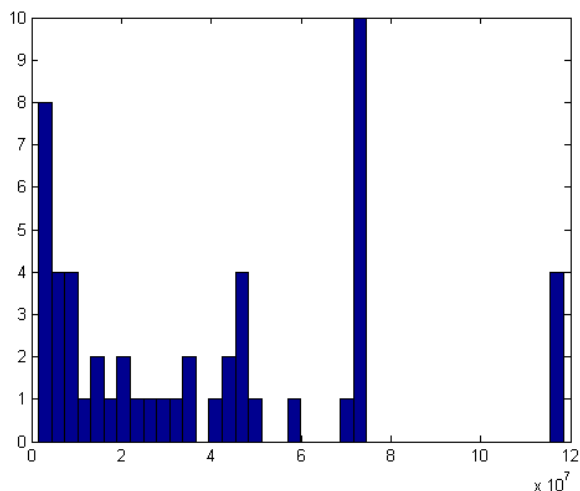


Рисунок 3 – Гистограммы значений детектора Харриса для изображений рис. 1, б и рис. 2

что величины потоков из-за ограниченных размеров фрагментов, на которых они строятся, чувствительны к ошибкам дискретизации (особенно при поворотах), значения эталонных признаков усреднялись по серии изображений в пределах углов поворота $0-90^\circ$. При усреднении в целях снижения влияния хвостов распределения применялась ранговая обработка. Для этого отбрасывалось 10 % наименьших и 10 % наибольших значений. В результате для трех конкретных эталонов рис. 1 получены оценки матожидания и дисперсии признаков, приведенные в табл. 1. В целом диапазон значений инвариантных локальных потоков составил: для признака v_1 : 9,4... 4320, для v_2 : -6,2e5...-71,6, для v_3 : -432... -32, для v_4 : 503...93620.

Что касается подтверждения нормального закона распределения значений локальных потоков, то следует отметить, что для анализируемого в эксперименте количества ХП (20-30 для разных эталонов) нельзя принять однозначное решение о нормальном законе распределения, т. к. данных недостаточно. Однако, как показали дополнительные исследования, с увеличением количества ХП для одного эталона (в пределах 200-300) принятие гипотезы о нормальности можно считать вполне правомочным.

Как видим из табл. 1, значения математических ожиданий для признаков разных эталонов имеют различия, что в целом позволяет применить для распознавания методы ближайшего среднего или ближайшего соседа, опираясь на оценки μ_i , σ_i . Заметим при этом, что ни один из признаков в отдельности не дает возможности надежного распознавания, т. к. их значения для некоторых пар эталонов близки между собой. Как видно из таблицы, значения потоков для 2-го и 3-го эталонов достаточно близки. Кроме того, дисперсия признаков достаточно велика по отношению к среднему, что не дает возможности установить четкие границы для значений эталонных признаков.

Усредненная по всем характерным точкам 2-го эталона матрица ковариации имеет вид, показанный в табл. 2.

Анализ значений матриц ковариации в пространстве локальных потоков показал, что их в данном эксперименте нельзя считать ни одинаковыми для всех классов, ни одинаковыми для каждого из классов в отдельности, что подтверждает целесообразность использования выбранной модели независимых ХП с индивидуальной матрицей ковариации.

Оценим теперь порог для значений сходства ХП. Учитывая, что квадрат расстояния Махаланобиса между двумя векторами с нормально распределенными компонентами - это случайная величина с распределением χ^2 [2], можно при заданном уровне значимости установить оптимальное значение порога для принятия решения об эквивалентности двух значений ХП. Это выглядит следующим образом. Если при заданном уровне значимости α и числе степеней свободы γ расстояние (4) (эмпирическое значение критерия) меньше заданного порога (критического значения из таблицы χ^2 -распределения), т. е. $d < \chi_{kr}^2[\alpha; \gamma]$, то нет оснований отвергать гипотезу об эквивалентности значений ХП и они считаются одинаковыми. Величина γ равна количеству слагаемых, т. е. размеру n вектора признаков $\gamma = n$. Например, $\chi_{kr}^2[0,05;4] = 9,49$. Таким образом, вектор размерности 4 признаков объекта, имеющий величину расстояния Махаланобиса $d < 9,49$ до эталонного вектора ХП, с вероятностью 0,95 может быть отнесен к данному эталону. Эти теоретические расчеты полностью подтверждаются экспериментом.

Посредством оптимального порога при заданном уровне значимости имеется возможность регулировать надежность распознавания отдельного ХП, а на ее основе оценивать и вероятность правильного глобального решения.

Приведенный подход анализа на основе χ^2 -распределения можно использовать и для метрики эвклидова

Таблица 1 - Параметры распределений характерных признаков

Эталон	v_1		v_2		v_3		v_4	
	μ_1	σ_1^2	μ_2	σ_2^2	μ_3	σ_3^2	μ_4	σ_4^2
<i>a</i>	1,05e3	1,05e6	-1,21e5	1,53e10	-224	9,87e3	2,99e4	5,95e8
<i>б</i>	1,55e3	1,22e6	-1,75e5	1,94e10	-235	7,61e3	3,12e4	4,33e8
<i>в</i>	1,50e3	1,83e6	-1,79e5	3,69e10	-216	9,98e3	2,81e4	5,17e8

Таблица 2 - Усредненная матрица ковариации для 2-го эталона

1,64e+06	-2,18e+08	-36734	2,10e+06
-2,18e+08	3,02e+10	6,06e+06	-5,65e+08
-36734	6,06e+06	11574	-2,74e+06
2,10e+06	-5,65e+08	-2,74e+06	6,93e+08

расстояния, что соответствует случаю некоррелированности компонентов вектора признаков между собой. При этом значения признаков нужно нормировать в виде $v_i^* = (v_i - \mu_i) / \sigma_i$, чтобы в итоге суммировать значения случайных нормальных величин с параметрами ($\mu_i = 0, \sigma_i = 1$).

Отдельной проблемой стоит выбор порогов для разделения значений признаков изображений разных классов. Для некоторых корреляционных подходов при структурном распознавании такая возможность имеется [1]. Для модели независимого анализа ХП, имеющиеся статистические данные говорят о том, что, учитывая разнообразие и диапазон значений, а также значительные величины дисперсий и ковариационных матриц, определение оптимальных порогов вызывает затруднения. Большую определенность в этой ситуации имеет вычисление порогов для величин локального сходства.

Типичные итоговые гистограммы голосования значений ХП за тот или иной эталон с использованием оптимальных порогов для сходства в приведенном примере задачи распознавания показаны на рис. 4. Цифры 1–3 соответствуют номерам эталонов, а значения гистограмм – количеству полученных голосов. На рис. 4, а изображена гистограмма при распознавании изображения рис. 1, б. Рис. 4, б соответствует распознаванию изображения рис. 2. Из диаграмм рис. 4 видим достаточно высокую схожесть между 2-м и 3-м эталонами в пространстве ХП, которая усиливается влиянием фона. Важным моментом при построении итоговых гистограмм является устранение учета откликов ХП на краях изображения. Эти отклики часто неразличимы

для разных эталонов. Результаты исследования помехоустойчивости исследуемого метода относительно аддитивных и локальных помех подробно описаны в [7].

ЗАКЛЮЧЕНИЕ

Новизна работы заключается в применении методов статистического анализа для оценки значений и величины сходства векторов ХП фрагментов изображений при решении задачи структурного распознавания объектов. Для фиксированного уровня значимости имеется возможность оптимальным образом устанавливать эквивалентность признаков по значению сходства. Для этой цели использованы метрические и статистические свойства векторов признаков.

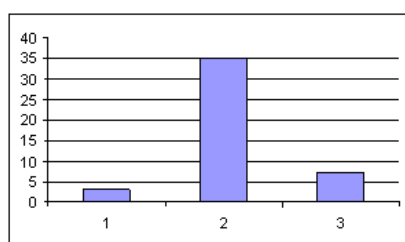
Важной практической особенностью используемой системы признаков есть общая вычислительная база (в виде компонент вектора градиента) как для формирования точек интереса, так и для вычисления в этих точках значений векторов-признаков.

Статистические оценки показывают, что более универсальными и приемлемыми в практическом плане моделями для представления множеств ХП являются те, в которых каждый ХП базы моделей имеет индивидуальную матрицу ковариации при участии в распознавании. Другие варианты приводят к ограничениям на базу моделей.

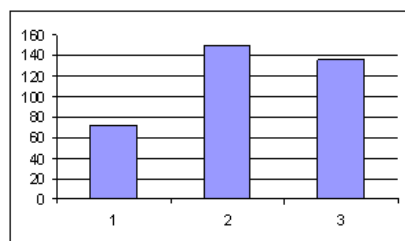
Практическая значимость обсуждаемых результатов состоит в экспериментальном подтверждении процедур проверки гипотез об эквивалентности признаков, а также таких принципиально важных свойств детекторов ХП, как инвариантность, помехоустойчивость, универсальность применения для разных видов признаков и базы распознаваемых моделей. Перспективы исследования состоят в статистическом обосновании решения по множеству отношений характерных признаков и их сравнение с решениями по одиночным признакам.

ПЕРЕЧЕНЬ ССЫЛОК

1. Гороховатский В. А. Распознавание изображений в условиях неполной информации. – Харьков: ХНУРЭ, 2003. – 112 с.
2. Schmid C., Mohr R. Local Greyvalue Invariants for Image Retrieval. – IEEE Transactions on Pattern Analysis and Machine Intelligence. – 1997. – 19(5). – P. 530–534.
3. Kothe, U. Edge and Junction Detection with an Improved Structure Tensor. – Pattern Recognition. Proc. 25th DAGM Symp. – Magdeburg-Germany: Springer, 2003. – P. 25–32.
4. Шапиро Л., Стокман Дж. Компьютерное зрение. Пер. с англ. – М.: Бином. Лаборатория знаний, 2006. – 752 с.
5. Дуда Р., Харт П. Распознавание образов и анализ сцен. – Москва: Мир, 1976. – 512 с.
6. Koenderink J.J., Doorn A.J. Representation of local geometry in the visual system. – Biol. Cybernetics. – 1987. – № 55. – P. 367–375.
7. Путятин Е. П., Гороховатский В. А., Кузьмин С. В. Распознавание изображений в пространстве инвари-



а)



б)

Рисунок 4 – Гистограммы голосования:

а – для изображения рис. 1, б;
б – для изображения рис. 2

- антных локальных признаков // Радиотехника и информатика. – 2006. – № 1(32). – С. 69–73.
8. Гонсалес Р., Вудс Р. Цифровая обработка изображений. – М: Техносфера, 2005. – 1072 с.

Надійшла 27.04.07

Приведені результати досліджень по застосуванню статистичного підходу при формуванні, оцінці значень, аналізі властивостей та встановленню еквівалентності характерних ознак зображень в цілях структурного розпізнання об'єктів. Експериментальні оцінки підтвер-

джують можливість побудови оптимальних локальних рішень.

Results of researches on application of the statistical approach are resulted at formation, an estimation of values, the analysis of properties and an establishment of equivalence of characteristic attributes of images with a view of structural recognition of objects. Experimental estimations confirm an opportunity of construction of optimum local decisions.

УДК 681.3.06

В. И. Долгов, А. В. Неласая

ГЕОМЕТРИЧЕСКИЙ ПОДХОД К СЛОЖЕНИЮ ДИВИЗОРОВ ГИПЕРЭЛЛИПТИЧЕСКОЙ КРИВОЙ

В статье приводится понятийный аппарат теории дивизоров гиперэллиптических кривых, а также рассматриваются примеры применения геометрического правила сложения для дивизоров различного типа гиперэллиптической кривой второго рода.

ВВЕДЕНИЕ

Современные мировые стандарты цифровой подписи, в том числе и украинский стандарт ДСТУ 4145–2002, основаны на арифметике эллиптических кривых над конечными полями. Сейчас интерес исследователей направлен еще на одну перспективную алгебраическую структуру – группу дивизоров гиперэллиптических кривых, размер основного поля которых может быть уменьшен в несколько раз в зависимости от рода кривой без потери стойкости. Однако групповая операция в этом случае усложняется.

Целью данной статьи является краткое изложение понятийного аппарата теории дивизоров гиперэллиптических кривых, а также примеров применения геометрического правила сложения различных типов дивизоров кривой второго рода с использованием специализированного математического пакета.

1 ПОНЯТИЙНЫЙ АППАРАТ ТЕОРИИ ДИВИЗОРОВ ГИПЕРЭЛЛИПТИЧЕСКИХ КРИВЫХ

В целях популяризации сведений из теории гиперэллиптических кривых, а также для большего понимания дальнейших рассуждений, следуя работе [1], приведем здесь некоторые важные понятия и определения.

Определение 1. Пусть F – конечное поле и пусть \bar{F} – алгебраическое замыкание F . Гиперэллиптическая кривая C рода $g \geq 1$ над F определяется как множество решений $(x, y) \in F \times F$ уравнения

$$C: y^2 + h(x)y = f(x), \quad (1)$$

где $h(x) \in F[x]$ – полином степени не более g , $f(x) \in F[x]$ – нормированный полином степени $2g + 1$, и при этом не существует решений $(x, y) \in \bar{F} \times \bar{F}$, которые бы одновременно удовлетворяли уравнению (1) и уравнениям с частными производными $2y + h(x) = 0$ и $h'(x)y - f'(x) = 0$.

Для конечной точки $P = (x, y)$ гиперэллиптической кривой C противоположной является точка $-P = (x, -y - h(x))$. Точка на бесконечности P_∞ противоположна сама себе.

Определение 2. Координатным кольцом C над F называется факторкольцо

$$F[C] = F[u, v]/(v^2 + h(u)v - f(u)),$$

где $(v^2 + h(u)v - f(u))$ – идеал в $F[u, v]$, порожденный полиномом $v^2 + h(u)v - f(u)$; $F[u, v]$ – кольцо многочленов от двух переменных, полученное последовательным присоединением к полю F сначала переменной u , а потом v .

Координатное кольцо C над \bar{F} определяется аналогично

$$\bar{F}[C] = \bar{F}[u, v]/(v^2 + h(u)v - f(u)).$$

Говорят, что рациональная функция $R(x, y)$ находится на кривой $C(F)$ если $R(x, y) = 0$ и $C(x, y) = 0$ имеет, по крайней мере, одно решение $P = (X, Y)$, где $X, Y \in F \cup \{\infty\}$, то есть P находится на обеих кривых $R(x, y) = 0$ и $C(x, y) = 0$, что обозначается $P \in R \cap C$.

Элементы $\bar{F}[C]$ рассматриваются как полиномиальные функции, определенные на кривой C .

Определение 3. Поле функций $F(C)$ кривой C над F это поле дробей из элементов $F[C]$. Аналогично, поле функций $\bar{F}(C)$ кривой C над \bar{F} – есть поле дробей из элементов $\bar{F}[C]$. Элементы $\bar{F}(C)$ называются рациональными функциями на C .

Определение 4. Пусть $R \in \bar{F}(C)$ – рациональная функция и пусть $P \in C, P \neq \infty$. Тогда R определена в P если существуют полиномиальные функции $G, H \in \bar{F}[C]$, такие что $R = G/H$ и $H(P) \neq 0$; если таких $G, H \in \bar{F}[C]$ не существует, то R не определена в P .

Если рациональная функция R определена в точке P , то значением R в P по определению является $R(P) = G(P)/H(P)$. Значение $R(P)$ не зависит от выбора G и H .

Определение 5. Пусть $R \in \bar{F}(C)^*$ и пусть $P \in C$. Тогда, если $R(P) = 0$, то R имеет нуль в точке P . Если R не определена в P тогда R имеет полюс в P , и, следовательно, $R(P) = \infty$.

Справедлива теорема:

Теорема 1. Пусть $P \in C$. Тогда существует функция $U \in \bar{F}(C)$, имеющая в точке P нуль кратности 1 с $U(P) = 0$, для которой выполняется свойство: для каждой полиномиальной функции $G \in \bar{F}[C]^*$ существует целое d и функция $S \in \bar{F}(C)$ такие, что $S(P) \neq 0, \infty$ и $G = U^d S$. Более того, число d не зависит от выбора U . Функция U называется униформизирующим параметром для P .

Определение 6. Пусть $G \in \bar{F}[C]^*$ и $P \in C$. Пусть $U \in \bar{F}(C)$ – униформизирующий параметр для P , и, следовательно, $G = U^d S$ где $S \in \bar{F}(C), S(P) \neq 0, \infty$. Порядок G в P определяется как $\text{ord}_P(G) = d$.

Из этого определения следует, что порядок функции R в точке P является мерой кратности нулей функции этой функции (см. рис. 1, заимствованный из [2], здесь функция обозначена f). Другими словами, порядок функции R в точке P есть кратность пересечения кривой C с $R = 0$.

Определение 7. Пусть $R = G/H \in \bar{F}(C)^*$ – рациональная функция на C и $P \in C$. Порядок R в P равен $\text{ord}_P(R) = \text{ord}_P(G) - \text{ord}_P(H)$.

Значение $\text{ord}_P(R)$ является порядком нуля поля функции R на кривой C , а $\text{ord}_P(R^{-1})$ тогда является порядком полюса функции R в точке P .

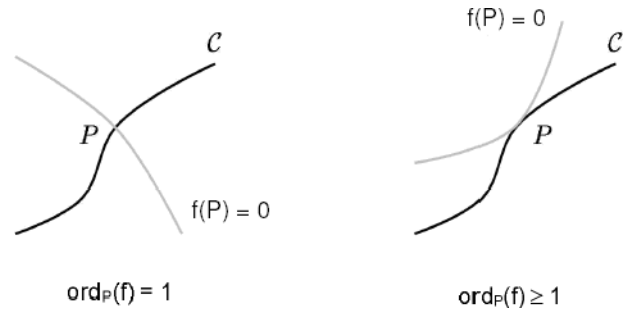


Рисунок 1 – Порядок функции f в точке P

Определение 8. Дивизором на алгебраической кривой называется конечная формальная сумма точек кривой $P_i \in C$

$$D = \sum_{P_i \in C} m_i [P_i], \quad m_i \in \mathbb{Z},$$

где только конечное число m_i не равно нулю.

Степень дивизора D , обозначаемая как $\text{deg} D$, есть целое число, определяемое суммой $\sum_{P \in C} m_p$. Порядок дивизора D в точке P – это целое число m_p : $\text{ord}_P(D) = m_p$. Количество точек дивизора называется весом дивизора.

Множество всех дивизоров формирует аддитивную группу с правилом сложения:

$$\sum_{P_i \in C} m_i [P_i] + \sum_{P_i \in C} n_i [P_i] = \sum_{P_i \in C} (m_i + n_i) [P_i].$$

Символом D^0 обозначается подгруппа, состоящая из дивизоров степени 0.

Ненулевая рациональная функция R , заданная на алгебраической кривой, может иметь лишь конечное число нулей и полюсов. Ее можно задать с точностью до константы списком нулей и полюсов с учетом их кратностей.

Для рациональной функции R на C понятие дивизора определяется следующим образом:

$$\text{div}(R) = \sum_{P \in C} \text{ord}_P(R) [P].$$

Дивизор функции называется главным дивизором. Известно, что $\text{div}(R)$ удовлетворяет условию

$$\sum_{P \in C} \text{ord}_P(R) = 0.$$

Тогда произведению функций на кривой будет соответствовать сумма их дивизоров.

Определение 9. Дивизор $D \in D^0$ называется главным дивизором, если $D = \text{div}(R)$ для некоторых рациональных функций $R \in \bar{F}(C)^*$. Множество всех главных дивизоров обозначим \mathcal{P} .

Дивизор $\text{div}(R)$ может быть разложен в разность двух дивизоров:

$$\text{div}(R) = \text{div}_0(R) - \text{div}_\infty(R),$$

где $\text{div}_0(R)$ соответствует пересечению C с кривой $R = 0$, а $\text{div}_\infty(R)$ – пересечению C с функцией $1/R = 0$.

На рис. 2 (здесь обозначение функции f), заимствованным из работы [2], показан пример главного дивизора. Для рассматриваемого примера $\text{div}(R) = P_1 + P_2 + P_3 + P_4 - (2Q_1 + 2Q_2)$.

Функции на гиперэллиптической кривой образуют группу по умножению, поэтому их гомоморфные образы – дивизоры функций образуют группу главных дивизоров и имеют степень 0.

Определение 10. Факторгруппа $J = D^0 / \mathcal{P}$ (группа дивизоров степени 0 по подгруппе главных дивизоров) называется Якобианом кривой C .

Если $D_1, D_2 \in D^0$ и $D_1 - D_2 \in \mathcal{P}$, то D_1 и D_2 называются эквивалентными дивизорами (обозначение $D_1 \sim D_2$). В соответствии с приведенным выше определением, каждый элемент Якобиана – это класс эквивалентности дивизоров.

Определение 11. Дивизор, представленный в виде

$$D = \sum_{P \in C} m_i [P_i] - (\sum_{P \in C} m_i [P_i])_\infty$$

называется приведенным, если:

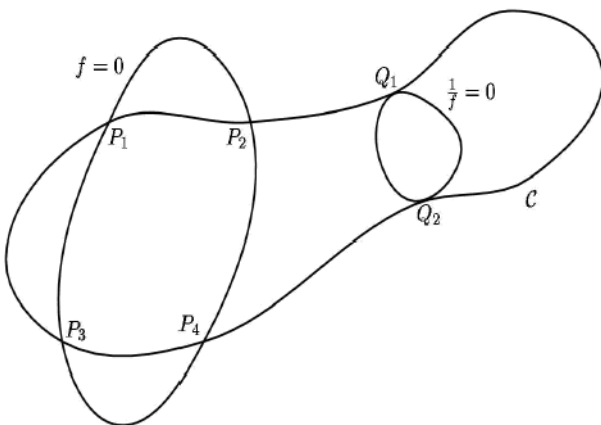


Рисунок 2 – Пример главного дивизора

1. Все m_i неотрицательны, и если $P = -P_i$ тогда $m_i \leq 1$.

2. Если $P_i \neq -P_i$, тогда P_i и $-P_i$ встречаются в сумме одновременно.

3. $\sum m_i \leq g$.

Каждый смежный класс якобиана обладает ровно одним представителем в виде приведенного дивизора.

Объектами, соответствующими дивизорам, являются идеалы координатного кольца $F[C]$ (см. определение 2) а объектами, соответствующими классам дивизоров, являются классы идеалов $F[C]$. По теореме Римана – Роха главные дивизоры отображаются в главные идеалы. Следовательно, можно рассматривать соответствие между классами дивизоров и классами идеалов. Это позволяет для удобства вычислений использовать интерпретацию дивизоров через классы идеалов полиномиальных порядков. Такое представление в виде пары полиномов было предложено Мамфордом [5].

Лемма 1. Пусть $P = (x, y)$ – обычная точка гиперэллиптической кривой C . Тогда для каждого $k \geq 1$ существует уникальный полином $b_k(u) \in \bar{F}[u]$ такой, что:

1. $\deg_u b_k < k$;

2. $b_k(x) = y$;

3. $b_k^2(u) + b_k(u)h(u) \equiv f(u) \pmod{(u-x)^k}$.

Теорема 2. Пусть $D = \sum m_i [P_i] - (\sum m_i) [\infty]$ – приведенный дивизор, где $P_i = (x_i, y_i)$. Пусть $a(u) = \prod (u-x_i)^{m_i}$. Существует уникальный полином $b(u)$, удовлетворяющий:

1. $\deg_u b < \deg_u a$;

2. $b(x_i) = y_i$ для всех i , для которых $m_i \neq 0$;

3. $a(u)$ делит $b(u)^2 + b(u)h(u) - f(u)$.

При выполнении условий теоремы 2

$$D = \text{gcd}(\text{div}(a(u)), \text{div}(b(u) - v)).$$

Обычно запись $\text{gcd}(\text{div}(a(u)), \text{div}(b(u) - v))$ сокращается до $\text{div}(a(u), b(u) - v)$, или, более просто до $\text{div}(a, b)$.

С использованием приведенных понятий и определений рассмотрим особенности выполнения арифметических операций с дивизорами гиперэллиптических кривых.

2 ПРИМЕРЫ ВЫПОЛНЕНИЯ АРИФМЕТИЧЕСКИХ ОПЕРАЦИЙ С ДИВИЗОРАМИ ГИПЕРЭЛЛИПТИЧЕСКИХ КРИВЫХ

Арифметика на гиперэллиптической кривой (операции сложения и дублирования элементов якобиана) реализуется с помощью функций на кривой. Как следует из вышеприведенных рассуждений, якобиан ги-

перэллиптической кривой второго рода включает дивизоры, образованные либо одной, либо двумя точками. Правило сложения хорд и касательных, применимое для эллиптической кривой, в этом случае уже не работает. Исходя из определения якобиана (определение 10), для того, чтобы построить группу, мы должны образовать факторгруппу сумм точек на кривой по подмножеству сумм тех точек, которые лежат на функции, определенной на кривой. В качестве функции на кривой возьмем полином, как обобщение понятия хорды, проходящей более чем через две точки. В общем случае, для нахождения n коэффициентов полинома $y = a_1x^{n-1} + a_2x^{n-2} + \dots + a_{n-1}x + a_n$ степени $n-1$ достаточно n точек. Сумма всех точек пересечения полученного полинома и заданной эллиптической кривой по определению является дивизором функции и дает нуль рассматриваемой факторгруппы. Пусть $D_1 = P_1 + P_2$ и $D_2 = Q_1 + Q_2$. Необходимо вычислить $D_3 = D_1 + D_2$. Показанная на рис. 3 [2], сумма точек пересечения полученного кубического полинома и гиперэллиптической кривой $P_1 + P_2 + Q_1 + Q_2 + (-R_1) + (-R_2)$ образует нуль группы. На рис. 3, заимствованном из работы [2], показана процедура вычисления суммы точек пересечения кубического полинома и гиперэллиптической кривой $P_1 + P_2 + Q_1 + Q_2 + (-R_1) + (-R_2)$. Эта сумма образует нуль группы. Отсюда очевидно, что противоположной к $(-R_1) + (-R_2)$ суммой $R_1 + R_2$ будет $P_1 + P_2 + Q_1 + Q_2$.

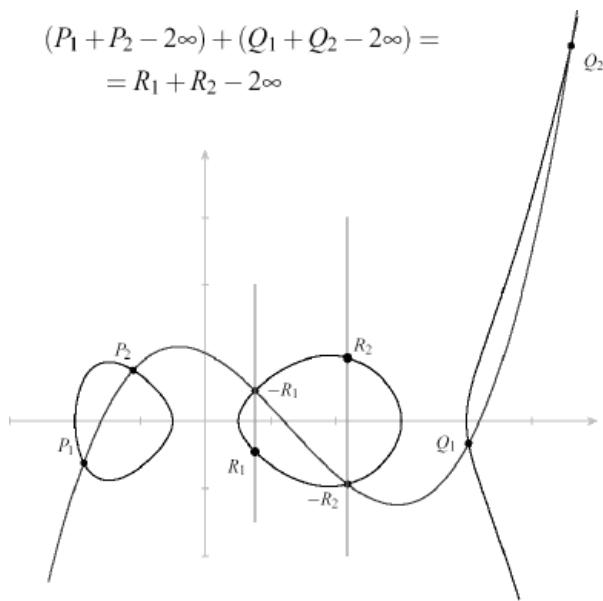


Рисунок 3 – Геометрическая визуализация сложения на кривой второго рода

В зависимости от вида входных дивизоров возможны различные варианты использования описанного правила сложения. В качестве примера рассмотрим гиперэллиптическую кривую

$$y^2 = x^5 + 5x \pmod{73}. \quad (2)$$

Она содержит 49 конечных точек: [27, 4], [53, 1], [53, 72], [36, 23], [36, 50], [56, 28], [56, 45], [37, 36], [37, 37], [32, 54], [32, 19], [21, 41], [9, 57], [12, 20], [12, 53], [9, 16], [41, 2], [41, 71], [61, 29], [64, 67], [64, 6], [29, 22], [29, 51], [72, 33], [44, 10], [44, 63], [27, 69], [46, 35], [46, 38], [72, 40], [50, 68], [24, 14], [61, 44], [24, 59], [21, 32], [49, 60], [49, 13], [50, 5], [0, 0], [1, 15], [1, 58], [17, 26], [17, 47], [52, 12], [52, 61], [20, 27], [20, 46], [23, 11], [23, 62] и бесконечно удаленную точку.

Поскольку данная кривая является кривой Фурукавы, ее порядок можно легко определить по приведенному в [6] алгоритму с помощью программы, разработанной авторами [7]. Полученный порядок, то есть количество приведенных дивизоров якобиана кривой, равен 3842. Все дальнейшие вычисления выполняются в поле $GF(73)$.

Случай 1. Сложение двух дивизоров веса 2. Все точки различны.

- $P_1 = (27, 4)$,
- $P_2 = (53, 1)$,
- $P_3 = (56, 45)$,
- $P_4 = (12, 20)$.

Представим дивизор $D_1 = P_1 + P_2 - 2\infty$ в форме Мамфорда в виде пары полиномов $\langle u, v \rangle$.

$$U = (x - 27)(x - 53) = x^2 + 66x + 44.$$

Полином $v = cx + d$ находим из условий

$$\begin{cases} 27*c + d = 4 \\ 53*c + d = 1 \end{cases}$$

откуда имеем $v = 42x + 38$. Следовательно,

$$D_1 = P_1 + P_2 - 2\infty = \langle x^2 + 66x + 44, 42x + 38 \rangle.$$

Аналогично

$$D_2 = P_3 + P_4 - 2\infty = \langle x^2 + 5x + 15, 52x + 53 \rangle.$$

Для того, чтобы получить дивизор, равный сумме дивизоров D_1 и D_2 построим кубический полином, проходящий через четыре точки P_1, P_2, P_3, P_4 (четыре точки однозначно определяют кубический полином, который имеет четыре коэффициента: коэффициенты

```

> # Все точки различны
restart;
p:=73:
x1:=27:
y1:=4:
x2:=53:
y2:=1:
x3:=56:
y3:=45:
x4:=12:
y4:=20:
S:=msolve ({y1=a1*x1^3+a2*x1^2+a3*x1+a4,
            y2=a1*x2^3+a2*x2^2+a3*x2+a4,
            y3=a1*x3^3+a2*x3^2+a3*x3+a4,
            y4=a1*x4^3+a2*x4^2+a3*x4+a4},p):
a:=subs (S,[a1,a2,a3,a4]):
msolve ({y^2=x^5+5*x,y=a [1]*x^3+a [2]*x^2+a [3]*x+a [4]},p);
{x = 9, y = 57}, {x = 12, y = 20}, {x = 27, y = 4}, {x = 53, y = 1}, {x = 56, y = 45}, {x = 72, y = 33}

```

при первой, второй и третьей степенях x и свободный член). Для их нахождения необходимо составить четыре уравнения. Это можно сделать, подставив в уравнение кубики поочередно координаты каждой из четырех входных точек. Выполним вычисления с помощью пакета компьютерной алгебры Maple.

Решая совместно систему уравнений, содержащую уравнение кривой и уравнение полученного кубического полинома определим две других точки пересечения (9, 57) и (72, 33). Противоположные к ним точки (9, 16) и (72, 40) и образуют дивизор $\langle x^2 + 65x + 64, 56x + 23 \rangle$, равный сумме $D1$ и $D2$.

Случай 2. Сложение двух дивизоров веса 2. $P1$ входит в сумму дважды.

$$P1 = (27, 4),$$

$$P2 = (53, 1),$$

$$P3 = (32, 19),$$

$$D1 = P1 + P2 - 2\infty = \langle x^2 + 66x + 44, 42x + 38 \rangle,$$

$$D2 = P1 + P3 - 2\infty = \langle x^2 + 14x + 61, 3x + 69 \rangle.$$

Если среди четырех исходных точек одна встречается дважды, то подстановкой координат исходных точек в уравнение искомого кубического полинома можно получить только три уравнения. Следовательно, необходимо получить еще одно дополнительное условие, связывающее коэффициенты искомого полинома, для получения четвертого уравнения. Это можно сделать, если учесть тот факт, что в двойной точке полином пересекает заданную гиперэллиптическую кривую дважды, то есть касается ее.

Для кривой, заданной уравнением $f(x, y) = 0$ касательная в точке $P(x_1, y_1)$ имеет вид [3]:

$$\left(\frac{\partial f}{\partial x}\right)\Big|_P (x - x_1) + \left(\frac{\partial f}{\partial y}\right)\Big|_P (y - y_1) = 0.$$

Воспользуемся тем фактом, что в двойной точке касательная к заданной гиперэллиптической кривой совпадает с касательной к искомому полиному. Обозначим $f(x, y): x^5 + 5x - y^2 = 0$ уравнение заданной гиперэллиптической кривой и $\varphi(x, y): a_1x^3 + a_2x^2 + a_3x + a_4 - y = 0$ уравнение искомого полинома. Тогда

$$\frac{\partial f}{\partial x} = (5x^4 + 5)\Big|_{(27, 4)} = 10,$$

$$\frac{\partial f}{\partial y} = (-2y)\Big|_{(27, 4)} = 65$$

и уравнение касательной к заданной гиперэллиптической кривой в точке (27, 4) имеет вид

$$10(x - 27) + 65(y - 4) = 0. \quad (3)$$

Аналогично, для искомого кубического полинома

$$\frac{\partial \varphi}{\partial x} = (3a_1x^2 + 2a_2x + a_3)\Big|_{(27, 4)} = 3a_127^2 + 2a_227 + a_3,$$

$$\frac{\partial \varphi}{\partial y} = (-1)\Big|_{(27, 4)} = -1$$

и уравнение касательной к $\varphi(x, y)$ в точке (27, 4) имеет вид

$$(3a_127^2 + 2a_227 + a_3)(x - 27) - 1(y - 4) = 0. \quad (4)$$


```

> # Точка [x1,y1] входит в сумму дважды
p:=73:
x1:=27:
y1:=4:
x2:=53:
y2:=1:
x3:=32:
y3:=19:
f_x:=5*x^4+5:
f_x:=modp1(Convertln(f_x,x),p);
f_y:=-2*y:
f_y:=modp1(Convertln(f_y,y),p);
f_x_P:=modp1(Eval(f_x,x1),p);
f_y_P:=modp1(Eval(f_y,y1),p);
S:=msolve({y1=a1*x1^3+a2*x1^2+a3*x1+a4,
            y2=a1*x2^3+a2*x2^2+a3*x2+a4,
            y3=a1*x3^3+a2*x3^2+a3*x3+a4,
            (3*a1*x1^2+2*a2*x1+a3)*(-f_y_P)=f_x_P},p);
a:=subs(S,[a1,a2,a3,a4]);
msolve({y^2=x^5+5*x,y=a[1]*x^3+a[2]*x^2+a[3]*x+a[4]},p);
{x = 27, y = 4}, {x = 32, y = 19}, {x = 44, y = 10}, {x = 49, y = 60}, {y = 1, x = 53}

```

Умножим обе части равенства (4) на число (-65) , чтобы урвать коэффициенты при $(y - 4)$ в равенствах (3) и (4). Поскольку два полученных уравнения представляют собой одну и ту же прямую, приравняв коэффициенты при $(x - 27)$, мы можем записать:

$$(3a_1 27^2 + 2a_2 27 + a_3)(-65) = 10.$$

Это и будет четвертое, недостающее условие, связывающее коэффициенты искомого полинома. Выше приведена реализация в Maple.

Решая совместно систему уравнений, содержащую уравнение кривой и уравнение полученного кубического полинома определим две других точки пересечения $(44, 10)$ и $(49, 60)$. Противоположные к ним точки $[44, 63]$ и $[49, 13]$ образуют дивизор $\langle x^2 + 53x + 39, 63x + 65 \rangle$, равный сумме $D1$ и $D2$.

Аналогичный результат можно получить, если входные дивизоры сформированы несколько иным способом:

$$D1 = 2P1 - 2\infty = \langle x^2 + 19x + 72, 56x + 25 \rangle,$$

$$D2 = P2 + P3 - 2\infty = \langle x^2 + 61x + 17, 20x + 36 \rangle.$$

В этом случае для представления дивизора $D1$ в форме Мамфорда воспользуемся уравнением касательной к кривой в точке $P1$.

$$u = (x - 27)^2 = (x + 46)^2 = x^2 + 19x + 72.$$

В этом случае полином $v = cx + d$ степени ≥ 1 определяется как уравнение касательной к кривой в заданной точке. Пусть кривая C задана уравнением

$\varphi(x, y) = 0$. Тогда касательная к C в точке $P(x_1, y_1)$ имеет вид [3]:

$$\left(\frac{\partial\varphi}{\partial x}\right)\Bigg|_P (x - x_1) + \left(\frac{\partial\varphi}{\partial y}\right)\Bigg|_P (y - y_1) = 0.$$

Для нашего примера имеем:

$$\varphi(x, y) = x^5 + 5x - y^2,$$

тогда

$$\frac{\partial\varphi}{\partial x} = (5x^4 + 5)\Big|_{(27,4)} = 10, \quad \frac{\partial\varphi}{\partial y} = (-2y)\Big|_{(27,4)} = 65.$$

Уравнение касательной к кривой в точке $(27, 4)$ имеет вид:

$$9(x - 27) + 65(y - 4) = 0;$$

$$y = 56x + 25.$$

Следовательно, $D1 = 2P1 - 2\infty = \langle x^2 + 19x + 72, 56x + 25 \rangle$.

При этом $D2 = P2 + P3 - 2\infty = \langle x^2 + 61x + 17, 20x + 36 \rangle$.

Их сумма также равна $\langle x^2 + 53x + 39, 63x + 65 \rangle$, следовательно, сумма двух дивизоров не зависит от порядка вхождения в них точек кривой.

Рассуждая аналогично, можно получить сумму двух двойных точек или выполнить дублирование, используя условия, определяемые касательными дважды.

```

> # В сумму входят три различных точки
p:=73:
x1:=27:
y1:=4:
x2:=53:
y2:=1:
x3:=32:
y3:=19:
S:=msolve ({y1=a1*x1^2+a2*x1+a3,
            y2=a1*x2^2+a2*x2+a3,
            y3=a1*x3^2+a2*x3+a3},p);
a:=subs (S,[a1,a2,a3]);
msolve ({y^2=x^5+5*x,y=a [1]*x^2+a [2]*x+a [3]},p);
{y = 15, x = 1}, {x = 27, y = 4}, {x = 29, y = 51}, {y = 19, x = 32}, {x = 53, y = 1}

```

Случай 3. Сумма состоит из трех различных точек, каждая из которых встречается в сумме ровно один раз.

$$P1 = (27, 4),$$

$$P2 = (53, 1),$$

$$P3 = (32, 19),$$

$$D1 = P1 - \infty = \langle x - 27, 4 \rangle,$$

$$D2 = P2 + P3 - 2\infty = \langle x^2 + 61x + 17, 20x + 36 \rangle.$$

В этом случае мы строим полином второго порядка $y = a_1x^2 + a_2x + a_3$, требующий наличия трех условий для нахождения трех коэффициентов. Выше приведена реализация в Maple.

Решая совместно систему уравнений, содержащую уравнение кривой и уравнение полученного полинома определим две других точки пересечения (1, 15) и (29, 51). Противоположные к ним точки (1, 58) и (29, 22) образуют дивизор $\langle x^2 + 43x + 29, 30x + 28 \rangle$, равный сумме $D1$ и $D2$.

Используя приведенные рассуждения, можно получить решения и для других возможных частных случаев в зависимости от вида входных дивизоров. Если же входные дивизоры содержат противоположные точки, которые аннулируют друг друга, необходимо сначала исключить их из рассмотрения, а затем продолжить сложение оставшихся точек по описанным выше правилам.

ЗАКЛЮЧЕНИЕ

Существует универсальный алгоритм Кантора [4], который можно использовать для сложения дивизоров на гиперэллиптических кривых любого рода. Однако рассмотрение частных случаев, зависящих от вида входных дивизоров, позволяет разработать явные формулы, являющиеся гораздо более быстрыми при про-

граммной и аппаратной реализациях. Приведенная в статье иллюстрация геометрического закона сложения дивизоров дает достаточно полное представление о возможных частных случаях, требующих детальной разработки в явном виде.

ПЕРЕЧЕНЬ ССЫЛОК

1. A. Menezes, Y. Wu, R. Zuccherato. An Elementary Introduction to Hyperelliptic Curves. – Springer-Verlag, Berlin, Germany, 1998. – 31 p.
2. T. Wollinger. Software and Hardware Implementation of Hyperelliptic Curve Cryptosystem. Dissertation for the Degree of Doctor-Ingenieur. – Bochum, Germany, 2004. – 201 p.
3. Г. Корн, Т. Корн. Справочник по математике. Для научных работников и инженеров. – М., 1974. – 832 с.
4. D. G. Cantor. Computing in Jacobian of a Hyperelliptic Curve // Mathematics of Computation. – Volume 48 (177), January. – 1987. – P. 95–101.
5. D. Mumford. Tata Lectures on Theta II // Prog. Math. – Volume 43. – Birkhäuser, 1984. – P. 61–75.
6. E. Furukava, M. Kawazoe, T. Takahashi. Counting Points for Hyperelliptic Curves of type $y^2 = x^5 + ax$ over Finite Prime Fields, [Электронный ресурс], (14 p), 2004 – Режим доступа, <http://eprint.iacr.org/181.pdf>, свободный.
7. Неласая А. В., Долгов В. И., Зайцев С. А. Определение порядка группы дивизоров гиперэллиптической кривой // Международная научно-техническая конференция «Компьютерное моделирование и интеллектуальные системы» (КМИС–2007), Запорожье, ЗНТУ, 26–27 марта 2007. – С. 173–177.

Надійшла 11.10.07

У статті приводиться понятійний апарат теорії дивизорів гіпереліптичних кривих, а також розглядаються приклади застосування геометричного правила для додавання дивизорів різного типу гіпереліптичної кривої другого роду.

The conceptual means of divisors hyperelliptic curves theory are considered. The examples of geometrical rule for adding of various kind divisors are given.

Т. С. Дьячук, Р. К. Кудерметов

РАСПРЕДЕЛЕНИЕ ЗАДАНИЙ В МНОГОПРОЦЕССОРНОЙ СИСТЕМЕ

Разработана математическая модель оптимального распределения заданий для параллельной вычислительной системы с учетом загруженности ее вычислительных узлов. Представлены результаты экспериментальных исследований модели, которые демонстрируют значительный прирост производительности по сравнению с равномерным распределением параллельного задания по узлам кластера.

ВВЕДЕНИЕ

Производительность распределенных и параллельных компьютерных систем, предназначенных для выполнения сложных вычислительных задач, существенно зависит от правильного распределения вычислительных заданий между узлами этих систем. Распределение заданий между параллельно работающими вычислительными узлами называют балансировкой нагрузки. Существует множество стратегий, методов и программных средств балансировки нагрузки в кластерных системах (PBS, MOSIX и др.) [1, 2]. Однако многообразие задач и структур вычислительных систем вызывают необходимость разработки и исследования все новых подходов к балансировке, которые более тонко учитывают особенности решаемых вычислительных задач и средств для их решения.

В данной работе предложена модель балансировки нагрузки для Beowulf-кластера при решении на нем одновременно нескольких задач (заданий), реализованных в виде параллельных программ. Кластер представляет собой множество компьютеров (узлов), объединенных сетью, работающих под управлением операционной системы Red Hat Linux 7.2. Для организации параллельных вычислений используется стандарт MPI 1, реализованный в библиотеке mpich 1.2.4 Аргоннской национальной лаборатории США.

ПОСТАНОВКА ЗАДАЧИ УПРАВЛЕНИЯ ЗАДАНИЯМИ

Под заданием будем понимать параллельную программу, части которой могут выполняться параллельно на различном количестве вычислительных узлов кластера, причем на одном узле может выполняться несколько частей параллельной программы, в зависимости от их распределения между узлами кластера. Задания могут возникать в произвольные моменты времени,

в зависимости от потребностей пользователей, получающих доступ к ресурсам кластера.

Эффективное распределение множества заданий зависит от различных факторов, в частности загруженности и объема свободной памяти узлов кластера, интенсивности обмена узлами через сеть, а так же и от характера самих заданий, например, сложности, их приоритета, количества параллельных частей. Распределение будет эффективным, если каждое из заданий выполняется за минимально возможное время.

Следует заметить, что время, затрачиваемое на распределение заданий, должно быть существенно меньшим по сравнению со временем выполнения заданий.

В данной работе авторами предложен подход, при котором учитывается только степень загруженности узлов кластера, и задача эффективного распределения сформулирована как задача оптимизации с нелинейной целевой функцией:

$$\Theta(\Omega(\beta_1, \beta_2, \dots, \beta_N), \Psi) \rightarrow \min, \quad (1)$$

где Θ – время выполнения задания; Ψ – функция, характеризующая задачу (например: степень ее параллелизма, требуемая память и т. д.); Ω – функция, характеризующая кластерную систему; β_i – загруженность i -го узла кластера; N – число узлов кластера.

МОДЕЛЬ РАСПРЕДЕЛЕНИЯ ЗАДАНИЙ

Назовем размерностью задания S_{\max} количество подзадач, на которое может быть разбито задание, и, соответственно, такое количество параллельных ветвей может иметь программа, реализующая задание. Таким образом, максимальное число узлов, которое может быть задействовано для выполнения задания, равно его размерности.

Введем обозначения:

N – число узлов кластера;

T – время выполнения задания на одном узле, при условии, что узел не загружен другими заданиями;

s – оптимальное количество ветвей задания для текущей загрузки узлов кластера ($s = 1 \dots S_{\max}$);

$t = T/s$ – время выполнения одной подзадачи на узле, при условии, что он не загружен другими заданиями;

x_i – число подзадач, выполняемых на i -м узле ($x_i \geq 0, i = 1, \dots, N$);

k_i – коэффициент замедления выполнения заданий на i -м узле, за счет уже выполняющихся заданий на данном узле. Коэффициент рассчитывается, исходя из процессорного времени, которое может быть выделено пользовательской подзадаче и зависит от приоритетов процессов, выполняющихся на данном узле.

Напомним, что любая программа, не являющаяся частью ядра, может выполняться в ОС Linux лишь в виде процесса. У каждого процесса есть приоритет планирования, который по умолчанию равен 20. Приоритет может быть изменен при помощи системного вызова `nice`, вычитающего свое значение из 20. Поскольку значение `nice` находится в диапазоне от -20 до $+19$, приоритеты всегда попадают в промежуток от 1 до 40. Чем выше приоритет, тем большая доля процессорного времени выделяется процессу [3].

Обозначим значение `nice` для уже выполняющихся процессов на i -м узле кластера как ni_{ig} , где g – порядковый номер процесса. Таким образом, каждый процесс на узле кластера имеет значение ni_{ig} в диапазоне от -20 до $+19$. Для пользовательских подзадач введем значение `nice` как ni , по умолчанию оно равно 0 (т. е. $ni = 0$).

Будем считать, что доля процессорного времени d , выделяемая процессу g , пропорциональна его приоритету:

$$d = (20 - ni_{ig}) \cdot \frac{1}{\sum_{j=0}^J (20 - ni_{ij})}, \quad (2)$$

где J – число процессов на i -м узле кластера.

В общем случае коэффициент замедления для процесса g находится как

$$k_i = \frac{1}{d} = \frac{\sum_{j=0}^J (20 - ni_{ij})}{(20 - ni_{ig})}. \quad (3)$$

Поскольку пользовательские подзадачи по умолчанию будут иметь значение `nice`, равное 0, то

$$k_i = \frac{\sum_{j=0}^J (20 - ni_{ij}) + x_i \cdot 20}{20} = \sum_{j=0}^J \left(1 - \frac{ni_{ij}}{20}\right) + x_i. \quad (4)$$

Введем переменную z_i , характеризующую степень загрузки узла кластера:

$$z_i = \sum_{j=0}^J \left(1 - \frac{ni_{ij}}{20}\right). \quad (5)$$

Данная переменная является суммой приоритетов всех процессов, выполняющихся на узле, разделенная на 20. Например, если значение `nice` для всех процессов, выполняющихся на узле, равно 0, то степень загрузки равна числу процессов на данном узле $z_i = J$. Максимальное значение $z_i = 2J$ достигается при значении `nice` равному -20 для всех процессов.

Таким образом, коэффициент замедления равен

$$k_i = z_i + x_i. \quad (6)$$

Время выполнения подзадач на i -м узле L_i

$$L_i = t \cdot k_i = k_i \cdot \frac{T}{s}. \quad (7)$$

Так как время выполнения параллельной задачи будет равно максимальному значению L_i и T является постоянным, то получаем функцию, которую необходимо минимизировать:

$$L_i = \max_i \left(\frac{k_i}{s}\right). \quad (8)$$

Таким образом, необходимо найти

$$L_{\text{опт}} = \min_s \max_i \left(\frac{z_i + x_i}{s}\right). \quad (9)$$

Решая задачу минимизации, находим оптимальное число подзадач s и число подзадач на каждой машине x_i , причем $s = \sum_i x_i$. После определения всех x_i задание распределяется между соответствующими i -ми узлами кластера.

РЕЗУЛЬТАТЫ ЭКСПЕРИМЕНТАЛЬНОЙ ПРОВЕРКИ МОДЕЛИ

Экспериментальная проверка предложенной модели распределения проводилась на пяти узлах кластера. Поскольку в данной модели из характеристик кластера учитывалась только степень загрузки узлов, задания были выбраны такими, которые не требуют интенсивного обмена данными между узлами кластера. Значение размерности S_{max} принято равным 22, то есть каждое задание могло состоять из 22 подзадач.

Для проверки эффективности системы было рассмотрено три случая:

– равномерная загрузка всех узлов кластера:

$$z_1 = z_2 = z_3 = z_4 = z_5 = z;$$

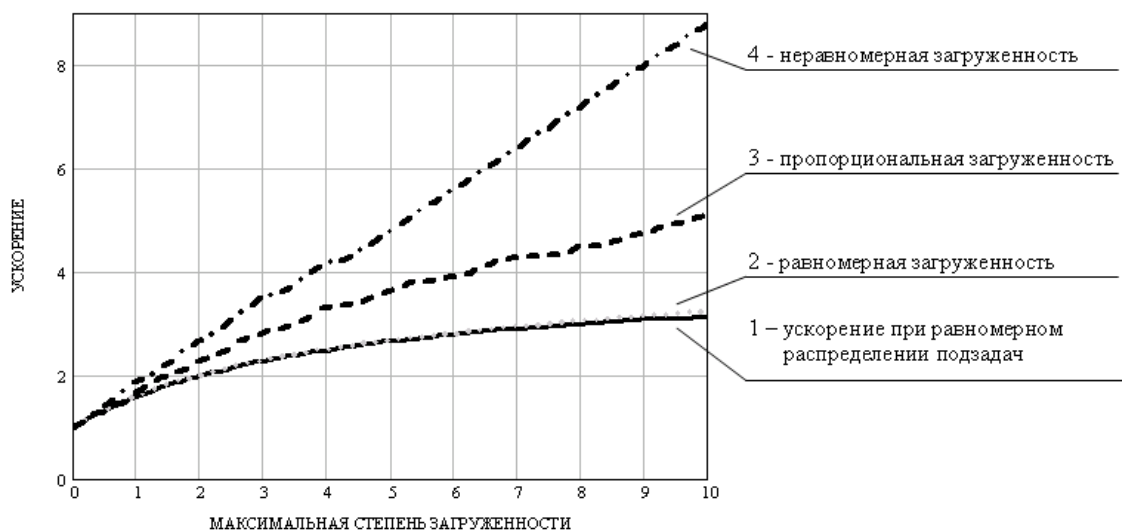


Рисунок 1 – Результаты исследований

– пропорциональная загрузка всех узлов кластера:

$$z_1 = 0, z_2 = z/4, z_3 = 2z/4, z_4 = 3z/4, z_5 = z;$$

– неравномерная загрузка всех узлов кластера:

$$z_1 = z_2 = z_3 = z_4 = 0, z_5 = z.$$

Значение z назовем максимальной степенью загрузки узлов кластера;

T_1 – время выполнения параллельной программы при простом распределении по одной подзадаче на каждый из узлов кластера, то есть $x_1 = x_2 = x_3 = x_4 = x_5 = 1$;

T_2 – время выполнения параллельной программы при равномерном распределении по четыре подзадачи на каждый из узлов кластера, то есть $x_1 = x_2 = x_3 = x_4 = x_5 = 4$;

T_3, T_4, T_5 – время выполнения параллельной программы по разработанному алгоритму для равномерной, пропорциональной и неравномерной загрузки узлов кластера соответственно.

Для каждого случая загрузки были рассмотрены следующие показатели:

– ускорение работы параллельного задания при равномерном распределении подзадач между пятью узлами кластера. При этом на каждом узле выполнялось по 4 подзадачи, то есть число подзадач $\approx S_{max}$, время выполнения T_2 . Ускорение равно T_1/T_2 . Так как T_1 и T_2 зависят от узла с максимальной загрузкой,

то для всех трех случаев загрузки значения ускорения будут одинаковы, график 1 рис. 1;

– ускорение работы параллельного задания при распределении подзадач по разработанному алгоритму (время выполнения T_3, T_4 и T_5 соответственно для каждого случая). Ускорение равно $T_1/T_3, T_1/T_4$ и T_1/T_5 , графики 2, 3 и 4 рис. 1.

Сравнения этих показателей и показывают эффективность разработанной системы.

На рис. 1 построена зависимость ускорения от максимальной степени загрузки для каждого случая.

Графики 1 и 2 совпали, так как при равномерной загрузке алгоритм равномерно распределяет подзадачи по всем узлам кластера. Наибольшее ускорение было получено при неравномерной загрузке кластера график 4.

ВЫВОДЫ

Разработанный алгоритм ориентирован на использование сведений о загрузке узлов, собранных непосредственно перед его работой. Он рассчитан на то, что нет значительной динамики в изменении загрузки за время его работы. Наибольший эффект достигается при неравномерной загрузке узлов кластера. В ходе дальнейших исследований планируется расширить данную модель, учитывая большее число факторов, разработать систему организации очередей заданий, предусмотреть гарантированное освобождение всех ресурсов, используемых задачами, от ее «обломков» при завершении, а также провести апробацию алгоритма при более сложных вариантах загрузки.

ПЕРЕЧЕНЬ ССЫЛОК

1. Лацис А. Как построить и использовать суперкомпьютер. – М.: Бестселлер, 2003. – 240 с.
2. Букатов А. А., Дацюк В. Н., Жегуло А. И. Программирование многопроцессорных вычислительных систем. – Ростов-на-Дону. Издательство ООО «ЦВВР», 2003. – 208 с.
3. Таненбаум Э. Современные операционные системы. 2-е изд. – СПб.: Питер, 2002. – 1040 с.

Надійшла 25.09.07

Розроблена математична модель оптимального розподілу завдань для паралельної обчислювальної системи

з урахуванням завантаженості її обчислювальних вузлів. Представлені результати експериментальних досліджень моделі, які демонструють значний приріст продуктивності в порівнянні з рівномірним розподілом паралельного завдання по вузлах кластера.

The mathematical model of tasks optimum distribution for the parallel computing system is developed in view its computing nodes load. Results the model experimental researches which show a significant increase of productivity in comparison with uniform distribution of the parallel task on nodes of a cluster are submitted.

УДК 004.087

А. А. Егошина

ФОРМАЛЬНАЯ МОДЕЛЬ СЛОВООБРАЗОВАТЕЛЬНОЙ СЕМАНТИКИ

В статье рассмотрен вопрос формализации семантики естественного языка. Предложена формальная модель словообразовательной семантики, учитывающая специфику словообразовательного значения в сравнении с заданными грамматическими и лексическими значениями производного слова.

ВВЕДЕНИЕ

Семантический компонент уже достаточно давно признается необходимой частью полного описания языка. Семантика как раздел лингвистики отвечает на вопрос, каким образом человек, зная слова и грамматические правила какого-либо естественного языка, оказывается способным передать с их помощью самую разнообразную информацию о мире (в том числе и о собственном внутреннем мире), даже если он впервые сталкивается с такой задачей, и понимать, какую информацию о ми-

ре заключает в себе любое обращенное к нему высказывание, даже если он впервые слышит его.

Семантика как информация, передаваемая языком или какой-либо его единицей (словом, грамматической формой слова), представляет собой не жестко детерминированную систему. Ячейка семантики (полнозначное слово) организована по принципу «семантического треугольника» [1] и может быть схематично представлена следующим образом.

Схема, представленная на рис. 1, резюмирует семантические отношения, т. е. аналогичным образом семантика организована во всех единицах языка.

Свой вклад в формирование общих принципов семантического описания вносят разные теории языка. Семантика как наука начала развиваться еще во второй половине 19 века. Фундаментальные лингвистико-гносеологические концепции В. Фон Гумбольта, А. А. Потебни, В. Вудта и др. определили первый этап развития



Рисунок 1 – Семантический треугольник

© Егошина А. А., 2007

семантики. Второй этап ознаменовался выделением семантики в особую часть языкознания (труды М. М. Покровского и других русских и немецких ученых). Третий этап развития семантики начинается приблизительно в двадцатых годах 20 века. Он характеризуется сближением семантики с логикой и ориентацией на синтаксис. К началу семидесятых годов прошлого века устанавливаются более гармоничный и комплексный подход к семантическим явлениям. В работах В. А. Звегинцева, Ю. Д. Апресяна и др. основной ориентацией становится анализ не абстрактного, изолированного предложения, а рассматривается предложение в реальной речи с учетом прагматики языка. А. В. Бондарко и Т. В. Бултина уже исследуют семантику морфологических форм.

ПОСТАНОВКА ЗАДАЧИ

Общеизвестен тот интерес к семантике, который характерен для современного этапа развития лингвистики. Интеллектуальные системы с естественно-языковым интерфейсом, который включает активное словообразование, в этом смысле не составляют исключения. Описание модуля словообразования является важной составной частью описания языковой системы в целом. Словообразование – одно из основных средств пополнения словарного состава языка новыми словами. Оно тесно связано с лексикой и грамматикой. Иногда в современной лингвистике словообразование рассматривают как часть лексикологии. Однако более характерен для русской лингвистической традиции грамматический подход к словообразованию, когда оно включается в состав грамматики. Основанием для этого является тот факт, что словообразовательные средства языка – это морфемные средства, и поэтому механизм словообразования во многом сближается с механизмом морфологического уровня языка, в результате чего проблемы словообразования тесно переплетаются с проблемами морфологии.

В [2] предлагается формализация семантики путем сведения описания производных (мотивированных) слов к описанию мотивирующих, более простых по морфемному составу. В работе [3] полагается, что если целью грамматики является построение правил порождения текста на данном языке, то в такую грамматику войдут лишь семантически регулярные конструкции. Однако автор [4] утверждает, что словообразовательная семантика может быть предметом описания как грамматики, порождающей текст, так и аналитической грамматики, служащей различным целям – теоретико-познавательным, типологическим и др. Подходы, предложенные в рассмотренных работах, не являются эффективными из-за необходимости хранения полных значений словоформ мотивирующих слов.

В предыдущих работах автора решены задачи разработки структуры лингвистического процессора интеллектуальной системы с естественно-языковым интерфейсом на основе многокомпонентного словаря [5]. Применение подсловарей морфем позволяет существенно сократить затраты памяти и время поиска [6]. Целью настоящего исследования является формализация словообразовательной семантики в интеллектуальной системе с морфемы словарями.

ФОРМАЛИЗАЦИЯ СЛОВООБРАЗОВАТЕЛЬНОЙ СЕМАНТИКИ

Формализация особенностей организации семантических единиц словообразования позволяет автоматизировать процесс построения производных слов с учетом заданной специфики грамматического и лексического значений.

Одной из основных трудностей формализации словообразовательной семантики является создание правил порождения семантически нерегулярных конструкций. Однако для решения таких задач как построение автоматизированных систем перевода, поисковых систем и интеллектуальных систем обучения языку автором предлагается рассматривать правила порождения семантически нерегулярных конструкций не в их реальном значении, а в «структуральном» значении, т. е. в значении, равном сумме значений составляющих слово морфем.

Для решения данной задачи необходимо знать значение этих морфем, в том числе и инвариантных. Значение инвариантных морфем – это тот участок семантико-словообразовательного поля, в котором находятся и контекстные значения данной морфемы. Этот участок очерчен теми семантическими границами, в пределах которых возможны новообразования с помощью данной морфемы.

Формально определим семантику каждого слова X языка как функцию $f(X)$. Слово x представляет собой упорядоченное множество морфем, т. е.

$$X = (X_1, X_2, X_3), \quad (1)$$

где X_1 – множество приставок, входящих в слово. $X_1 = (x_1^1, x_1^2, \dots)$. Так как слово может не содержать приставку, то $|X_1| \geq 0$ – мощность множества X_1 ;

X_2 – множество корней, входящих в слово. $X_2 = (x_2^1, x_2^2, \dots)$. Так как слово обязательно содержит хотя бы один корень, $|X_2| > 0$;

X_3 – множество суффиксов, входящих в слово. $X_3 = (x_3^1, x_3^2, \dots)$. Так как слово может не содержать суффикс, то $|X_3| \geq 0$.

Исходным пунктом классификации слов обычно является деление на части речи. Производные слова определенной части речи связаны между собой многообразными отношениями словообразовательного характера. лишь некоторые префиксы и единичные суффиксы (например, *-оньк*) выступают одновременно в разных частях речи. Остальные аффиксы присущи словам только одной части речи. разного вида семантические взаимодействия (совмещение значений и их отдельных компонентов), взаимоотношения между производными словами разных способов словообразования осуществляются только в пределах одной части речи. Таким образом, семантическое значение аффикса есть функция двух переменных: аффикса и части речи, т. е. $f(x_j^i, z_k)$, где z_k – элемент множества всех частей речи Z .

Пусть M – множество всех слов языка, тогда справедливо следующее

$$\forall x \in M f(x) = f(x_1^i, z_k) \cup f(x_2^j, z_m) \cup f(x_3^h, z_n). \quad (2)$$

В качестве примера рассмотрим семантику слова *белеть* (корень слова *бел-*, *-e* – суффикс, означающий абсолютное становление признака, заданного корнем, *-ть* – глагольный суффикс, указывающий на то, что выполняется действие).

X_1 – пустое множество, так как слово не имеет приставки;

$$X_2 = \{x_2^1\};$$

$$x_2^1 = \text{бел};$$

$$X_3 = \{x_3^1, x_3^2\};$$

$$x_3^1 = e;$$

$$x_3^2 = \text{ть};$$

$$Z = \{z_1\};$$

$$z_1 = \text{глагол};$$

$$f(x) = f(x_2^1, z_1) \cup f(x_3^1, z_1) \cup f(x_3^2, z_1);$$

$$f(x) = \text{«становится абсолютно белого цвета»}.$$

Учитывая изложенные выше позиции, можно определить причинно-следственную связь между семантической и морфологической формами слова. Определим процесс словообразования как

$$G = (I, O, \varphi), \quad (3)$$

где I – множество входных воздействий; O – множество мотивированных слов, образующих словообразовательное гнездо; φ – функция перехода (закон словообразования), задающая отображение $\varphi : I \rightarrow O$.

Множество входных воздействий определяется следующим образом:

$$I = (C, z, X), \quad (4)$$

где C – множество ключевых слов, определяющих семантику производного (мотивированного) слова;

z – часть речи;

X – исходное (мотивирующее) слово.

При выборе аффикса, необходимого для данного закона словообразования, учитывается его семантическое значение, ключевые слова которого должны соответствовать множеству C , т. е., если x^i – некоторый аффикс, то $f(x^i, z) \cap C \neq \emptyset$.

Рассмотрим предложенный подход на примере. Пусть задано следующее множество входных воздействий I :

$C = \{\text{лицо, выполняющее действие}\};$

$z = \{\text{существительное}\};$

$X = \{\text{преподавать}\};$

φ : «сохранение конечной гласной инфинитивной основы с добавлением суффикса»;

x^i – суффикс *-тель*;

$f(x^i, z) = \{\text{лицо или предмет, выполняющее или производящее действие}\};$

$$f(x^i, z) \cap C \neq \emptyset;$$

$$\varphi : I \rightarrow O = \text{преподаватель}.$$

ЗАКЛЮЧЕНИЕ

Таким образом, в настоящей работе построена формальная модель словообразовательной семантики, учитывающая специфику словообразовательного значения в сравнении с заданными грамматическими и лексическими значениями производного слова.

В дальнейшем планируется модификация модели с целью детализации словообразовательных категорий и учета синонимии и омонимии словообразовательных средств. Полученные результаты могут применяться при автоматизированном переводе и при разработке интеллектуальных информационно-поисковых и обучающих систем.

ПЕРЕЧЕНЬ ССЫЛОК

1. Лингвистический энциклопедический словарь / Гл. ред. В. Н. Ярцева. – М.: Сов. энциклопедия, 1990. – 685 с.: ил.
2. Тузов В. А. Математическая модель языка. – Л.: Изд-во Ленингр. ун-та, 1984. – 176 с.
3. R. Grzegorzczkova, J. Puzinina. Gramatyka opisowa a slowotworstwo // Biuletyn polskiego towarzystwa jezykownawczego. – XXX. – С. 82.
4. Улханов И. С. Словообразовательная семантика в русском языке. – М.: Наука, 1977. – 255 с.
5. Егошина А. А. Языковые и алгоритмические аспекты построения морфологических процессоров для интеллектуального поиска в полнотекстовых базах данных // VI международная конференция «Интеллектуальный анализ информации ИАИ-2006»: Киев, 16–19 мая 2006 г.: Сб. тр. / Рос. ассоц. искусств. интеллекта и др.; Под ред. Т. А. Таран. – К.: Просвіта, 2006. – 334 с.: ил.
6. Егошина А. А., Об одном способе построения статического словаря морфологического процессора // Материалы Седьмой Международной научно-техни-

чешской конференции „Искусственный интеллект. Интеллектуальные и многопроцессорные системы – 2006». – Т.2. – Таганрог: Изд-во ТРТУ, 2006. – 404 с.

Надійшла 13.11.07

В статті розглянуте питання формалізації природної мови. Надана формальна модель словотворчої семантики, яка ураховує специфіку словотворчого значення в порів-

нянні з заданими граматичними та лексичними значеннями похідного слова.

In the article the formalization of semantics of natural language is considered. The formal model of semantics of creation of words is offered which takes into account word-formation specificity of meaning in comparison with the given grammatic and lexical meanings of a derivative word.

УДК 681.391

В. Н. Журавлев, Е. А. Архипова

АНАЛИЗ МЕТОДА РАСЧЕТА ПАРАМЕТРА ЭФФЕКТИВНОСТИ МАСКИРОВАНИЯ РЕЧИ В ТЕХНИЧЕСКИХ КАНАЛАХ УТЕЧКИ

В статье проведен анализ метода цифровой корреляционной обработки контрольного фрагмента речи, который позволяет на основе расчета коэффициента корреляции обосновать аналитическую оценку параметра эффективности аддитивного маскирования речевых сигналов.

ВВЕДЕНИЕ. ПОСТАНОВКА ЗАДАЧИ

Среди разнообразных средств спецтехники в последние годы интенсивно развивается и совершенствуется рынок устройств, предназначенных для несанкционированного доступа (НСД) к конфиденциальной информации, содержащейся в речевом сигнале (РС), который является первичным общедоступным выражением результата мыслительного процесса человека.

При проектировании систем технической защиты речевой информации в выделенных помещениях (ВП) от утечки по каналам несанкционированного доступа (НСД) технической разведки противника (ТРП) основным параметром, определяющим, как показатель эффективности, так и категорию защиты, служит параметр словесной разборчивости $W(t)$. Данный параметр предлагается рассчитывать инструментально-расчетным методом [1–3], предполагающим следующие основные допущения.

1. Тестовый сигнал (модель РС) $si(\Delta\omega, t)$, маскирующий сигнал (МС) $sn(\Delta\omega, t)$ и сигнал канала утечки $su(\Delta\omega, t) = si(\Delta\omega, t) + sn(\Delta\omega, t)$ представляют собой случайные процессы с нормальным законом распределения плотности вероятности, при этом

$$\Delta\omega \times t \geq 10, \quad t \in [t_1, t_1 + T_a], \quad (1)$$

где $\Delta\omega$ – эффективная полоса частот РС, t_1 – время начала процесса анализа, T_a – интервал анализа эф-

фективности маскирования, который не превышает среднего интервала времени между ложными тревогами и выбирается из ряда фиксированных значений $T_a = (1, 5, 10, 30)$ секунд.

2. Тестовый и маскирующий сигналы представляют собой сигналы с равномерным распределением спектральной плотности мощности $N(\Delta\omega, t)$, т. е. со спектром «белого шума».

3. Точка канала НСД, в которой установлены средства ТРП, расположена на расстоянии 1 метр от источника тестового сигнала $si(\Delta\omega, t)$, при этом акустическое поле сигнала $su(\Delta\omega, t)$ полагается плоским с полностью активным акустическим сопротивлением среды распространения.

4. На основании предыдущего допущения, в части плоского акустического поля, а так же постоянства АЧХ рецептора средства ТРП на интервале времени T_a , отношение сигнал – помеха $SN(\Delta\omega, t)$ рассчитывается не как классическое [4] отношение эффективных энергий сигнала и помехи, поглощающихся рецептором средства ТРП за время интервала анализа:

$$SN(\Delta\omega, t) = 10 \lg \frac{\int_{t_1}^{T_a} p_{si}(\Delta\omega, t) dt}{\int_t^{T_a} p_{sn}(\Delta\omega, t) dt}, \quad t \in [t_1, t_1 + T_a], \quad (2)$$

где $p_{si}(\Delta\omega, t)$ и $p_{sn}(\Delta\omega, t)$ – мгновенные мощности сигнала и помехи, соответственно, а как логарифм отношения усредненных на интервале времени T_a акустических давлений, что возможно только для моногармонических анализируемых сигналов при активном входном сопротивлении рецептора давления.

В соответствии с отечественными руководящими документами [5], составляющими этого параметра для «белого» либо «розового» шумового сигнала, являются: с одной стороны – спектральная плотность мощности $N_{sn}(\Delta\omega, t)$ (для эталонного сигнала), с другой стороны – уровень звукового давления (акустические характеристики уровня выходного сигнала). Спектральную плотность мощности (СПМ) рекомендуется [6]

$$\text{рассчитывать по выражению: } N(\Delta\omega, t) = \frac{U_{\text{эфф}}^2}{1, 11\Pi_{\text{эфф}}},$$

где: $U_{\text{эфф}}$ – эффективное значение напряжения шумоподобного сигнала, $\Pi_{\text{эфф}}$ – значение эффективной полосы пропускания (терминология в соответствии с источником [6]). В рассматриваемом случае, сигналы $si(\Delta\omega, t)$ и $sn(\Delta\omega, t)$ содержат конечную сумму гармонических составляющих, рецептор звукового давления (микрофон либо слуховой аппарат человека) имеет комплексный характер входного сопротивления, а с учетом эффекта отоакустической эмиссии внутреннего уха, обладает существенной нестабильностью АЧХ на интервале времени T_a . На данном этапе можно акцентировать внимание на примере: квадрат амплитуды звукового давления идеального «белого» шума стремится к единице $p_{sn}^2(t) \rightarrow 1$, а его СПМ стремится к нулю:

$$N_{sn}(\Delta\omega, t) = \int_{t_1}^{t_1+T_a} sn(\Delta\omega, t)e^{-j\omega t} dt \rightarrow 0.$$

Инструментально-расчетный метод определения эффективности маскирования РС, изложенный в [1–3] трудоемок, содержит много эмпирических коэффициентов, которые не допускают его адаптации под РС конкретного диктора. Учитывая вышеизложенные основные допущения метода, можно констатировать, что его корректность вызывает сомнения, а исследование и анализ эффективности параметров маскирования РС в общем случае, представляет собой отдельную актуальную научно-техническую задачу, для которой объектом исследований является параметр информационной адекватности (разборчивости) сигнала $su(\Delta\omega, t)$ в точке НСД ТРП по отношению к тестовому сигналу $si(\Delta\omega, t)$.

1 ПРЕДЛАГАЕМЫЕ АЛГОРИТМЫ И РЕШЕНИЯ

Термины и определения. В связи с тем, что до настоящего времени в сфере активной защиты РС полностью не сформировался перечень терминов и определений, позволяющий углубленно анализировать объект исследований настоящей статьи, дополним, применив аксиоматический метод исследования, существующую

нормативную терминологию следующими терминами и определениями.

Шум – «беспорядочные звуковые колебания разной физической природы, характеризующиеся случайным изменением амплитуды, частоты и др.» [7]. В связи с тем, что на сигнал маскирования накладываются определенные ограничения [4], необходимо адаптировать данное определение к задачам СТЗИ.

Энергетическое маскирование – метод маскирования, при котором параметры МСП $sn(\Delta\omega, t)$ не адаптированы с параметрами РС $si(\Delta\omega, t)$, объем МСП $Qsn(t) = T_a\Delta\omega_{sn}(t)D_{sn}(t)$ равен объему канала утечки $Qku(t) = T_a\Delta\omega_{ku}(t)D_{ku}(t)$ и больше объема РС $Qsi(t) = T_a\Delta\omega_{si}(t)D_{si}(t)$ на всем интервале времени $t \in [t_1, t_1 + T_a)$ анализа $Qsn(t) = Qku(t) > Qsi(t)$, $t \in [t_1, t_1 + T_a)$, где $\Delta\omega_{sn}(t)$, $\Delta\omega_{si}(t)$, $\Delta\omega_{ku}(t)$ и $\Delta D_{sn}(t)$, $\Delta D_{si}(t)$, $\Delta D_{ku}(t)$ – эффективный диапазон частот сигналов $sn(\Delta\omega, t)$, $si(\Delta\omega, t)$ и канала утечки $su(\Delta\omega, t)$ и их динамические диапазоны, соответственно.

Параметрическое маскирование – метод маскирования, при котором параметры МСП $sn(\Delta\omega, t)$ адаптированы к параметрам РС $si(t)$ и объем МСП $Qsn(t)$ меньше объема канала утечки $Qku(t)$ и равен (либо больше) объема РС $Qsi(t)$ на интервале времени анализа $Qsn(t) < Qku(t) \geq Qsi(t)$, $t \in [t_1, t_1 + T_a)$.

Маскирующий сигнал помехи (МСП) – синтезированный СТЗИ сигнал $sn(\Delta\omega, t)$, который аддитивно воздействует на сигнал $si(\Delta\omega, t)$ в канале утечки.

$$N_{su}(\Delta\omega, t) = N_{sn}(\Delta\omega, t) + N_{si}(\Delta\omega, t), \tag{3}$$

где $N_{su}(\Delta\omega, t)$ – СПМ сигнала утечки в точке НСД ТРП. Данное воздействие вызывает эффект уменьшения объема $Qku(t)$ [8] канала утечки, который, в свою очередь, влияет на параметр информационной адекватности $W(t)$ в точке НСД ТРП.

Отношение сигнал – помеха (по МСП) – отношение спектральных плотностей мощности информационного сигнала и маскирующего сигнала помехи в полосе частот канала утечки, выраженное в децибелах (дБ).

$$SN(\Delta\omega, t) = 10\lg \frac{\int_{\omega_1}^{\omega_2} N_{si}(\Delta\omega, t)d\omega}{\int_{\omega_1}^{\omega_2} N_{sn}(\Delta\omega, t)d\omega}, \Delta\omega \in [\omega_1, \omega_2]. \tag{4}$$

Адекватность информации в точке НСД – вероятностный параметр $AD(t)$ соответствия, создаваемого ТРП, информативного смыслового образа $Oti(t)$, реальному информативному образу $Oi(t)$, который передается в выделенном помещении.

Прагматическая адекватность – логический параметр $PA(t)$, определяющий практическую ценность информации канала утечки для ТРП. В дальнейшем, по умолчанию, будем полагать, что данный параметр принимает значение «истина», $PA(t) = \text{true}$.

Синтаксическая адекватность – вероятностный параметр $SiA(t)$, формы представления сигнала РС $si(t)$ в канале утечки, не затрагивающий его смыслового содержания, например, язык общения.

Семантическая адекватность – вероятностный параметр $SeA(t)$ эффективности СТЗИ, который определяет соответствие передаваемых в выделенном помещении и получаемых ТРП смысловых образов. В случае определения семантической адекватности, как структурного параметра, то она будет характеризоваться изменением объема РС $Qsi(t)$ в точке НСД ТРП относительно объема канала утечки $Qku(t)$ на интервале времени анализа.

$$SeA(t) = \frac{\partial Qsi(t)}{\partial Qku(t)}, \quad t \in [t_1, t_1 + T_a]. \quad (5)$$

Системный параметр $SeA(t)$ на практике определяется в точке доступа ТРП к каналу НСД и измеряется в процентах разборчивости $W(t)$ счетного количества информационных элементов языка общения, которыми являются фонемы.

Для обеспечения логической последовательности анализа эффективности системных методов защиты РС произведем ранжирование предлагаемых параметров по эффективности их влияния на достоверность и надежность СТЗИ: 1) прагматическая адекватность; 2) семантическая адекватность; 3) синтаксическая адекватность.

Анализ волнового пакета акустического сигнала, распространяющегося в дисперсионной среде. В связи с тем, что, в общем случае, многолучевое и не плоское распространение сигналов $si(\Delta\omega, t)$ и $si(\Delta\omega, t)$ проявляется в их частотной и временной их дисперсии в объеме выделенного помещения, что искусственно повышает эффективность МСП. Худший случай минимальной эффективности маскирования будет наблюдаться для однолучевого распространения сферической звуковой волны сигнала $su(\Delta\omega, t)$, который должен являться мерой категорирования выделенного помещения. Для данного случая проведем анализ механизма распространения и взаимодействия акустических колебаний в воздушной среде.

Как известно, функция, описывающая гармоническую волну [9], распространяющуюся в бездисперсионной среде, записывается в виде:

$$u(x, t) = A \cos(kx - \omega t), \quad (6)$$

где $k = 2\pi/\lambda$ – волновой вектор, ω – круговая частота.

Анализ выражения (6) показывает, что можно ввести фазовую функцию $\phi(x, t)$ косинусоидальной бегущей волны, распространяющейся в положительном направлении оси OX , как аргумент волновой функции $\cos(\omega t - kx)$:

$$\phi(x, t) = \omega t - kx.$$

При анализе поведения гребня волны ($\cos\phi(x, t) \rightarrow \max$) или ее впадины ($\cos\phi(x, t) \rightarrow \min$), по мере увеличения времени необходимо переходить к большим значениям x так, чтобы фаза $\phi(x, t)$ была постоянной. Условие постоянства фазы с математической точки зрения означает, что полный дифференциал функции $\phi(x, t)$, имеющий вид

$$d\phi = \left(\frac{\partial\phi}{\partial t}\right)dt + \left(\frac{\partial\phi}{\partial x}\right)dx = \omega dt - kdx, \quad (7)$$

равен нулю. Приравнявая (7) нулю, находим условие постоянства фазы $\frac{dx}{dt} = v_f = \frac{\omega}{k}$, где v_f – фазовая скорость волны, которое дает связь между фазовой скоростью волны, частотой волны и волновым вектором. Условия распространения волны определяются свойствами среды. При этом ω , а, следовательно, и фазовая скорость, могут зависеть от волнового вектора k . Как указывалось, в общем случае рассматриваемые сигналы не являются плоскими волнами, их фазовая скорость v_f зависит от волнового вектора [10], поэтому их можно считать диспергирующими. Диспергирующая волна, представляющая собой суперпозицию бегущих волн с различными волновыми числами, будет менять свою форму по мере распространения в пространстве, так как составляющие с различными длинами волн распространяются с различными скоростями.

Рассмотрим задачу об описании движения волнового пакета, локализованного в момент времени $t = 0$ в пространстве. Данное условие означает, что огибающая волнового пакета является функцией достаточно быстро стремящейся к нулю при $x \rightarrow \pm\infty$. Введем обозначение $f(x) = u(x, 0)$. Условие «быстрого стремления функции к нулю» обеспечивает возможность ее разложения в ряд Фурье. Если функция $f(x)$ не периодическая функция по координате x , то ее суперпозиция составляющих ее функций непрерывна по переменной k , и выражается через интеграл Фурье $f(x) = \int_{-\infty}^{\infty} A(k)e^{ikx} dk$, где $A(k) = \int_{-\infty}^{\infty} f(x)e^{-ikx} dx$. Каждая гармоническая составляющая в данном выражении определяет собственную гармоническую волну с частотой $\omega = \omega(k)$, т. е. каждая частотная составляющая бегущей волны распространяется со своей собственной фа-

зовой скоростью $v_f = \frac{\omega(k)}{k}$. Искомая функция $u(x, t)$, описывающая бегущую волну, является суперпозицией гармонических бегущих волн, образующих исследуемые сигналы. Это означает, что найти $u(x, t)$ можно заменой kx на $[kx - \omega(k)t]$ в каждой гармонической составляющей суперпозиции $u(x, t) = \int_{-\infty}^{\infty} A(k)e^{i[kx - \omega(k)t]} dt$.

Однако, для того, чтобы использовать алгоритмы цифровой обработки сигналов при нахождении численных значений функции $u(x, t)$ вычисление интеграла необходимо проводить на интервалах конечной длительности, т. е. проводить разложение функции не в интеграл, а в ряд Фурье. В данном случае, как известно [11], при расчете значений функции $f(x)$, следует использовать следующие выражения:

$$f(x) = \sum_{n=1}^N A(n)e^{iknx}, \quad k = \frac{2\pi}{x_{\max} - x_{\min}},$$

$$A(n) = \frac{1}{x_{\max} - x_{\min}} \int_{x_{\min}}^{x_{\max}} f(x)e^{-iknx} dx, \quad n = 0, 1, \dots, N,$$

$$u(x, t) = \sum_{n=1}^N A(n)e^{i(kn(x - x_{\min}) - \omega(kn))}, \quad (8)$$

где N – число значений функции $f(x)$.

Таким образом, можно сделать вывод, что в точке НСД ТРП происходит суммирование спектральных плотностей мощности гармонических составляющих сигналов в соответствии с (3). Исследование отношения $SN(\Delta\omega, t)$ необходимо производить на интервале времени T_a , в течение которого основные спектральные параметры сигналов не претерпевают существенных изменений, т. е. на интервале стационарности.

В работе исследуются случайные процессы $si(\Delta\omega, t)$, $sni(\Delta\omega, t)$ и $sn(\Delta\omega, t)$, которые на интервале времени T_a обладают свойством стационарности [12]. Основываясь на результатах ранее проведенных исследований, положим, что исследуемые случайные процессы распределены по нормальному закону, т. о. известна их главная статистическая характеристика – плотность распределения вероятности. Важной характеристикой случайных процессов является их автокорреляционные функции (АКФ), определяемые как

$$R_{ss}(\tau) = \frac{1}{T_a} \int_0^{T_a} s(\Delta\omega, t)s(\Delta\omega, t + \tau) dt, \quad (9)$$

где $s(\Delta\omega, t)$ – исследуемый процесс.

Анализ АКФ определяет: полную мощность исследуемого процесса, равную значению $R_{ss}(0)$, мощность постоянной составляющей на интервале времени T_a ,

определяемой как $R_{ss}(T_a)$, и интервал корреляции τ_0 случайного процесса, который определяет верхнюю частоту ω_2 исследуемой полосы частот $\Delta\omega$.

Детерминированной характеристикой случайного процесса является СПМ $N(\Delta\omega, t)$, показывающая, как распределена мощность сигнала в исследуемой полосе частот $\Delta\omega$. СПМ случайного процесса и его АКФ связаны преобразованием Фурье (теорема Винера-Хинчина):

$$N(\Delta\omega, t) = \int_0^{T_a} R_{ss}(\tau)e^{i\omega\tau} d\tau. \quad (10)$$

Анализ амплитудного спектра и СПМ сигналов помехи. Представим исследуемые сигналы $s(\Delta\omega, t)$ моделью в виде упорядоченного набора прилегающих друг к другу фрагментов длиной T_a :

$$s(\Delta\omega, t) = \sum_{i=1}^N s_i(\Delta\omega, t), \quad (11)$$

где $s_i(t) =$

$$= \begin{cases} s(t), & t \in [t_i, t_i + T_a), \quad t_i = t_0 + (i-1)T_a, \quad i = \overline{1, N}, \\ 0, & t_i + T_a < t \leq t_i. \end{cases}$$

Для упрощения последующего анализа полагаем, что в пределах каждого отличного от 0 фрагмента значений сигнала $s_i(t)$ амплитуды центрированы и распределены по нормальному закону с дисперсией $D[s_i(t)]$.

Анализ спектрального коэффициента корреляции СПМ как параметра эффективности маскирования речи. В связи с тем, что зависимости речевого $si(\Delta\omega, t)$ и маскированного $su(\Delta\omega, t)$ сигналов не имеют функционального характера, т. е. равномерному изменению одного признака соответствует изменение другого признака в среднем, можно, для оценки тесноты связи между исследуемыми сигналами, применить методы корреляционного анализа СПМ, в частности коэффициент корреляционного отношения Пирсона [13]:

$$r = \frac{\sum_{i=1}^N (x_i - \bar{x})(y_i - \bar{y})}{\sqrt{\sum_{i=1}^N (x_i - \bar{x})^2} \sqrt{\sum_{i=1}^N (y_i - \bar{y})^2}}, \quad (12)$$

где x и y – исследуемые СПМ, \bar{x} и \bar{y} – выборочные средние значения, определяемые (в частности, для x)

$$\text{как } \bar{x} = \frac{1}{N} \sum_{i=1}^N x_i.$$

Таким образом, в качестве системного объективного параметра семантической адекватности (5) сигналов

в точке доступа ТРП к каналу НСД можно принять коэффициент корреляционного отношения Пирсона, определяемый в соответствии с (12), и рассчитываемый для функций СПМ $N(\Delta\omega, t)$ на т. н. «слоговых» временных интервалах T_a квазистационарного состояния речевых сигналов. Категорирование выделенных помещений можно производить в соответствии со шкалой Чеддока, которая позволяет укрупнено объединять коэффициенты корреляции в определенных интервалах тесноты связи между исследуемыми сигналами.

2 РЕЗУЛЬТАТЫ ЭКСПЕРИМЕНТАЛЬНЫХ ИССЛЕДОВАНИЙ

Методика эксперимента. Анализ сигналов проводился на персональном компьютере, который оснащен звуковой картой SB Audigy2, диапазоном квантования 16 бит и частотой дискретизации $f_s = 48$ кГц, отношением сигнал – шум порядка 80 dB. В качестве РС $si(\Delta\omega, t)$ анализировалось контрольное слово «лето», которое содержит гласные и согласные, а также вокализованные и невокализованные фонемы. Контрольное слово произносилось мужским голосом. Программирование алгоритма корреляционной обработки производилось в среде пакета программ MatLab 6.5. Параметр сегментации сигнала $si(\Delta\omega, t)$ на временные отрезки соответствовал слоговой постоянной времени $T_a = 23$ ms. Свойства центрированности реализации обеспечивается формой РС, симметричной относительно нуля. Интервал дискретизации $1/f_s = 20$ мкс значительно меньше максимального интервала корреляции $1/2\omega_2$ сигнала $si(\Delta\omega, t)$, что позволяет утверждать о состоятельности выборочных оценок математических ожиданий и дисперсий, рассчитываемых при анализе. Преобразования Фурье, корреляционные интегралы (10) и коэффициент корреляции (12) рассчитывались стандартными процедурами пакета MatLab.

В качестве сигналов маскирования $sn(\Delta\omega, t)$ исследовались два процесса: со спектральной плотностью мощности «белого» шума («HG»), который формировался массив значений стандартной процедурой пакета MatLab; и сигнал промышленного генератора типа «ANG2200» («Nang»). Перед проведением анализа задавалось интегральное значение отношения сигнал – помеха $SN(\Delta\omega, t)$, с постоянной интегрирования равной длительности контрольного слова, рассчитанное по общепринятой методике отношения среднеквадратических значений амплитуд сигнала и помехи. Маскировка сигнала $si(\Delta\omega, t)$ выполнялась на интервале длительности контрольного слова методом аддитивного маскирования СПМ в соответствии с выражением (3). Для проведения корреляционного анализа сигнал $N_{su}(\Delta\omega, t)$ сегментировался на N временных интервалов длительностью T_a в соответствии с (11). На каж-

дом интервале анализа T_a рассчитывались точечные оценки коэффициента корреляции $r(SN)$ (12).

Результаты экспериментов. На рис. 1 приведены результаты работы программы анализа амплитудного спектра и СПМ для двух сигналов маскирования – «HG» (левые рисунки) и «Nang» (правые рисунки). На верхних рисунках приведены временные диаграммы сигналов $sn(t)$, на средних – их амплитудные спектры, на нижних – СПМ.

Проведя анализ спектров (рис. 1) можно сделать выводы, что характеристики амплитудных спектров отличаются от СПМ как распределением амплитуд спектральных составляющих в исследуемом диапазоне частот, так и амплитудой спектральных составляющих, которая отличается более чем на 20 dB. Для сигнала «Nang» наблюдается изменение спектрального состава в высокочастотной области СПМ (>2 kHz), что говорит о корреляционной связи его высокочастотных составляющих, которые не участвуют в процессе маскирования РС.

На рис. 2 приведены результаты работы программы анализа точечных параметров эффективности маскирования контрольного слова «лето» для двух сигналов маскирования – «HG» (верхние рисунки) и «Nang» (нижние рисунки). На левых рисунках приведены зависимости коэффициента корреляции $r = f(T_a)$, на средних – зависимость амплитудного отношения сигнал – помеха от временного интервала анализа контрольного слова $SN[A(t)] = f(T_a)$, на правых – зависимость отношения сигнал – помеха СПМ от временного интервала анализа контрольного слова $SN[N(t)] = f(T_a)$.

Проведя анализ графиков (рис. 2), можно сделать выводы, что характеристики зависимости коэффициента корреляции $r = f(T_a)$ показывают большую эффективность сигнала «Nang» по сравнению с сигналом «белый шум», что подтверждается результатами ранее проведенных исследований [14]. Оценка параметра сигнал – помеха завышена, при расчете ее по методу [1–3], форма характеристик зависимостей $SN[A(t)] = f(T_a)$ и $SN[N(t)] = f(T_a)$ отличаются друг от друга, что говорит о различных соотношениях параметров суммируемых сигналов, что подтверждается предыдущим экспериментом настоящей статьи и сферичностью модели исследуемых акустических волн. Параметр сигнал – помеха является переменным на интервале длительности контрольного слова, его приблизительное соответствие интегральному (–10 dB) наблюдается только на временных интервалах вокализованных фонем.

На рис. 3 приведены результаты работы программы анализа точечных параметров зависимости коэффициента корреляции $r = f(SN, T_a)$ при маскировании контрольного слова «лето» двумя сигналами – «HG»

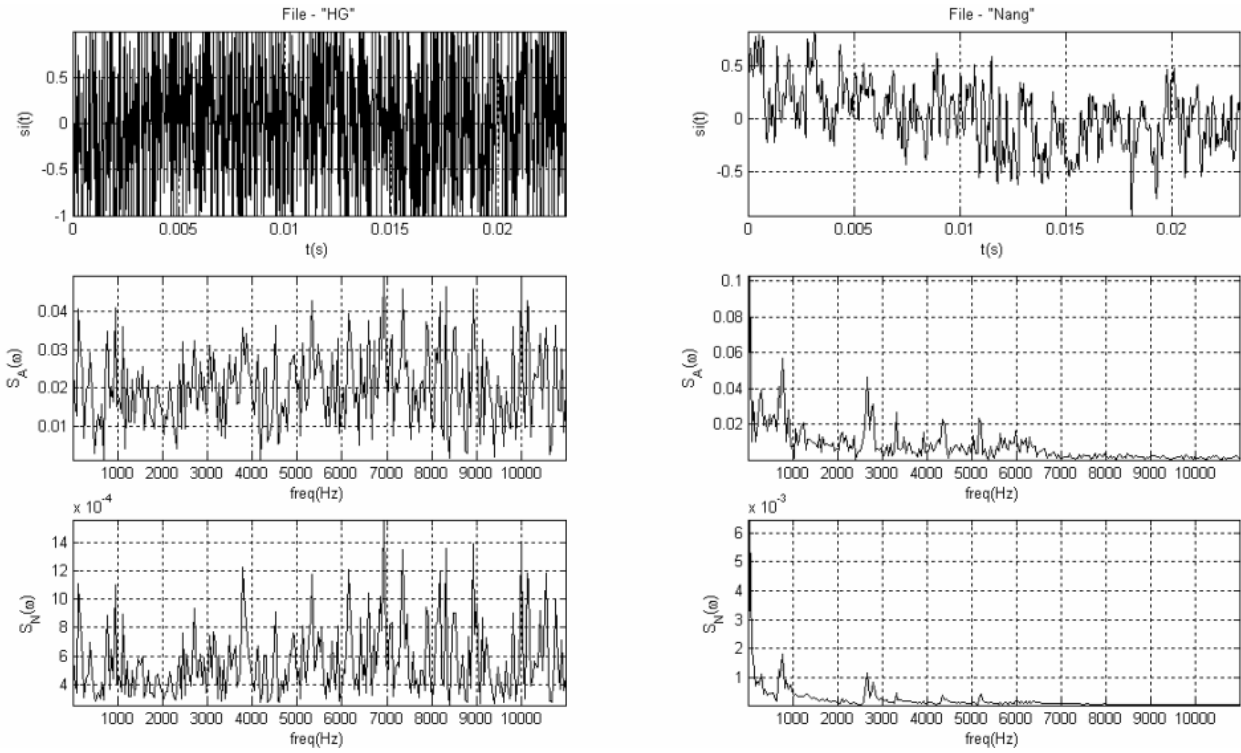


Рисунок 1 – Графики исследования амплитудных спектров и СПМ сигналов маскирования

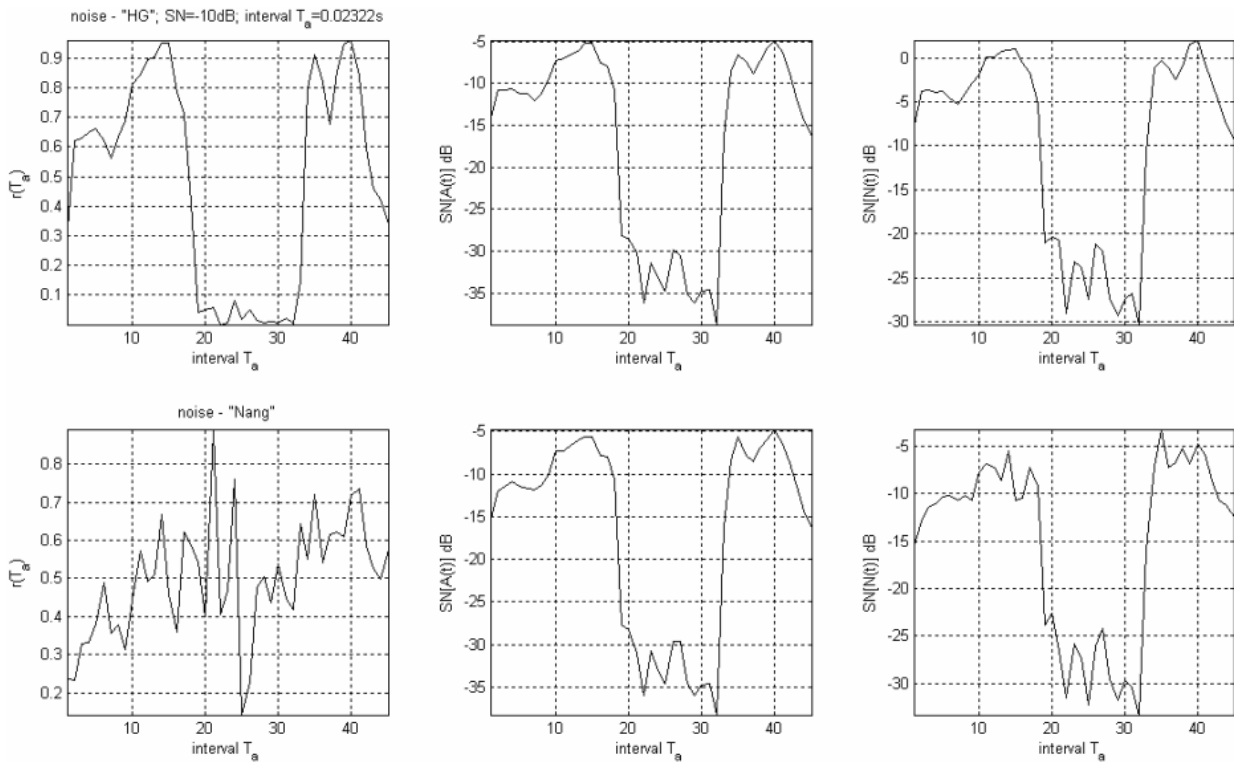


Рисунок 2 – Графики исследования зависимости коэффициента корреляции $r = f(T_a)$ и точечных отношений сигнал – помеха (амплитудного спектра и СПМ) аддитивного маскирования контрольного слова с интегральным отношением сигнал – помеха $SN = -10$ dB

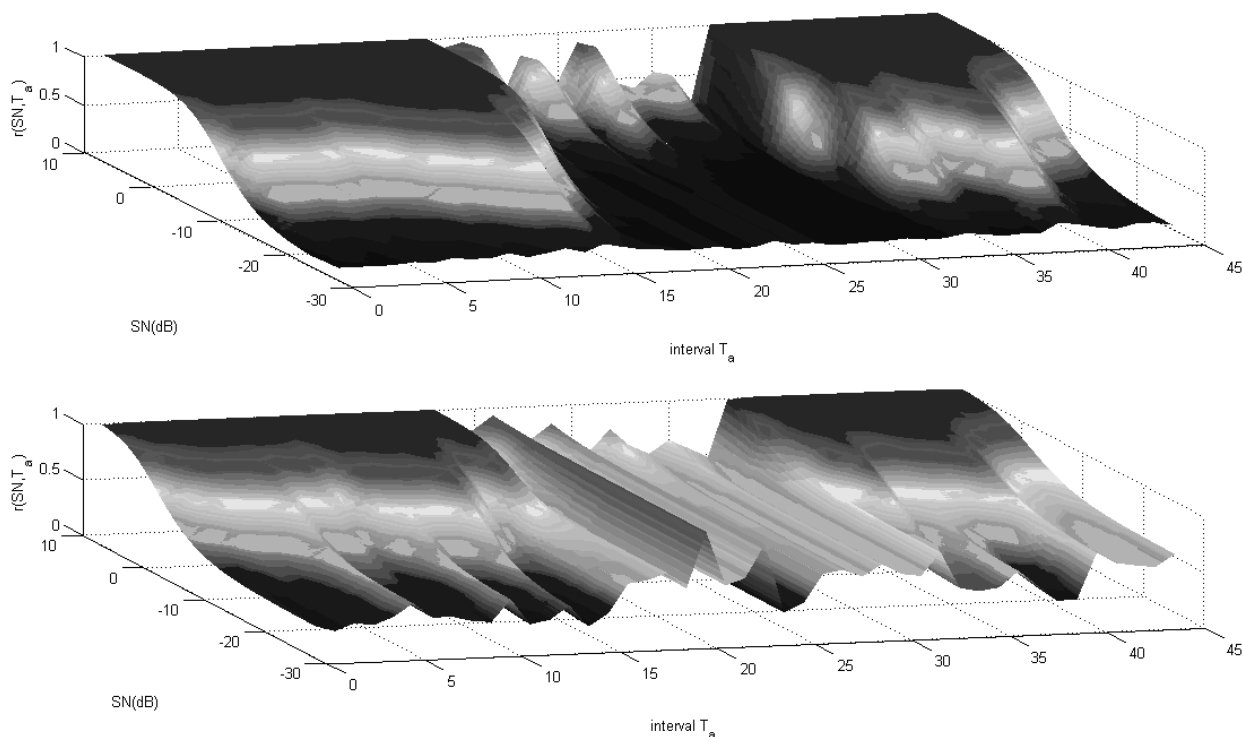


Рисунок 3 – Графики зависимостей значений точечных оценок коэффициента корреляции $r = f(SN, T_a)$ от интегрального значения отношения сигнал – помеха $SN[N(t)]$ от 10 до -30 dB

(верхний рисунок) и «Nang» (нижний рисунок) при изменении интегрального параметра сигнал – помеха $SN[N(t)]$ от 10 до -30 dB.

Проведя анализ графиков (рис. 3) можно сделать выводы, что характеристики зависимости коэффициента корреляции $r = f(SN, T_a)$ показывают большую эффективность сигнала «Nang» по сравнению с сигналом «белый шум» во всем диапазоне изменений интегрального параметра сигнал – помеха, что обобщает результаты предыдущего эксперимента. Характер изменения зависимости $r = f(SN, T_a)$ для сигнала «Nang» показывает внутренние корреляционные связи сигналов контрольного слова и маскирующего сигнала, особо это проявляется не интервале времени активности фонемы «т».

ВЫВОДЫ И ПРЕДЛОЖЕНИЯ

1. Предложенный и исследованный аналитический корреляционный параметр эффективности маскирования $r = f(T_a)$, определяемый в соответствии с выражением (13), является мерой оценки тесноты связи между исследуемыми сигналами $si(\Delta\omega, t)$ и $su(\Delta\omega, t)$, т. е. семантической адекватности в точке НСД ТРП $SeA(t) = r(t)$, $t \in [t_1, t_1 + T_a)$.

2. Приведенный анализ зависимости параметра отношения сигнал – помеха от вокализации фонем контрольного слова позволяет сделать вывод о том, что:

- в связи с тем, что РС на интервале длительности слова (или предложения) не стационарен по параметру постоянства дисперсии, при измерении и расчете $SN(t)$ необходимо учитывать только СПМ на стационарных интервалах времени вокализованных фонем, которые обладают природной помехозащищенностью;

- постоянная времени измерения СПМ должна укладываться в интервал стационарности, не может быть менее «слогового интервала» $T_a \geq 15$ ms и не превосходить времени активности вокализованной фонемы, что при эффективной полосе частот $\Delta\omega = 3000$ Hz соответствует требованиям (1).

3. Предложенный метод позволяет применять в качестве тестовых сигналов ключевые слова конкретных дикторов, и, таким образом, адаптировать СПМ для проведения конфиденциальных переговоров ограниченного количества собеседников.

4. Корректность исследованных в проведенном эксперименте характеристик изменения параметра эффективности маскирования с общепринятыми параметрами разборчивости РС при наличии помех необходимо подтвердить проведением артикуляционных испытаний по стандартной методике [15].

В статті проведено критичний аналіз методу розрахунку параметра ефективності маскування речі в технічних каналах утечки інформації. Предложено і підтверджено експериментальними дослідженнями метод цифрової кореляційної обробки контрольованого фрагмента речі, який усуває недоліки існуючого методу і дозволяє на основі розрахунку коефіцієнта кореляції АКФ сегментів сигналу контрольованого слова і сигналу каналу утечки розрахувати параметр ефективності адитивного маскування речевих сигналів.

ПЕРЕЧЕНЬ ССЫЛОК

1. Дворянкин С. В., Макаров Ю. К., Хорев А. А. Обоснование критериев эффективности защиты речевой информации от утечки по техническим каналам // Защита информации. INSIDE. – 2007. – № 2. – С. 18–25.
2. Хорев А. А. Оценка эффективности защиты информации от утечки по техническим каналам // Специальная техника, 2006. – № 6. – С. 53–61.
3. Хорев А. А. Оценка эффективности защиты информации от утечки по техническим каналам // Специальная техника, 2007. – № 1. – С. 51–64.
4. Шеннон. Связь при наличии шума. В кн. Работы по теории информации и кибернетике. – М.: Издательство иностранной литературы, 1963. – 827 с.
5. Засоби активного захисту мовної інформації з акустичними та віброакустичними джерелами впроміювання. Класифікація та загальні технічні вимоги. НД ТЗІ-Р-001-2000. ДСТСЗІ СБ України. – Київ.: – 2000. – 9 с.
6. НД ТЗІ 2.3-003-2001. Технічний захист мовної інформації в симетричних абонентських аналогових телефонних лініях. Засоби активного приховування мовної інформації. Генератори спеціальних сигналів. Методика випробувань. ДСТСЗІ СБ України. – Київ.: – 2001. – 21 с.
7. Советский энциклопедический словарь. / Гл. ред. А. М. Прохоров. 4-е изд. – М.: Сов. энциклопедия, 1989. – 1632 с.
8. Отт Г. Методы подавления шумов и помех в электронных системах. /пер с англ. – М.: Мир, 1978. – 317 с.
9. Крауфорд Ф. Волны. – М.: Наука, 1974. – 358 с.
10. Акустика: Справочник. Под ред. М.А. Сапожкова. – М.: Радио и связь, 1989. – 336 с.
11. Цифровая обработка сигналов / А. Б. Сергиенко – СПб.: Питер, 2003. – 608 с.
12. Цвикер Э., Фельдкеллер Р. Ухо как приемник информации. Пер. с нем. Под редакцией Б. Г. Белкина. – М.: Связь, 1971. – 255 с.
13. Гайдышев И. Анализ и обработка данных: специальный справочник. – СПб: Питер, 2001. – 752 с.
14. Журавлев В. Н., Прокофьев М. И. Анализ результатов артикуляционных и сегментальных испытаний сигналов маскування речі. Правове, нормативне та метрологічне забезпечення системи захисту інформації в Україні. – 2006. – № 13. – С. 36–48.
15. ГОСТ Р 50840-95. Государственный стандарт Российской Федерации. Передача речи по трактам связи. Методы оценки качества, разборчивости и узнаваемости. Издание официальное. – М.: Госстандарт России, 1997. – 12 с.

Надійшла 3.10.07
Після доробки 17.10.07

В статті проведений аналіз методики цифрової кореляційної обробки контрольованого фрагменту мови, що дозволяє на основі розрахунку коефіцієнта кореляції обґрунтувати аналітичну оцінку параметра якості передачі сигналу мови по каналам зв'язку та проводити аналіз ефективності адитивного маскування мовних сигналів.

The method of check utterance digital correlation processing, which allows to substantiate the communication path quality criteria analytic estimation on the base of correlation coefficient calculation and to analyze the speech signal additive masking effectiveness is under review.

УДК 004.75

М. Б. Ильяшенко

АЛГОРИТМ ОПТИМАЛЬНОГО РЕЗЕРВИРОВАНИЯ РАСПРЕДЕЛЕННЫХ ВЫЧИСЛИТЕЛЬНЫХ РЕСУРСОВ ПО КРИТЕРИЮ ЭФФЕКТИВНОГО ИСПОЛЬЗОВАНИЯ ВЫЧИСЛИТЕЛЬНЫХ УЗЛОВ

В работе предлагается алгоритм оптимального резервирования вычислительных ресурсов, способный осуществлять поиск областей резервирования, согласно критериям оптимальности по числу и максимальной загруженности задействованных вычислительных узлов. Приводится детальное описание алгоритма и набора предварительных условий, которые сужают область поиска переробной части алгоритма.

© Ильяшенко М. Б., 2007

ВВЕДЕНИЕ

Алгоритмы управления распределенными вычислительными ресурсами, как правило, используют одну из двух возможных парадигм: централизованное управление ресурсами [1, 2] или контроль ресурсов локальными приложениями [3, 4]. Каждая парадигма имеет свои достоинства и недостатки, в частности, централи-

зованное управление ресурсами позволяет производить глобальную оптимизацию выделения, использования и переназначения ресурсов сети, но этот подход недостаточно масштабируем как в терминах производительности, так и устойчивости сети в целом. С другой стороны, второй подход, основанный на локальных менеджерах, управляющих ресурсами одного или нескольких компьютеров, может приводить к нестабильному выделению ресурсов и не позволяет получать оптимальные решения в масштабах всей сети по такому параметру, как например, средняя загрузка ресурсов сети. Каждый подход имеет свою область применения. Локальное управление ресурсами рационально использовать в сетях с частыми изменениями в структуре использования ресурсов и большим количеством мелких задач, выполняющихся в рамках сети GRID, в отличие от централизованного подхода, который лучше применять при небольшом количестве ресурсоемких задач, работающих в течение длительного времени и не требующих частого перераспределения ресурсов сети.

Кроме того, большинство подходов к распределению ресурсов сети не учитывают необходимость совместного выделения ресурсов (co-allocation), необходимого для исполнения параллельных задач в рамках сетей GRID. Наиболее важными видами ресурсов для параллельных задач являются продуктивность процессоров, объем доступной оперативной памяти и пропускная способность или латентность каналов связи для каждого вычислительного узла сети. Алгоритм глобального резервирования распределенных вычислительных ресурсов должен поддерживать все типы ресурсов, которые могут быть зарезервированы и при этом влияют на производительность программ. В рассмотренном алгоритме базовыми ресурсами, подлежащими резервированию, являются минимальная производительность процессов и минимальная необходимая полоса пропускания сети.

В работе использован графо-аналитический подход к решению задачи оптимального резервирования распределенных вычислительных ресурсов в сетях GRID. Недавний прогресс в развитии алгоритмов точной проверки графов на изоморфность [5–8] позволяет приступить к разработке новых алгоритмов глобальной оптимизации выделения ресурсов, способных работать с GRID сетями реальных размеров. Алгоритм оптимального резервирования распределенных вычислительных ресурсов, представленный в работе, основан на алгоритме нахождения граф-подграф изоморфизма на взвешенных графах [9] с модификациями, направленными на возможность распределения нескольких узлов параллельной задачи на один вычислительный узел сети.

1 ПОСТАНОВКА ЗАДАЧИ

Пусть даны два графа G_N и G_T . Пусть граф $G_N = (E_N, V_N, I_N, J_N)$ — представляет собой граф сети, где E_N — множество физических сетевых соединений (линий связи), V_N — множество вычислительных узлов сети (процессоров), I_N — множество весов, приписанных ребрам графа, представляющих собой пропускную способность сетевых соединений, измеренную в Mbps или Gbps и J_N — множество весов, приписанных вычислительным узлам графа сети, представляющих собой производительность процессоров, измеренную в MIPS или MFLOPS. Пусть граф $G_T = (E_T, V_T, I_T, J_T)$ — представляет собой граф параллельной задачи, для которой необходимо зарезервировать часть вычислительных ресурсов сети до начала вычислений, где E_T — множество ребер графа, соответствующих сетевым взаимодействиям между модулями параллельной программы, V_T — множество вершин, соответствующих модулям параллельной программы, I_T — множество весов, приписанных ребрам графа задачи, выражающих потребность в пропускной способности сети для данного сетевого взаимодействия, измеренную в Mbps или Gbps, J_T — множество весов, приписанных вершинная графа, соответствующих потребности модулей в вычислительной производительности процессорных узлов, измеренной в MIPS или MFLOPS.

В терминах определений, введенных выше, задача резервирования распределенных вычислительных ресурсов формализуется следующим образом. Необходимо найти подстановку $\varphi: V_T \rightarrow V_N$, такую, что для каждой пары вершин (v_i, v_j) , если $(v_i, v_j) \in E_T \Rightarrow (\varphi(v_i), \varphi(v_j)) \in E_N$ и для $(i_i, i_j) \in I_T \Rightarrow (\varphi(i_i), \varphi(i_j)) \in I_N$, и $j_i \in J_T \Rightarrow \varphi(j_i) \in J_N$.

Задача оптимизации резервирования ресурсов по критерию минимального количества задействованных вычислительных узлов сети означает, что среди всех $\varphi: V_T \rightarrow V_N$, выбирается та подстановка, при которой $|V_N| \rightarrow \min$.

2 АЛГОРИТМ РЕЗЕРВИРОВАНИЯ РЕСУРСОВ

2.1 Структура алгоритма

Как следует из постановки задачи, проблема схожа с задачей нахождения граф-подграф изоморфизма на взвешенных графах, стой лишь разницей, что допускается совмещение нескольких вершин графа задачи с одной и той же вершиной графа сети. Веса, приписанные вершинам и ребрам графа сети, являются ограничениями «сверху», налагаемыми на возможное решение.

Описание алгоритма удобно производить в терминах поиска в пространстве состояний. Каждому состоянию соответствует частичная подстановка $\phi(s)$, содержащая лишь часть вершин графов, которые уже были совмещены.

Алгоритм состоит из предварительной и основной части.

2.2 Предварительная часть алгоритма

Основной задачей предварительной части алгоритма является выполнение всех проверок, которые основаны на данных известных до начала работы алгоритма и не базируются на информации о уже совмещенных вершинах частичной подстановки $\phi(s)$. На их основе формируется матрица возможных совмещений. Так же в предварительной части алгоритма производится сортировка вершин графов.

Центральным элементом предварительной части алгоритма является матрица возможных совмещений. Это бинарная таблица $M_{i,j}$, каждая ячейка которой хранит агрегированное значение о возможности либо не возможности совмещения вершин $V_{N,i}$ и $V_{T,j}$. Значения матрицы формируются на основании предварительных проверок и могут принимать значения «истина» или «ложь», в зависимости от того перспективно или нет совмещение соответствующих вершин в основной части алгоритма на основании проверок, выполненных в предварительной части алгоритма.

В представленном алгоритме реализовано несколько базовых проверок, формирующих матрицу возможных совмещений $M_{i,j}$. Если результат проверки условия «истина», то значение матрицы, соответствующее проверяемым вершинам $V_{N,i}$ и $V_{T,j}$, устанавливается равным «ложь».

Пусть $|V_X|$ — степень вершины V_X . Тогда условие, основанное на сравнении степеней вершин, выглядит следующим образом:

$$|V_{N,i}| < |V_{T,j}| \Rightarrow M_{i,j} = \text{false}.$$

Условие, основанное на сравнении весов, приписанных вершинам графов:

$$J_{N,i} < J_{T,j} \Rightarrow M_{i,j} = \text{false}.$$

Пусть V_{in_X} — подмножество, содержащее все вершины, связанные с вершиной V_X входящими ребрами, тогда условие, основанное на этом определении:

$$|V_{in_{N,i}}| < |V_{in_{T,j}}| \Rightarrow M_{i,j} = \text{false}.$$

Пусть I_{in_X} — подмножество весов, приписанных ребрам, входящим в вершину V_X . Условие, основанное на значении I_{in_X} :

$$\sum I_{in_{N,i}} < \sum I_{in_{T,j}} \Rightarrow M_{i,j} = \text{false}.$$

Пусть V_{out_X} — подмножество, содержащее все вершины, связанные с вершиной V_X исходящими ребрами, тогда условие, основанное на этом определении:

$$|V_{out_{N,i}}| < |V_{out_{T,j}}| \Rightarrow M_{i,j} = \text{false}.$$

Пусть I_{out_X} — подмножество весов, приписанных ребрам, исходящим из вершины V_X . Условие, основанное на значении I_{out_X} :

$$\sum I_{out_{N,i}} < \sum I_{out_{T,j}} \Rightarrow M_{i,j} = \text{false}.$$

Приведенный выше список простейших условий, не требующих сложных вычислений, тем не менее, часто приводит к значительному сокращению области поиска основной, переборной, части алгоритма. Но в дополнение к рассмотренным выше, предлагается использование условий, основанных на волновом разложении графов, более подробно описанном в работе [10]. Метод волнового разложения графов формирует много структурной информации о графе, которая может быть использована для построения более эффективных условий ограничения области поиска, используемых в предварительной части алгоритма. Наиболее важной частью волнового разложения графов, используемой в данном алгоритме, является подграф окружения вершины. Это подграф, состоящий из вершин, находящихся на расстоянии k ребер от исходной, где k — параметр, определяющий как много соседних вершин будет задействовано при формировании подграфа окружения. Другим важным параметром волнового разложения, использованным в работе, является число вершин, вошедших в каждую волну разложения графа, начиная с заданной вершины.

Пусть $W_{X,Y,k}$ — множество вершин, вошедших в волновое разложение графа X , начиная с вершины Y с максимальным реберным расстоянием до вершин k . Условие, основанное на этом определении:

$$|W_{N,i,k}| < |W_{T,j,k}| \Rightarrow M_{i,j} = \text{false},$$

где $k = 1 \dots |W_{T,j}|$.

Пусть $U_{X,Y,k}$ — множество весов, соответствующих вершинам множества $W_{X,Y,k}$:

$$\sum U_{N,i,k} < \sum U_{T,j,k} \Rightarrow M_{i,j} = \text{false},$$

где $k = 1 \dots |W_{T,j}|$.

Пусть $Q_{X,Y,k}$ — подмножество ребер, которые вошли в волновое разложение графа X , начиная с вершины Y с максимальным расстоянием до вершин в k ребер и пусть $|Q_{X,Y,k}|$ — количество ребер в подмножестве $Q_{X,Y,k}$. Тогда формируется условие:

$$|Q_{N,i,k}| < |Q_{T,j,k}| \Rightarrow M_{i,j} = \text{false},$$

где $k = 1 \dots |Q_{T,j}|$.

Пусть $R_{X,Y,k}$ — множество весов, приписанных ребрам, вошедшим в подмножество $Q_{X,Y,k}$:

$$\sum R_{N,i,k} < \sum R_{T,j,k} \Rightarrow M_{i,j} = \text{false},$$

где $k = 1 \dots |Q_{T,j}|$.

Условия, основанные на волновом разложении графов, могут быть просчитаны для некоторого константного значения k или для всех возможных значений k . Во втором случае условия будут иметь большую вычислительную сложность, но так же лучше ограничивать область поиска переборной части алгоритма. Поскольку все условия предварительной части алгоритма имеют полиномиальную вычислительную сложность, то для больших графов имеет смысл выполнить все перечисленные выше условия. Некоторое дополнительное время, затраченное на условия предварительной части алгоритма, приведет к более значительному сокращению времени основной, переборной, не полиномиальной, части.

Еще одним важным действием, выполняемым в рамках предварительной части алгоритма, является сортировка вершин графов. Целью является перестановка вершин графов таким образом, чтобы получить более сильное ограничивающее условие, основанное на частичных подстановках, получаемых в процессе работы основной, переборной, части алгоритма. Идея основана на том, чтобы вначале совмещать вершины, имеющие больше внутренних реберных связей, и, как результат, формирующих более сильное ограничивающее условие переборной части. В программе сортируются только вершины графа G_T , описывающего требования параллельной задачи к распределенным ресурсам, в то время, как порядок следования вершин графа сети G_N остается без изменений, т. к. соответствие вершин графа G_N вершинам графа G_T будет устанавливаться уже в переборной части алгоритма.

Пусть $T_{T,i}$ — число ребер, инцидентных вершинам, имеющим индекс меньше, чем i , и пусть $P_{T,i} = \sum M_j$ — число возможных совмещений, доступных для вершины $V_{T,i}$ графа G_T . Тогда порядок следования вершин графа G_T определяется следующими выражениями:

$$V_{T,i} = V_{T,k},$$

где $T_{T,k} = \min(T_{T,j})$ для $j = (i+1) \dots |V_T|$.

В случае, если $T_{T,i} = T_{T,k}$ применяется другое условие:

$$V_{T,i} = V_{T,k},$$

где $P_{T,k} = \min(P_{T,i}, P_{T,j})$.

2.3 Переборная часть алгоритма

Переборная часть алгоритма объединяет в себе все действия, направленные на поиск полной подстановки ϕ , являющейся решением задачи. Алгоритм представлен рекурсивной функцией поиска в пространстве состояний, которая на каждом шаге вложенности генерирует новую частичную подстановку $\phi_{i+1}(s)$ из предыдущей частичной подстановки $\phi_i(s)$, путем добавления одной вершины в частичную подстановку.

Начальное состояние $\phi_0(s) = 0$. На каждом шаге функция перебирает все вершины графа G_N , которые имеют пометку «истина» в строке матрицы возможных совмещений, соответствующей вершине $V_{T,i}$, используемой на текущем уровне вложенности алгоритма i . Для каждой вершины $V_{N,j}$, алгоритм производит несколько проверок, имеющих целью удостовериться, что все необходимые условия для включения новой вершины в частичную подстановку $\phi_{i+1}(s)$ выполнены. Далее приведен полный список условий, используемых в переборной части алгоритма с более подробным описанием.

Наиболее важным является условие, основанное на значении соответствующего элемента матрицы возможных совмещений M :

$$M_{i,j} = \text{true}.$$

Пусть $T_{N,i}$ — число ребер, связывающих вершину $V_{N,i}$ с как минимум одной вершиной из вошедших в частичную подстановку $\phi_i(s)$:

$$T_{N,i} > T_{N,j}.$$

Условие, основанное на множестве ребер, входящих в текущую частичную подстановку:

$$(v_i, v_k) = (\phi_i(v_i), \phi_i(v_k)),$$

где $k = 1 \dots i$.

Условие, основанное на весах этих ребер:

$$(i_i, i_k) \geq (\phi_i(i_i), \phi_i(i_k)),$$

где $k = 1 \dots i$.

Если все условия, описанные выше, истинны, то очередная пара вершин $(V_{T,i}, V_{N,j})$ будет добавлена в частичную подстановку $\varphi_i(s)$ и будет сформирована новая частичная подстановка $\varphi_{i+1}(s)$.

В этом списке последние два условия следуют из определения задачи и являются формальными строгими условиями, позволяющими очередной вершине войти в частичную подстановку. Когда эти условия последовательно применяются ко всем парам вершин, входящим в частичные подстановки $\varphi_i(s)$, для всех значений i от 1 до $|V_T|$, в результате формируется суммарное условие, прямо следующее из постановки задачи и гарантирующее, что полученная полная подстановка будет удовлетворять всем условиям, оговоренным в постановке задачи, т. е. будет являться решением задачи.

Перебор вариантов производится методом поиска в глубину.

2.4 Поиск оптимальной области резервирования

Для поиска оптимальной области резервирования используется особенность реализации алгоритма, заключающаяся в том, что перебор возможных подстановок $\varphi(s)$ можно продолжать, после нахождения первой полной подстановки и перебрать все полные подстановки, которых может быть несколько. Из всех полных подстановок, которые будут найдены алгоритмом, выбирается та, что соответствует критерию оптимальности.

Пусть подстановка $\varphi(s)$ ставит в соответствие вершинам графа задачи G_T часть вершин графа сети G_N : $\varphi(s): V_T \rightarrow V'_N \in V_N$.

Оптимальная область резервирования по критерию числа задействованных вычислительных узлов определяется следующим условием, налагаемым на полную подстановку:

$$\varphi(s): V_T \rightarrow V'_N \in V_N, \text{ где } |V'_N| \rightarrow \min.$$

Практический смысл оптимизации области резервирования по числу задействованных узлов заключается в том, что используется минимально возможное число физических компьютеров, что позволяет выключить остальные компьютеры или использовать их для совершенно других задач. Так же это упрощает администрирование, предоставляя возможность выдать права пользователям только на определенных компьютерах и передать минимально необходимое, но достаточное количество компьютеров в пользование или аренду.

Оптимальная область резервирования по критерию максимальной загруженности вычислительных узлов сети определяется следующим условием:

$$\varphi(s): V_T \rightarrow V'_N \in V_N,$$

при этом

$$\sum (J'_N - J_T) \rightarrow \min,$$

где J'_N — множество весов, приписанных вершинам графа сети V'_N , задействованным в полной подстановке.

Практический смысл оптимизации по критерию максимальной загруженности вычислительных узлов сети в наиболее полном использовании имеющихся вычислительных мощностей. Если такая оптимизация производится в масштабах всей сети для составного графа содержащего все задачи, выполняемые в сети, то как следствие получается глобальная оптимизация распределения зарезервированных областей, если же критерий оптимальности используется при добавлении каждой следующей задачи в сеть, то оптимизация зарезервированных областей производится жадным алгоритмом, который так же дает хорошие результаты оптимизации, без необходимости увеличения размерности задачи до масштабов всей сети.

ЗАКЛЮЧЕНИЕ

В работе представлен переборный алгоритм оптимизации резервирования распределенных вычислительных ресурсов по критериям числа и максимальной загруженности задействованных вычислительных узлов. В алгоритме использованы составные условия, ограничивающие область поиска переборной части алгоритма, в том числе, основанные на волновом разложении графов, значительно сокращающие время поиска оптимального решения.

Дальнейшие усилия будут приложены к оптимизации областей резервирования по критериям оптимального использования пропускной способности сети и возможности применения алгоритма для сверхбольших сетей, посредством разбиения задачи оптимизации на несколько подзадач, меньшей размерности.

ПЕРЕЧЕНЬ ССЫЛОК

1. *I. Foster, A. Roy, and V. Sander.* A quality of service architecture that combines resource reservation and application adaptation // Proceedings of the 8th International Workshop on Quality of Service (IWQOS). – Pittsburgh, PA. – June 2000. – P. 181–188.
2. *Sander V.* A Metacomputer Architecture Based on Cooperative Resource Management // Proceedings of High Performance Computing and Networking Europe 1997 (HPCN 1997). – Wien. – April (1997). – P. 28–30.
3. *F. Berman, R. Wolski, S. Figueira, J. Schopf, and G. Shao.* Application level scheduling on distributed heterogeneous networks. // Proceedings of Supercomputing – 1996. – P. 39.
4. *Zeng Wandan, Chang Guiran, Zhang Dengke, Zheng Xiu-ying.* G-RSVP: A Grid Resource Reservation Model // First International Conference on Semantics, Knowledge and Grid (SKG'05). – Guilin Guanxi, China – 2005. – P. 79.

5. L. P. Cordella, P. Foggia, C. Sansone, M. Vento. Performance evaluation of the VF Graph Matching Algorithm // Proc. of the 10th ICIAP, IEEE Computer Society Press. – 1999. – P. 1172–1177.
6. Bunke H., Vento M. Benchmarking of graph matching algorithms. // Proceedings of the 2nd Workshop on Graph-based Representations. – Haindorf. – 1999. – P. 109–114.
7. Cordella L. P., Foggia P., Sansone C., Vento M. An improved algorithm for matching large graphs. // Proc. of the 3rd IAPR TC-15 Workshop on Graph-based Representations in Pattern Recognition. – Italy. – 2001. – P. 149–159.
8. P. Foggia, C. Sansone, M. Vento. A performance comparison of five algorithms for graph isomorphism. // Proc. of Ильяшенко М. Б. Разработка и исследование параллельного алгоритма проверки граф-подграф изоморфизма. // Радиотехника. Информатика. Управление. – 2006. – № 1. – С. 63–69.
9. Пинчук В. П. Табличные инварианты на графах и их применение // Кибернетика и системный анализ. – 2001. – № 4. – С. 33–45.

Надійшла 20.07.07
Після доробки 26.10.07

В роботі запропоновано алгоритм оптимального резервування обчислювальних ресурсів, що здатен виконувати пошук областей резервування, згідно критеріям оптимальності по числу та максимальному завантаженню задіяних обчислювальних вузлів. Наводиться детальний опис алгоритму та набору попередніх умов, що звужують область пошуку у переборній частині алгоритму.

This paper presents algorithm for optimal distributed resources reservation, that allow to find reservation areas according to minimal amount of used computational nodes optimization criteria and maximal productivity usage of equipped computational nodes criteria. Paper contains detailed algorithm description, including set of preliminary conditions that reduce computational complexity of enumerating part of algorithm.

УДК:681.142.2

С. В. Курапов

КОНСТРУКТИВНЫЙ АЛГОРИТМ ДЛЯ РАСКРАСКИ КУБИЧЕСКИХ ГРАФОВ

В данной работе представлен конструктивный алгоритм для раскраски кубических графов с применением теории вращения вершин. Алгоритм основан на фундаментальной теореме Петерсена выделения 1-факторов и 2-факторов в кубических графах.

ВВЕДЕНИЕ

В 1913 г. Биркгоф ввел понятие неприводимого графа и доказал ряд теорем о свойствах таких графов. Пользуясь этими результатами, американский математик Франклин доказал, что гипотеза четырех красок верна для всех плоских графов с числом вершин до 31. Французский ученый Майер довел это число до 96. Хееш в 1969 г. свел вопрос о справедливости гипотезы четырех красок к исследованию достаточно большого так называемого неустраняемого множества конфигураций. Хеешу удалось доказать, что после первого шага метода нейтрализации остается около 8900 положительных конфигураций, большинство из которых неприводимы [1]. В 1977 году доказательство гипотезы четырех красок было наконец получено К. Аппелем и У. Хакеном (Appel, Haken) последователями Хееша, и опубликовано в двух статьях [2].

Значительную часть рутинных проверок выполнил компьютер, и это революционное нововведение в сложившуюся практику дедуктивных рассуждений в чис-

той математике служит основанием для некоторого естественного скептицизма по отношению к данному доказательству, и по сей день.

Читатель данного доказательства, должен разобратся в 50 страницах текста и диаграмм, 85 страницах с почти 2500 дополнительными диаграммами, 400 страницами микрофишей, содержащими еще диаграммы, а также тысячи отдельных проверок утверждений, сделанных в 24 леммах основного текста. Вдобавок читатель узнает, что проверка некоторых фактов потребовала 1200 часов компьютерного времени, а при проверке вручную потребовалось бы гораздо больше. Статьи устрасшаюци по стилю и длине, и немногие математики прочли их сколько-нибудь подробно [3].

В литературе [4] приведено мнение большинства специалистов о том, что нет более быстрого метода раскрашивания карты, чем перебор всех вариантов. Там же приведена Фортран-программа на основе алгоритма полного перебора. Следует также заметить, что во всех этих подходах к решению проблемы, отсутствует понятие планарности, хотя гипотеза доказывается для плоских графов, отсутствует также и конструктивный алгоритм раскраски. Поэтому, исследования направленные на создание более эффективных алгоритмов раскраски карты актуальны.

1 ПОСТАНОВКА ЗАДАЧИ

Будем рассматривать конечные трехсвязные неориентированные графы без петель, мостов и кратных ребер. Пусть $G(X, U; P)$ – граф с множеством вершин $X = \{x_1, x_2, \dots, x_n\}$ и ребер $U = \{u_1, u_2, \dots, u_m\}$, где n – количество вершин графа и m – количество ребер графа G , а P – трехместный предикат ставящий в соответствие одному ребру две вершины записанный в виде матрицы смежностей или матрицы инцидентий. Пусть \mathfrak{F} – множество всех суграфов этого графа. Относительно операции сложения

$$(X, U_1; P) \oplus (X, U_2; P) = (X, (U_1 \cup U_2) \setminus (U_1 \cap U_2); P) \quad (1)$$

это множество, как известно, образует абелеву 2-группу, которую можно рассматривать как векторное пространство над полем из двух элементов $GF(2)$. Размерность этого пространства, называемого пространством суграфов графа G , конечно и равно m ($\dim \mathfrak{F} = m$). В качестве базиса этого пространства можно выбрать множество односторонних суграфов (u_1, u_2, \dots, u_m) . Тогда в этом базисе каждому элементу Y пространства \mathfrak{F} однозначно сопоставляется последовательность координат (a_1, a_2, \dots, a_m) , где $a_i \in \{0, 1\}$. При этом оказывается, что ребро u_i входит в суграф Y , если $a_i = 1$, и не входит в данный суграф в противном случае. В дальнейшем для удобства будем отождествлять пространство суграфов \mathfrak{F} и его координатное пространство.

Пространство суграфов, в свою очередь, содержит два подпространства: подпространство разрезов \mathfrak{F}_s и подпространство циклов \mathfrak{F}_c . Базис подпространства разрезов может состоять из так называемых единичных разрезов. Где под единичным разрезом подразумевается множество ребер инцидентных данной вершине. Размерность этого подпространства определяется

$$\eta(G) = n - 1. \quad (2)$$

Подпространство циклов составляют квазициклы. Напомним, что суграф называется квазициклом, если все его вершины имеют четную валентность. Множество графа G образует, как легко можно видеть, подпространство квазициклов \mathfrak{F}_c пространства \mathfrak{F} . Известно, что размерность подпространства \mathfrak{F}_c совпадает с цикломатическим числом

$$v(G) = m - n + 1 \quad (3)$$

графа G , а порядок группы \mathfrak{F}_c равен $2^{v(G)}$. Характерным квазициклом является простой цикл графа у которого валентность равна двум. Однако, существует подмножество простых циклов с мощностью еще

меньшей, чем подмножество простых циклов, обладающее определенными характерными свойствами. С этой целью вводится определение τ -цикла графа (единичного цикла).

Определение [5]: единичным циклом (τ -циклом) называется простой цикл, между двумя любыми несмежными вершинами которого в соответствующем графе не существует маршрутов меньшей длины, чем маршруты, принадлежащие данному циклу.

Подмножество, состоящее из единичных циклов, будем называть подмножеством единичных циклов и обозначать C_E .

Понятие τ -цикла графа G тесно связано с минимальными $(s - t)$ маршрутами графа на основании теоремы связности графа принадлежащей Менгеру [6] и верна следующая теорема.

Теорема [5]. Для любого трехсвязного графа без петель, мостов и кратных ребер линейное подпространство квазициклов содержит базис из τ -циклов.

Из всего множества связных графов можно выделить подмножество планарных графов, для которых существует базис состоящий из единичных циклов графа и удовлетворяющий нулевому значению функционала Мак-Лейна для этого базиса [5]

$$F(C) = \sum_{i=1}^m S_i^2 - 3 \sum_{i=1}^m S_i + 2m, \quad (4)$$

где S_i – количество базисных единичных циклов проходящих по i -му ребру.

Для любого плоского графа исходя из выбранного базиса единичных циклов определяется цикл называемый ободом графа, как сумма независимых векторов подпространства циклов C

$$C_{об} = \sum_{i=1}^{v(G)} C_i. \quad (5)$$

Имея множество независимых векторов подпространства \mathfrak{F}_c и обод графа $C_{об}$ строится вращение вершин δ , определяемое как циклический порядок (или циклическая перестановка) всех ребер инцидентных вершине введенное Г. Рингелем [7].

И тогда запись (G, δ) , будет обозначать рисунок графа G на плоскости с некоторым вращением вершин индуцированное базовой системой циклов и ободом.

Возможно и обратное, заданное вращение индуцирует систему циклов удовлетворяющую нулевому значению функционала Мак-Лейна.

Соответственно для множества непланарных графов не удается выделить базис состоящий из единичных циклов удовлетворяющий нулевому значению функционала Мак-Лейна. Для непланарных графов существу-

ет два непересекающихся подмножества относительно вращения вершин.

Первое подмножество представляет собой вращение вершин индуцированное подмножеством циклов из базиса подпространства C_E для которого функционал Мак-Лейна равен нулю, и обода для этого подмножества. Здесь по каждому ребру проходит ровно два цикла, но мощность этого подмножества меньше чем цикломатическое число графа. Этот случай подробно описан в монографии Г. Рингеля [7] и тесно связан с понятием рода поверхности и рода графа.

Второе подмножество представляет собой совокупность вращений вершин характеризующих рисунки графов, тоже индуцированные циклами, совокупность которых удовлетворяет нулевому значению функционала Мак-Лейна, но в данной совокупности существуют замкнутые ориентированные маршруты дважды проходящие по одному ребру, причем в различных направлениях. Это так называемые ω -маршруты.

Например, для плоского графа K_4 с вращением (рис. 1, а) имеем следующую систему индуцированных τ -циклов: $\{1, 2, 4\}$, $\{1, 3, 5\}$, $\{2, 3, 6\}$ и $C_{06} = \{4, 5, 6\}$.

Для неплоского графа K_5 с вращением (представленного на рис. 2, а) имеем следующую систему индуцированных циклов $c_1 = \{1, 3, 2\}$, $c_2 = \{1, 8, 5\}$, $c_3 = \{2, 9, 10\}$, $c_4 = \{3, 4, 9, 6, 7, 4, 5, 6\}$, $c_{06} = \{7, 8, 10\}$. Как видно c_4 включает в себя дважды ребро 4 и ребро 6, это ω -маршрут (см. рис. 2, б).

Рассмотрение свойств перечисленных выше вращений вершин выходит за рамки данной работы, здесь будем рассматривать только плоские графы с вращением вершин индуцирующие базовые единичные циклы удовлетворяющие нулевому значению функционала Мак-Лейна.

Плоский граф с треугольными гранями может быть построен из любого плоского графа путем введения новых ребер, в этом случае полученный граф называется максимально плоским графом.

Если в максимально плоском графе правильно раскрасить ребра, принадлежащие треугольным граням в три цвета (см. рис. 3), то тогда удастся раскрасить вершины данного плоского графа в четыре цвета согласно теореме Тэйта.

Теорема. (Тэйт) [1]. Необходимым и достаточным условием раскраски максимально плоского графа G с треугольными гранями четырьмя цветами состоит в том, чтобы хроматический класс двойственного кубического графа H был равен трем.

Если поставить в соответствие треугольным граням вершины в максимально плоском графе, то построится плоский кубический граф H двойственный максимально плоскому графу.

Переход от максимально плоского графа G с треугольными гранями к двойственному кубическому графу H осуществляется путем замены матрицы циклов

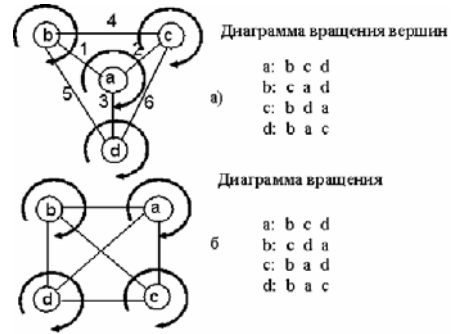


Рисунок 1 – Различные диаграммы вращений для различных рисунков графа G

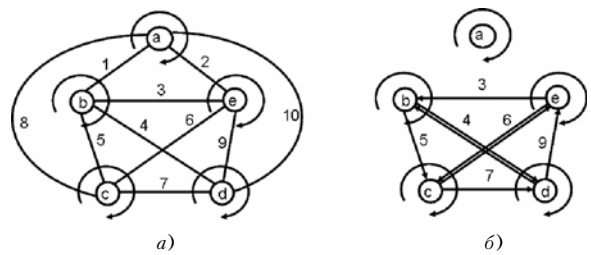


Рисунок 2 – Вращение вершин в непланарном графе G: а – вращение вершин; б – ω -маршрут

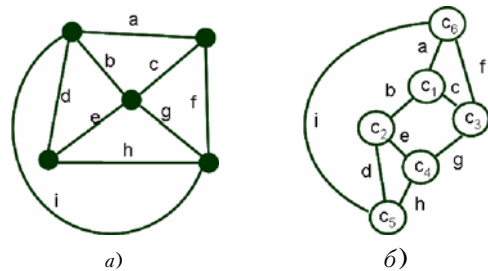


Рисунок 3 – Плоские графы: а – максимально плоский граф G; б – двойственный кубический граф H

максимально плоского графа G с треугольными гранями на матрицу инцидентий кубического графа H .

Следует заметить, что единичные циклы (независимые вектора подпространства циклов ξ_c) в кубическом графе H дуальны вращению вершин в максимально плоском графе G . Вращение вершин в кубическом графе H дуально базовым треугольным циклам в максимально плоском графе G .

Например, для максимально плоского графа G (см. рис. 3, а) треугольные циклы запишутся в виде:

$$C_1 = \{a, b, c\}; C_2 = \{b, d, e\}; C_3 = \{c, f, g\}; C_4 = \{e, g, h\}; C_5 = \{d, h, i\}; C_6 = \{a, f, i\}.$$

Вращение вершин в максимально плоском графе G (см. рис. 3, а) запишется в виде:

- 1: $a b d i$
- 2: $a f c$
- 3: $g f i h$
- 4: $e d h$
- 5: $b c g e$

Для кубического плоского графа H (см. рис. 3, б) единичные циклы запишутся в виде:

$$C_{H_1} = \{a, b, d, i\}; C_{H_2} = \{a, f, c\}; C_{H_3} = \{g, f, i, h\};$$

$$C_{H_4} = \{e, d, h\}; C_{H_5} = \{b, c, g, e\}.$$

Вращение вершин в кубическом графе H (см. рис. 3, а) запишется в виде:

- $C_1: a c b$
- $C_2: b e d$
- $C_3: c f g$
- $C_4: e g h$
- $C_5: d h i$
- $C_6: a i f$

Будем говорить, что n -фактор – это регулярный су-граф степени n . Тогда для кубических графов справедлива теорема Петерсена.

Теорема (Петерсен) [6]. Любой кубический граф, не содержащий мостов, можно представить в виде суммы 1-фактора и 2-фактора.

Будем называть кубический граф, у которого хроматический класс равен трем, правильно раскрашенным кубическим графом. Обозначим ребра, раскрашенные в красный цвет, буквой R (или цифрой 3), ребра, раскрашенные в зеленый цвет, буквой Q (или цифрой 2), ребра, раскрашенные в синий, буквой B (или цифрой 1). Тогда переокраска ребер осуществляется по следующему закону сложения.

+	0	B (1)	Q (2)	R (3)
0	0	B (1)	Q (2)	R (3)
B (1)	B (1)	0	R (3)	Q (2)
Q (2)	Q (2)	R (3)	0	B (1)
R (3)	R (3)	Q (2)	B (1)	0

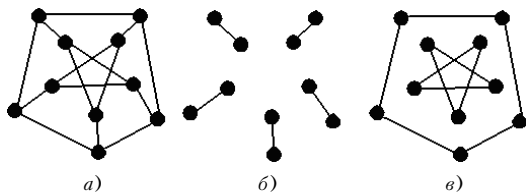


Рисунок 4 – Кубический граф Петерсена:
 а – граф Петерсена; б – 1-фактор графа Петерсена;
 в – 2-фактор графа Петерсена

В правильно раскрашенном кубическом графе обязательно образуется цветная конфигурация состоящая из цветных 2-факторов имеющих диски (простые циклы) только четной длины и цветных 1-факторов. Для такой цветной конфигурации справедливы следующие соотношения:

$$R_c \oplus Q_c \oplus B_c = \emptyset; \tag{6}$$

$$R_f \oplus Q_f \oplus B_f = H; \tag{7}$$

$$R_f \oplus R_c = H; \tag{8}$$

$$Q_f \oplus Q_c = H; \tag{9}$$

$$B_f \oplus B_c = H. \tag{10}$$

Здесь индекс c обозначает 2-факторы, индекс f обозначает 1-факторы, а H означает исходный кубический граф.

Любой цветной гамильтоновский квазицикл (2-фактор) состоит из дисков. Если диск один, то это уже гамильтонов цикл.

Таким образом, задача раскраски ребер кубического графа H свелась к задаче нахождения 2-фактора имеющего диски только четной длины.

2 АЛГОРИТМ ПОЛУЧЕНИЯ 2-ФАКТОРА КУБИЧЕСКОГО ГРАФА С ЧЕТНЫМИ ДЛИНАМИ ДИСКОВ

Для построения алгоритма раскраски плоских кубических графов рассмотрим свойства плоского биквадратного графа B_3 как дуально-реберного к кубическому графу H (см. рис. 5, где вершины закрашены более темным цветом, а ребра пунктирные линии).

Матрица смежностей биквадратного графа B_3 образуется как произведение транспонированной матрицы инцидентий и матрицы инцидентий для кубического графа H :

$$A = B^T * B. \tag{11}$$

Здесь сложение и умножение производится по закону:

$$1 + 1 = 0; 1 + 0 = 1; 0 + 0 = 0;$$

$$1 * 1 = 1; 1 * 0 = 0. \tag{12}$$

Часть единичных циклов, записанных в виде циклического обхода вершин, в биквадратном графе B_3 дуальна треугольным циклам максимально плоского графа G . Такие циклы в биквадратном графе B_3 будем называть базовыми треугольными циклами. Другую

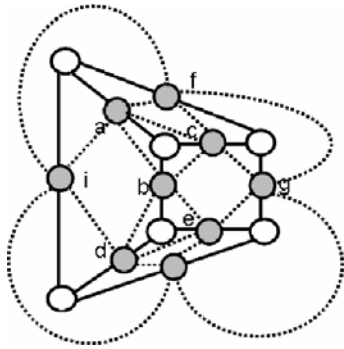


Рисунок 5 – Построение плоского биквадратного графа B_3 как дуально-реберного к плоскому кубическому графу H , изображенному на рис. 3, а

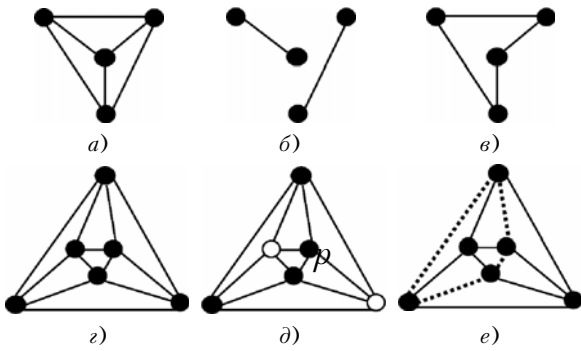


Рисунок 6 – Двойственность кубического графа H и биквадратного графа B_3 :

a – кубический граф H ; $б$ – 1-фактор графа H ;
 $в$ – 2-фактор графа H ; $г$ – биквадратный граф B_3 ,
 дуальный к графу H ; $д$ – вершины графа B_3 ,
 дуальные 1-фактору графа H ; $е$ – циклы графа B_3 ,
 дуальные 2-фактору графа H

часть единичных циклов, записанных в виде циклического обхода вершин, в биквадратном графе B_3 дуальна базовым единичным циклам плоского кубического графа H .

Например, для биквадратного графа B_3 (см. рис. 5) базовые треугольные циклы записаны в виде циклических кортежей:

$$C_1 = \langle b, a, c \rangle; C_2 = \langle b, e, d \rangle;$$

$$C_3 = \langle c, f, g \rangle; C_4 = \langle e, g, h \rangle;$$

$$C_5 = \langle d, h, i \rangle; C_6 = \langle a, i, f \rangle.$$

Другая часть базовых единичных циклов биквадратного графа B_3 запишется в следующем виде:

$$C_7 = \langle a, b, d, i \rangle; C_8 = \langle a, f, c \rangle;$$

$$C_9 = \langle g, f, i, h \rangle; C_{10} = \langle e, d, h \rangle;$$

$$C_{11} = \langle b, c, g, e \rangle.$$

Вращение вершин биквадратного графа определится как сумма по модулю 2 кортежей базовых треугольных циклов с учетом цикличности:

$$\text{Вращение вершины } \delta_a = \langle i, f, a \rangle \oplus \langle a, c, b \rangle = \langle i, f, c, b \rangle a: i f c b;$$

$$\text{Вращение вершины } \delta_b = \langle e, d, b \rangle \oplus \langle b, a, c \rangle = \langle e, d, a, c \rangle b: e d a c;$$

$$\text{Вращение вершины } \delta_c = \langle b, a, c \rangle \oplus \langle c, f, g \rangle = \langle b, a, f, g \rangle c: b a f g;$$

$$\text{Вращение вершины } \delta_d = \langle b, e, d \rangle \oplus \langle d, h, i \rangle = \langle b, e, h, i \rangle d: b e h i;$$

$$\text{Вращение вершины } \delta_e = \langle d, b, e \rangle \oplus \langle e, g, h \rangle = \langle d, b, g, h \rangle e: d b g h;$$

$$\text{Вращение вершины } \delta_f = \langle a, i, f \rangle \oplus \langle f, g, c \rangle = \langle a, i, g, c \rangle f: a i g c;$$

$$\text{Вращение вершины } \delta_g = \langle c, f, g \rangle \oplus \langle g, h, e \rangle = \langle c, f, h, e \rangle g: c f h e;$$

$$\text{Вращение вершины } \delta_h = \langle e, g, h \rangle \oplus \langle h, i, d \rangle = \langle e, g, i, d \rangle h: e g i d;$$

$$\text{Вращение вершины } \delta_i = \langle d, h, i \rangle \oplus \langle i, f, a \rangle = \langle d, h, f, a \rangle i: d h f a.$$

Так как биквадратный граф B_3 образуется как реберный граф к кубическому графу, то имеет место двойственность (см. рис. 6). Циклы биквадратного графа B_3 дуальные 2-фактору графа H будем называть 2-дуальные циклы и обозначать – C_w .

Как мы убедились ранее, 2-факторы в кубическом графе H и 2-дуальные циклы C_w в биквадратном графе B_3 могут состоять из дисков (представляющих собой простые циклы) нечетной длины. Для правильной раскраски необходимо, чтобы все диски представляли собой циклы только четной длины. Рассмотрим следующую операцию, которая позволяет переводить диски нечетной длины в диски четной длины.

Рассмотрим кубический граф H и дуально-реберный к нему биквадратный граф B_3 представленные на рис. 6. Выделим в биквадратном графе B_3 2-дуальный цикл C_w , состоящий из двух нечетных дисков.

$$C_w = \{1, 3, 6, 11, 16, 26, 27, 28, 29, 30\}.$$

Выделим в данном графе произвольный цикл четной длины, включающий только по одному ребру базовых треугольных циклов, не пересекающийся по ребрам с 2-дуальным циклом и соединяющий диски 2-дуальных циклов. На рис. 7, а это цикл $\{20, 21, 23, 25\}$. Относительно выбранного цикла построим цикл, включая в него оставшиеся два ребра в выбранных базовых треугольных циклах. Построенный таким образом цикл будем называть четным соединяющим циклом в биквадратном графе B_3 и обозначать C_v . На рис. 7, б представлен четный соединяющий цикл $C_v = \{6, 7, 16, 17, 22, 24, 28, 30\}$, выделенный сплошными линиями.

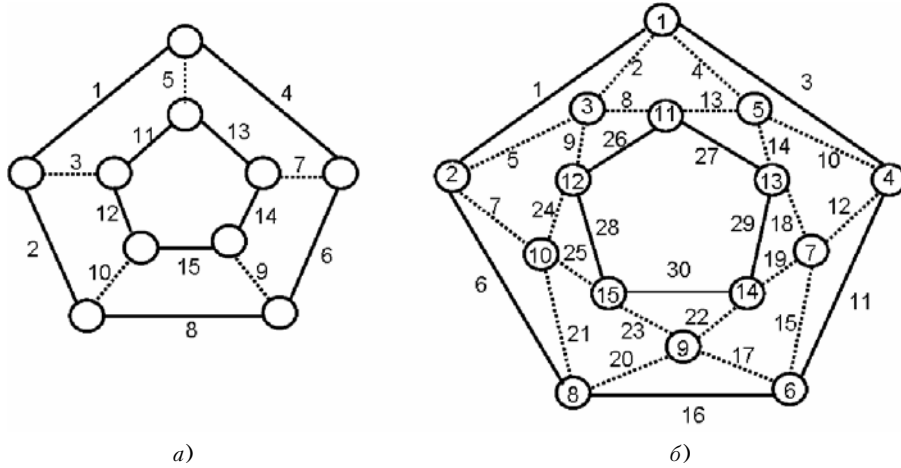


Рисунок 7 – Кубический граф H и его 2-фактор (а), биквадратный граф B_3 и его 2-дуальный цикл (б)

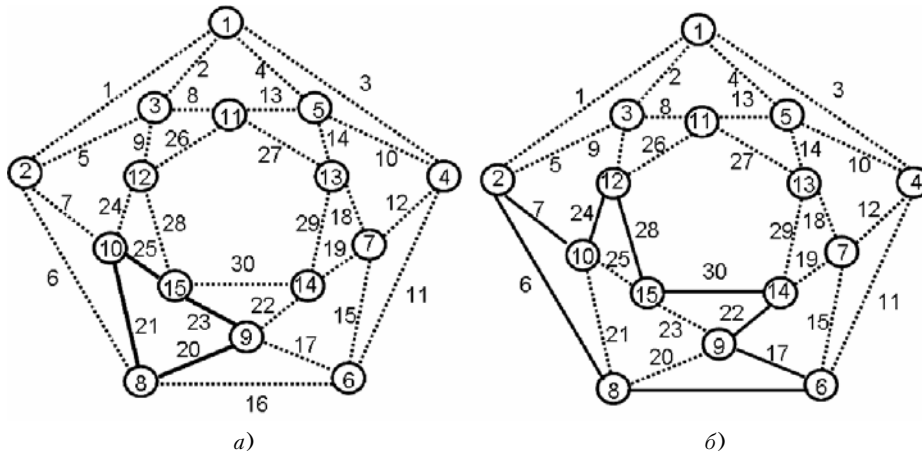


Рисунок 8 – Выделение четного соединяющего цикла в биквадратном графе B_3 (а), цикл C_v построенный на базе выделенного четного цикла (б)

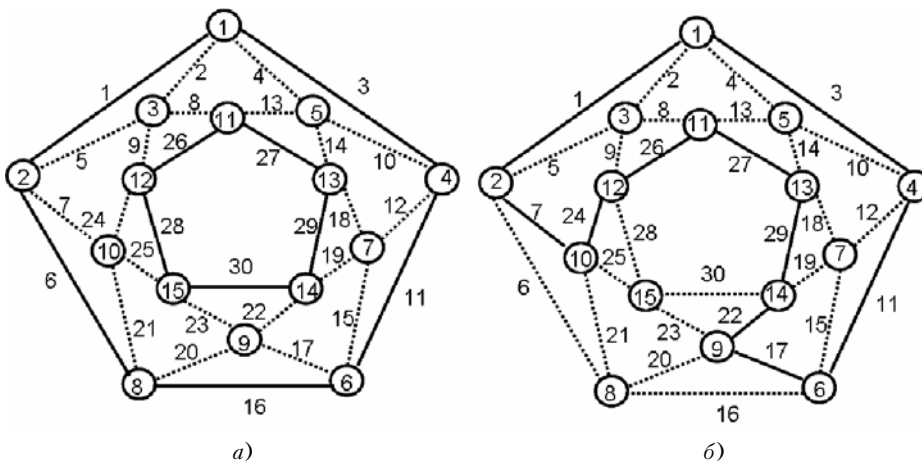


Рисунок 9 – Биквадратный граф B_3 и его 2-дуальные циклы с дисками нечетной длины (а) и 2-дуальные циклы с дисками четной длины (б)

Суммируем 2-дуальные циклы C_w , имеющие нечетную длину дисков и четный соединяющий цикл C_v . В результате получим новые 2-дуальные циклы с дисками четной длины:

$$\begin{aligned} C &= C_w \oplus C_v = \{1, 3, 6, 11, 16, 26, 27, 28, 29, 30\} \oplus \\ &\quad \oplus \{6, 7, 16, 17, 22, 24, 28, 30\} = \\ &= \{1, 3, 7, 11, 17, 22, 24, 26, 27, 29\}. \end{aligned}$$

Следует заметить, что операция перевода 2-дуальных циклов с дисками нечетной длины в 2-дуальные циклы с дисками четной длины возможно только тогда, когда выделенный четной соединяющей цикл C_v в биквадратном графе B_3 не есть W -маршрут. Введенная операция сложения 2-дуальных циклов с дисками нечетной длины и цикла C_v в биквадратном графе B_3 , переводящая в 2-дуальных циклы с дисками четной длины является общим свойством для плоских биквадратных графов B_3 .

Опишем алгоритм получения хроматического класса кубического графа H , равного трем.

Пусть задан кубический граф H и его рисунок.

Шаг 1. [Построение дуального биквадратного графа B_3]. Строим дуально-реберный биквадратный граф B_3 к кубическому графу H . Идем на шаг 2.

Шаг 2. [Построение 2-дуальных циклов]. Строим 2-дуальные циклы в биквадратном графе B_3 . Если выделенные 2-дуальные циклы содержат только диски четной длины, то конец работы алгоритма. В противном случае идем на шаг 3.

Шаг 3. [Выделение четного соединяющего цикла C_v]. Выделяем четные соединяющие циклы C_v . Если выделенные четные соединяющие циклы не W -маршруты, то суммируем их с 2-дуальными циклами и полученные новые 2-дуальные циклы имеют четные диски. Следовательно, биквадратный граф B_3 имеет хроматическое число равное трем, а соответствующий кубический граф имеет хроматический класс равный трем. Если выделенные четные соединяющие циклы W -маршруты, то получение новых 2-дуальных циклов имеющих четные диски не возможно. Следовательно, биквадратный граф B_3 не имеет хроматического класса равного трем. Конец работы алгоритма.

ВЫВОДЫ

1. Представлена новая операция, отражающая свойства класса плоских биквадратных графов B_3 , позволяющая производить перевод 2-фактора в соответствующем дуальном плоском кубическом графе H с дисками нечетной длины в 2-факторы с дисками чет-

ной длины. Применение данной операции позволяет конструктивно произвести раскраску ребер любого плоского кубического графа H графа в три цвета. Так как выделенные 2-дуальные циклы биквадратного графа B_3 содержат только диски четной длины, то тогда возможно раскрасить вершины, принадлежащие 2-дуальному циклу в два цвета, а в третий цвет раскрасить вершины дуальные 1-фактору кубического графа. Вершины биквадратного графа находятся на ребрах кубического графа, следовательно, хроматический класс кубического графа будет равен трем.

2. На основании введенной операции построен алгоритм с линейной вычислительной сложностью $O(m)$ для поиска хроматического класса плоского кубического графа H . Вычислительная сложность алгоритма для шага 2 равна $O(n)$. Вычислительная сложность алгоритма шага 3 также линейна, так как данный алгоритм основан на методе поиска в глубину, а его вычислительная сложность равна $O(m)$.

3. Данная работа носит концептуальный характер подхода к решению поставленной задачи, так как детальное описание пошаговых алгоритмов описано в монографии [8].

ПЕРЕЧЕНЬ ССЫЛОК

1. Донец Н. Г., Шор Н. З. Алгебраический подход к проблеме раскраски плоских графов – К.: Наук. думка, 1982. – 144 с.
2. Appel K., Haken W. Every Planar Map Is Four Colorable // Contemporary Mathematics. – Providence (R.I.): Amer. Math Soc., 1989. – Vol. 98. – P. 308.
3. Самохин А. В. Проблема четырех красок: Неоконченная теория доказательства // Соровский образовательский журнал. – 2001. – Т. 6, № 7. – С. 91–96.
4. Узерем Ч. Этюды для программистов: Пер. с англ. – М: Мир. – 1982. – 288 с.
5. Курапов С. В., Перепелица В. А. Идентификация рисунка графа на плоскости // Актуальные проблемы развития общественной мысли и практики управления: Сборник научных трудов. – Запорожский институт государственного и муниципального управления. – 1996. – Вып.2. – С. 225–232.
6. Харари Ф. Теория графов: Пер. с англ. Козырева В. П. / Под ред. Гаврилова В. Г. – М.: Мир, 1973. – 300 с.
7. Рингель Г. Теорема о раскраске карт. – М.: Мир. – 1977. – 126 с.
8. Курапов С. В., Савин В. В. Векторная алгебра и рисунок графа. – Запорожье: ЗГУ, 2003. – 200 с.

Надійшла 15.01.07
Після доробки 2.04.07

У даній роботі наведено конструктивний алгоритм для розфарбування кубічних графів із застосуванням теорії обертання вершин. Алгоритм заснований на фундаментальній теоремі Петерсена виділення 1-факторів і 2-факторів у кубічних графах.

The constructive algorithm for coloring the cubic graph with using the theory of the tops' rotation is presented. The algorithm is founded on the fundamental Petersen's theorem of separations of 1-factors and 2-factors in the cubic graphs.

А. И. Липчанский, А. А. Давыдов

КЛАССИФИКАЦИЯ ОСНОВНЫХ ПОКАЗАТЕЛЕЙ ДЛЯ ОЦЕНКИ НАДЕЖНОСТИ И ДИАГНОСТИРУЕМОСТИ КОМПЬЮТЕРНЫХ СИСТЕМ

В данной статье рассматривается классификация основных видов показателей для оценки надежности и диагностируемости компьютерных систем. Рассмотрены показатели безотказной работы, ремонтпригодности, долговечности, эксплуатационной готовности, комплексные показатели диагностируемости. Приведена классификация отказов.

ВВЕДЕНИЕ

За последние два десятилетия сложность и быстродействие компьютеров увеличилось в десятки тысяч раз. Задача, требовавшая год вычислений в 1983 г., сегодня решается за час, а современный карманный компьютер намного мощнее настольного тех времен. Но за прогресс и усовершенствования компьютерных систем (КС) приходится платить. Чем сложнее становятся КС, тем более неустойчиво и ненадежно они функционируют. Персональные компьютеры регулярно зависают, интернет-сайты довольно часто перестают работать. Новое программное обеспечение (ПО), разработанное для увеличения производительности КС, зачастую с точки зрения надежности только ухудшает ситуацию. В результате ежегодные затраты на поддержку и ремонт КС намного превышают общую стоимость оборудования и ПО как для индивидуальных пользователей, так и для корпораций. Традиционные меры по увеличению надежности КС рассчитаны на то, что операторы действуют безошибочно, и сейчас во многих случаях именно из-за их оплошности система выходит из строя на более длительное время, чем из-за любых других неполадок с оборудованием и ПО. Вопросы, связанные с оценкой надежности и диагностируемости КС, рассматривались во многих работах. Например, в [1] приведены различные аспекты гарантийности, принципы построения гарантийных КС, приводятся определения первичных свойств системы: безотказность, готовность, обслуживаемость, достоверность, функциональная безопасность, живучесть, целостность, конфиденциальность. В [2] выполнен аналитический обзор существующих экспертных систем (ЭС), встроенных в пассивные средства диагностирования вычислительных сетей, рассмотрены

преимущества и недостатки ЭС в целях определения целесообразности их использования для решения задач диагностирования корпоративных компьютерных сетей. В статье [3] рассмотрены применение экспертных систем реального времени для диагностики компьютерных систем.

ПОСТАНОВКА ЗАДАЧИ

Простой КС обходятся слишком дорого, а иногда просто недопустимы. Комплекс с надежностью в пять «девяток» работает без сбоев 99,999 % времени, т. е. за 25 лет в общем простаивает примерно 2 часа. Переход от двух «девяток» к пяти позволяет сохранить почти 90 часов рабочего времени в год. Поэтому классификация основных видов показателей для оценки надежности и диагностируемости КС на различных этапах жизненного цикла является актуальной задачей.

1 КЛАССИФИКАЦИЯ ОТКАЗОВ КС

Одной из важнейших характеристик надежности КС является безотказность – свойство объекта непрерывно сохранять работоспособное состояние в течение некоторого времени или наработки на отказ. Отказ КС рассматриваем как событие, заключающееся в прекращении способности объекта исполнять требуемую функцию. Ниже приведена классификация основных видов отказов КС:

- возникновение отказа: в аппаратной или программной части;
- изменения параметров КС во времени до отказа: внезапный (характеризующийся скачкообразным изменением значений параметров объекта) или постепенный (возникающий в результате постепенного изменения значений параметров);
- характер существования отказа во времени: сбой (самоустраняющийся или однократный отказ, устраняемый незначительным вмешательством оператора), перемежающийся (многократно возникающий самоустраняющийся отказ одного и того же характера), устой-

чивый (не прекращающийся до устранения его причины);

- возможность обнаружения: явный или скрытый;
- обусловленность другими отказами: зависимый или независимый;
- возможность восстановления работоспособности после отказа: неустранимый, устранимый на месте эксплуатации, устранимый на специализированном ремонтном предприятии;
- причина возникновения: конструктивный, производственный, эксплуатационный;
- тяжесть последствий: критический (при котором возникает угроза для жизни и здоровья людей, для окружающей среды, значительные экономические потери или невыполнение ответственного задания) или некритический.

2 ПОКАЗАТЕЛИ БЕЗОТКАЗНОЙ РАБОТЫ КС

Поскольку отказы компьютерных систем являются случайными событиями, то характеристики для оценки диагностируемости КС имеют вероятностный характер, а их численные значения определяются и анализируются статистическими и вероятностными методами. Рассмотрим основные показатели безотказной работы КС:

– средняя частота отказов, а (t_1, t_2) – отношение числа отказавших изделий в единицу времени к первоначальному числу испытываемых изделий при условии, что все вышедшие из строя изделия не восстанавливаются;

– наработка – продолжительность или объем работы объекта, измеренная в любых неубывающих величинах;

– средняя наработка до первого отказа (MTTF (Mean operating Time To Failure)), $T_{cp} = \int_0^{\infty} P(t)dt$ – математическое ожидание наработки объекта до первого отказа (имеет смысл только для восстанавливаемых систем);

– интенсивность отказов, $\lambda(t) = \frac{1}{P(t)} \frac{dP(t)}{dt}$ – условная плотность вероятности возникновения отказа объекта, определяемая при условии, что до рассматриваемого момента времени отказ не возник;

– средняя наработка на отказ (MTBF (Mean operating Time Between Failures)), $T_0 = t/M\{r(t)\}$ – отношение суммарной наработки восстанавливаемого объекта к математическому ожиданию числа его отказов в течение этой наработки;

– вероятность безотказной работы, $P(t_{0.p.}) = P\{\tau > t\}$ – вероятность того, что в пределах заданной наработки отказ объекта не возникнет.

3 ПОКАЗАТЕЛИ РЕМОНТОПРИГОДНОСТИ КС

Ремонтопригодность – свойство объекта, заключающееся в приспособленности к поддержанию и восстановлению работоспособного состояния путем технического обслуживания и ремонта при условии соблюдения условий эксплуатации объекта, а также условий проведения процедур технического обслуживания и ремонта. Рассмотрим основные показатели ремонтпригодности КС:

– вероятность восстановления, $P(t_b)$ – вероятность того, что время восстановления работоспособного состояния объекта не превысит заданное значение;

– среднее время ремонта (MRT (Mean Restoration Time)) – математическое ожидание времени восстановления работоспособного состояния объекта после отказа;

– среднее время восстановления (MTTR (Mean Time To Repair)), T_b – математическое ожидание времени восстановления работоспособного состояния объекта после отказа.

4 ПОКАЗАТЕЛИ ДОЛГОВЕЧНОСТИ КС

Долговечность – свойство объекта сохранять работоспособное состояние до наступления предельного состояния при установленной системе технического обслуживания и ремонта. Ниже представлены основные показатели долговечности КС:

– ресурс – суммарная наработка объекта от начала его эксплуатации или ее возобновления до перехода в предельное состояние;

– срок службы, T_{cc} – календарная продолжительность эксплуатации от начала эксплуатации объекта до перехода в предельное состояние.

5 ПОКАЗАТЕЛИ ЭКСПЛУАТАЦИОННОЙ ГОТОВНОСТИ

Эксплуатационная готовность – свойство объекта, отражающее как появление отказов, так и время неисправного состояния. Свойство эксплуатационной готовности является сочетанием свойств безотказности и ремонтпригодности объекта. Рассмотрим основные показатели эксплуатационной готовности:

– средний коэффициент эксплуатационной готовности, $\bar{A}(t_1, t_2)$ – усредненная на заданном интервале вероятность того, что объект окажется работоспособным в заданный момент времени, отсчитываемый от начала работы;

– средний коэффициент эксплуатационной неготовности, $\bar{U}(t_1, t_2)$ – усредненная на заданном интервале вероятность того, что объект окажется неработоспособным в заданный момент времени, отсчитываемый от начала работы;

– среднее время неисправного состояния (MDT) – математическое ожидание времени нахождения объекта в неисправном состоянии.

6 КОМПЛЕКСНЫЕ ПОКАЗАТЕЛИ ДИАГНОСТИРУЕМОСТИ

Рассмотрим следующие комплексные показатели диагностируемости:

– коэффициент готовности, $K_{Г} = T_0 / (T_0 + T_B)$ – вероятность того, что объект окажется в работоспособном состоянии в произвольный момент времени, кроме планируемых периодов, в течение которых применение объекта по назначению не предусматривается;

– коэффициент оперативной готовности, $K_{ОГ} = K \cdot P(t_{б.р.})$ – вероятность того, что система окажется в работоспособном состоянии в произвольный момент времени, кроме планируемых периодов, в течение которых применение системы по назначению не предусматривается, и начиная с этого момента будет работать безотказно в течение заданного времени;

– коэффициент технического использования, $K_{ТИ} = T_0 / (T_0 + T_B + T_{II})$ – отношение математического ожидания интервалов времени пребывания системы в работоспособном состоянии за некоторый период эксплу-

атации к сумме математических ожиданий интервалов времени пребывания системы в работоспособном состоянии, простоев, обусловленных техническим обслуживанием, и ремонтов за тот же период эксплуатации;

– коэффициент сохранения эффективности, $K_{эф}$ – отношение значения показателя эффективности за определенную продолжительность эксплуатации к номинальному значению этого показателя, вычисленному при условии, что отказы в системе в течение того же периода эксплуатации не возникают.

7 ПОКАЗАТЕЛИ НАДЕЖНОСТИ ПРОГРАММНОГО ОБЕСПЕЧЕНИЯ КС

Рассмотрим следующие показатели надежности программного обеспечения КС:

– количество ошибок в программе – число ошибок, обнаруженных в программе с момента начала ее эксплуатации до рассматриваемого момента;

– вероятность безотказной работы программы – вероятность того, что в пределах заданной наработки ошибка в программе не проявится;

– интенсивность потока ошибок программы – условная плотность вероятности возникновения ошибки в программе, определяемая при условии, что до рассматриваемого момента времени ошибка не возникла.

Таблица 1 – Выбор номенклатуры показателей для объектов вида I

Классификация объектов по признакам, определяющим выбор показателей				
По назначению	По режиму функционирования	По возможности восстановления и обслуживания		
		Восстанавливаемые		Невосстанавливаемые
		Обслуживаемые	Необслуживаемые	Обслуживаемые и необслуживаемые
Объекты конкретного назначения	Объекты непрерывного длительного применения	$K_{Г}$ или $K_{Т.и.}, T_0, T_B$	$K_{Г}, T_0, T_B$	$P(t_{б.р.})$ или $T_{ср}$
	Объекты многократного циклического применения	$K_{ОГ(t_{б.р.})} = K_{Г} \cdot P(t_{б.р.}), T_B$		P_0 и $T_{ср}$
Объекты общего назначения	Объекты непрерывного длительного применения и объекты многократного циклического применения	$K_{Т.и.}, T_0, T_B$	$K_{Г}, T_0, T_B$	$T_{ср}$
	Объекты однократного применения	–	–	P_0

Таблица 2 – Выбор номенклатуры показателей для объектов вида II

Классификация объектов по признакам, определяющим выбор показателей			
По назначению	По возможности восстановления и обслуживания		
	Восстанавливаемые		Невосстанавливаемые
	Обслуживаемые	Необслуживаемые	Обслуживаемые и необслуживаемые
Объекты конкретного назначения	$K_{эф}, T_B$		$K_{эф}$
Объекты общего назначения	$K_{Т.и.}, T_0$	$K_{Г}, T_0$	$T_{ср}$

8 ОБОБЩЕННАЯ СХЕМА ВЫБОРА ЗАДАВАЕМЫХ ПАРАМЕТРОВ ДИАГНОСТИРУЕМОСТИ

В табл. 1 приводится выбор номенклатуры показателей безотказности, ремонтпригодности или комплексных для объектов вида I.

В табл. 2 приводится выбор номенклатуры показателей безотказности, ремонтпригодности или комплексных для объектов вида II.

ВЫВОДЫ

Рассмотренные показатели надежности и диагностируемости КС целесообразно использовать на этапе автоматизированного проектирования диагностического обеспечения КС. Это позволит более эффективно осуществлять поиск возможных возникающих дефектов на этапе эксплуатации КС.

ПЕРЕЧЕНЬ ССЫЛОК

1. Харченко В. С. Гарантоспособность и гарантоспособные системы: элементы методологии // Научно-тех-

нический журнал «Радиоэлектронные и компьютерные системы». – 2006. – № 5 (17). – С. 7–20.

2. Кривуля Г. Ф., Липчанский А. И., Механна Сами, Зидат Хабис. Диагностика компьютерных сетей с использованием экспертных систем // Вестник ХГТУ. – 2004. – № 1 (19). – С. 11–16.
3. Кривуля Г. Ф., Бабич А. В., Липчанский А. И., Шкиль А. С. Применение экспертных систем реального времени для диагностики компьютерных систем // Наукотехнічний журнал «Інформаційно-керуючі системи на залізничному транспорті». – Харків. – 2004. – № 1. – С. 67–73.

Надійшла 2.03.07

У статті розглядається класифікація основних видів показників для оцінки надійності та здатності до діагностування комп'ютерних систем. Розглянуті показники безвідмовної роботи, ремонтоздатності, довговічності, експлуатаційної готовності, комплексні показники здатності до діагностування. Наведена класифікація відмов.

In given article is considered the classification of the basic kinds of parameters for a rating of reliability and diagnosability of computer systems. Such parameters were considered: of non-failure operation, of maintainability, of durability, of operational readiness, complex parameters diagnosability. Was given the classification of failures.

УДК 621.3

М. А. Новотарский

АЛГЕБРА ПРОЦЕССОВ ДЛЯ ИМИТАЦИОННОГО МОДЕЛИРОВАНИЯ СЛОЖНЫХ ДИСКРЕТНЫХ СИСТЕМ

В работе представлена алгебра процессов, как инструмент для формального описания сложных дискретных систем. Рассмотрена структура множеств, на которых определены переменные и выражения алгебры процессов. Описаны синтаксические конструкции и семантические правила выполнения операций, позволяющие задавать формальное описание имитационных моделей с реальной рабочей нагрузкой. Для определения адекватности модели предложено использовать принципы эквивалентности, базирующиеся на конгруэнтности, строгом и слабом взаимном подобии.

ВВЕДЕНИЕ

Ежедневно мы сталкиваемся с огромным количеством сложных дискретных систем, которые, большей частью, являются продуктом человеческой деятельности. Это и телекоммуникационные сети, и персональные компьютеры, и многое другое. Создавать или исследовать такие системы практически невозможно без пред-

варительной проверки принимаемых технических решений, поскольку компоненты систем, как правило, дорогостоящи и требуют существенных затрат на их производство. Одним из эффективных инструментов для проверки и предварительного анализа сложных систем является компьютерное моделирование. Построение моделей особенно актуально, когда речь идет о разработке мультипроцессорных систем, ориентированных на реализацию параллельных асинхронных алгоритмов. При этом не всегда удается эффективно использовать традиционные подходы к формализации исходной задачи, поскольку в данном случае модели должны отображать особенности параллелизма и асинхронного взаимодействия компонент. Например, ориентация на известные языки моделирования заведомо ставит исследователя в зависимость от способов функционирования и взаимодействия компонент, принятых в том или ином языке, а попытка эмулировать

в данном языке свои способы взаимодействия существенно снижает адекватность модели. В связи с этим значительно возрос интерес к поиску новых подходов к формализации, которые основываются на понятиях, способных более адекватно отображать различные аспекты функционирования параллельных компонент и их взаимодействие. Все они лежат в русле дальнейшего совершенствования уже известных технологий программирования компьютерных систем.

Первые программисты напрямую использовали машинные коды. Поэтому им приходилось иметь дело с сумматорами, регистрами, ячейками памяти и другими элементами, отображающими архитектуру системы. Существенным шагом к повышению сложности формализации стали языки программирования высокого уровня. Они позволили заменить понятия, связанные с архитектурой, на более близкие к тем, которыми пользуются при описании проблемы. Вместо байтов и регистров стало возможным использовать числа и строки. Языки моделирования добавили к традиционному перечню еще и специализированные структуры данных, облегчающие описание сложных компонент и их взаимодействие. Очевидно, что упомянутые сущности подняли формализацию на более высокий уровень. Однако существенный рост сложности объектов моделирования и нетривиальный характер их взаимодействия привели к ситуации, когда при использовании языков высокого уровня стало весьма затруднительно строить корректные модели. Возникла объективная необходимость введения новых понятий – процессов, событий, активностей – позволяющих представлять модель на более высоком уровне абстракции. Теории, использующие эти понятия в качестве базовых, получили название алгебр процессов. Главной отличительной чертой алгебры процессов является возможность представления в виде математически строгих выражений систем, состоящих из параллельно развивающихся и взаимодействующих процессов. Основы алгебры процессов, называемой «исчисление коммуникационных систем», впервые изложены в [1] и послужили толчком к бурному развитию этого направления. В результате были созданы алгебры процессов, позволяющие исследовать различные аспекты поведения сложных систем. Общим для этих алгебр является то, что все они созданы для описания сложных дискретных систем, базируются на ряде аксиом, задающих синтаксис выражений и семантические правила, и содержащих признаки поведенческой эквивалентности, позволяющие сравнивать системы, поведение которых неразличимо для внешнего наблюдателя.

Эта работа посвящена проблемам дальнейшего развития теории алгебр процессов, ориентированных на описание параллельных процессов, характеризующихся широким спектром взаимодействий, включая асинхронный характер взаимодействия. Еще одна суще-

ственная черта предлагаемой алгебры процессов состоит в ее ориентации на создание имитационных моделей с реальной рабочей нагрузкой. Такой подход позволяет значительно повысить поведенческую эквивалентность модели объекту моделирования. Простые и прозрачные способы представления реальной нагрузки осуществлены путем расширения спектра синтаксических конструкций, а новые семантические правила отображают не только традиционные для алгебр процессов коммуникационные аспекты функционирования модели, но и допускают использование прикладных алгоритмов обработки входных данных. Адекватность моделей, построенных с использованием данной алгебры процессов, строго очерчена путем доказательства конгруэнтности выражений, а также обоснования правил строгого и слабого взаимного подобия.

1 БАЗОВЫЕ ПОНЯТИЯ

Алгебра процессов для моделирования с дискретными событиями использует некоторые свойства классических базовых алгебр процессов [1], а также новые тенденции развития, возникшие в результате применения алгебр процессов к решению проблем имитационного моделирования с дискретными событиями [2, 3]. Парадигмы алгебры процессов предполагают наличие в системе активных компонент и взаимодействие между ними. В [1] активные компоненты называются процессами, которые совершают активности, представляющие дискретные действия системы. Любая активность в данной алгебре процессов может принадлежать множеству взаимодействия Λ или множеству внутренних активностей Ψ . Множество взаимодействия Λ содержит подмножество активностей ввода Δ и подмножество активностей вывода $\bar{\Delta}: \Lambda = \Delta \cup \bar{\Delta}$.

Собственно акт взаимодействия имеет место только при наличии источника и приемника информации, что обуславливает его бинарный характер. Поэтому каждая активность $a(x) \in \Delta$ всегда имеет сопряженную активность $\bar{a}(x) \in \bar{\Delta}$, и это именно та сопряженная активность, с которой готова синхронизироваться активность $\bar{a}(x)$. Скобки (x) содержат значения переменной, подлежащей вводу или выводу, и могут упускаться в случаях, если важным есть только факт взаимодействия.

Множество внутренних активностей Ψ состоит из подмножеств обработки Φ и подмножества задержек $T: \Psi = \Phi \cup T$.

Подмножество обработки включает активности обработки $\phi \in \Phi$. Каждая активность такого типа представляет определенные действия процесса по преобразованию информации. Исходя из этого определения, очевидно, что активность обработки может также представлять процессы, и тем самым обеспечивает невяную

агрегативность структур описания процессов. В алгебре процессов можно исключить внутреннее поведение процесса и рассматривать его как «черный ящик». В этом случае возникает возможность построить сложную имитационную модель, исключив из описания более простые части, поведение которых не представляет интереса. Такое невидимое поведение моделируется с помощью активностей $\tau \in T$, заменяющих внутренние действия задержкой, выраженной в единицах модельного времени. Активности типа τ также используют в случае явного выделения модельного времени для выполнения активностей обработки.

Таким образом, полный алфавит активностей $Act = \Lambda \cup \Psi$ объединяет множество активностей взаимодействия Λ и множество внутренних активностей Ψ .

Процесс в алгебре процессов задается выражением, содержащим активности как независимые или зависимые переменные. Простейший константный процесс, освобождающий все ресурсы, задается активностью nil , где $nil \in \Phi$. Если процесс содержит более одной активности, то последовательность их свершения регулируется префиксным представлением $\varphi.P$. Смысл префиксного представления состоит в том, что сначала свершается активность φ , а затем процесс ведет себя как процесс P . Примером элементарного процесса P_1 служит выражение: $P_1 := begin.end.nil$, где $\{begin, end, nil\} \subset \Phi$.

Символ $:=$ задает присвоение значения выражения в правой части процессу P_1 . Действия процесса, согласно данному выражению, состоят из прямой последовательности трех активностей. Активность $begin$ выполняет действия, связанные с началом функционирования процесса P_1 , end выполняет действия, связанные с завершением процесса P_1 , nil освобождает ресурсы и удаляет процесс P_1 из модели.

2 СИНТАКСИС

Синтаксис алгебры процессов для имитационного моделирования систем с реальной рабочей нагрузкой базируется на использовании типов переменных для параметров Var и типов переменных для процессов Pvar .

Множество переменных типа Var обеспечивает временное хранение параметров и включает простые переменные x, y, z, \dots ; индексированные переменные $x\{i\}$, $(y\{i\}, z\{i\}, \dots)$, где i – индекс; ссылочные переменные $\hat{x}, \hat{y}, \hat{z}, \dots$. Множество переменных типа Pvar временно хранит экземпляры процессов и включает простые переменные X, Y, Z, \dots ; индексированные переменные $X\{i\}, Y\{i\}, Z\{i\}, \dots$, где i – индекс; ссылочные переменные $\hat{X}, \hat{Y}, \hat{Z}, \dots$.

Выражения для представления функциональных зависимостей и времени составляют множество выраже-

ний Expr , заданное на множестве элементарных математических операций.

Выражения для процессов объединены во множество Pexpr , включающее следующие операции:

- префиксация – $\varphi.P, \langle \tau \rangle.P, \langle \tau \leftarrow t \rangle.P, a(x).P, \bar{a}(x).P$;
- выбор – $\sum_{i \in I} P_i, \sum_{i \in I} v_i!P_i, \sum_{i \in I} w_i\#P_i$;
- параллельная композиция и взаимодействие – $P|Q|\dots|C, P \leftrightarrow Q|\dots|C, P \triangleright Q|\dots|C, P \triangleleft Q|\dots|C$;
- переименование и рестрикция – $P[K], P \setminus A$;
- клонирование, вставка и сокрытие – $k \times P, C\{P\}, \varphi_P(C)$;
- рекурсия и цикл – $\mathfrak{R}(X = P), cycle[x](X = P)$.

Рассмотрим кратко смысл упомянутых операций. Префиксная задержка $\langle \tau \rangle.P$ задает процесс с начальной задержкой на τ единиц модельного времени, а параметрическая префиксная задержка $\langle \tau \leftarrow t \rangle.P$ описывает действия процесса, для которого сначала определяется значение параметра τ путем присвоения ему результата вычисления выражения t . После задержки τ в обоих случаях процессы ведут себя как процесс P .

Префиксный ввод $a(x).P$ задает действия процесса по вводу некоторого значения k с последующим присвоением его переменной x . Префиксный вывод $\bar{a}(x).P$ задает действия процесса по выводу того значения переменной x , которое она приобрела на момент свершения активности вывода $\bar{a}(x)$.

Операции выбора предназначены для выбора одного процесса из множества процессов с целью его дальнейшего развития. Недетерминированный выбор $\sum_{i \in I} P_i$ обеспечивает определение одного из процессов $P_1 + \dots + P_i + \dots + P_I$ в момент времени, который называется моментом выбора. Управляемый выбор $\sum_{i \in I} v_i!P_i$ реализует выбор процесса P_i , снабженного логической переменной v_i , принявшей истинное значение. Вероятностный выбор $\sum_{i \in I} w_i\#P_i$ обеспечивает выбор процесса P_i с вероятностью w_i при условии $w_1 + \dots + w_i + \dots + w_I = 1$.

Операции параллельной композиции и взаимодействия определяют взаимное поведение процессов, которые развиваются параллельно. Операцию $P|Q|\dots|C$ применяют в случаях, если нет необходимости акцентировать внимание на характере взаимодействия между параллельными процессами или если такого взаимодействия не существует. Операция может применяться для определения параллелизма группы процессов без детализации характера связей между ними. Операция взаимодействия определяет поведение совокупного

процесса, который состоит из двух процессов, развивающихся параллельно.

Синхронным взаимодействием $P \leftrightarrow_L Q|\dots|C$ процесса P с параллельной композицией процессов $Q|\dots|C$ будем называть систему, реализующую передачу информации от передатчика P к приемникам $Q|\dots|C$ в случае одновременного свершения комплементарных событий из множества $L \subseteq \Lambda$.

В отличие от операции выбора, при взаимодействии рассматривают процессы, активности которых пользуются собственными скрытыми ресурсами. Это разрешает им развиваться независимо и использовать свои ресурсы в случае, если события, которые их вызывают, не входят во множество событий взаимодействия. При возникновении события, которое входит во множество событий взаимодействия, происходит синхронизация развития параллельных процессов с целью выполнения общей активности.

Управляемой передачей $P \triangleright_L Q|\dots|C$ от процесса P к параллельной композиции $Q|\dots|C$ называют систему, состоящую из процесса-передатчика P и множества процессов-приемников $Q|\dots|C$, которые параллельно развиваются и взаимодействуют в момент свершения в процессе P активности из множества взаимодействия $L \in \Lambda$. Главное отличие этой операции от операции синхронного взаимодействия состоит в том, что передача данных происходит непосредственно в момент свершения активности вывода $\bar{a}(x)$ в процессе P , а композиция процессов $Q|\dots|C$ всегда готова к приему информации.

Асинхронным взаимодействием $P \triangleleft_L Q|\dots|C$ процесса P и параллельной композиции $Q|\dots|C$ будем называть систему, которая состоит из процесса P и множества параллельно развивающихся процессов $Q|\dots|C$, способных передавать или принимать информацию от процесса P в соответствии с множеством событий взаимодействия $L \in \Lambda$ независимо друг от друга. При асинхронном взаимодействии каждый из процессов, входящих в систему, может инициировать активность ввода или вывода независимо от готовности сопряженной активности. Поэтому задержки для синхронизации при такой организации взаимодействия процессов не возникает.

Операции переименования и сжатия относят к операциям группового редактирования, поскольку они обеспечивают модификацию группы активностей процесса одновременно. Операция переименования $P[K]$ производит замену всех активностей процесса P с помощью функции K , а операция сжатия $P \setminus A$ блокирует в процессе P все активности, входящие во множество A .

Операции клонирования, вставки и сокрытия предназначены для динамической модификации множества

компонент системы в ходе моделирования. Операция клонирования $k \times P$ представлена системой, которая состоит из k идентичных процессов P , развивающихся параллельно. Операция вставки $C\{P\}$ задает процесс C , включающий процесс P как subprocess. Операция сокрытия $\phi_P(C)$ описывает систему, состоящую из процесса C , в котором все активности, входящие в subprocess P , заменяются внутренней активностью ϕ .

Операции рекурсии и цикла предназначены для многократного повторения некоторой последовательности активностей, входящих в процесс. Операция рекурсии $\mathfrak{R}(X = P)$ образует систему, состоящую из рекурсивно повторяемой последовательности активностей процесса P , присвоенных переменной типа Pvar . Операция цикла $\text{cycle}[x](X = P)$, подобно рекурсии, состоит из циклически повторяемой последовательности процесса P , но количество таких повторений ограничено значением переменной x .

3 СЕМАНТИКА

Операционная семантика процессов моделирования, представленных данной алгеброй, сводится к определению переходов, представляющих изменения состояний имитационной модели. Определим динамику выполнения операций, используя нотацию Плоткина [4], представляющую операцию в виде дроби, знаменатель которой содержит общую запись операции, а числитель – возможные варианты ее выполнения.

$$\begin{aligned} & \text{Операции префиксации } \overline{\phi.P \xrightarrow{\phi} P}, \overline{\langle \tau \rangle.P \xrightarrow{\langle \tau \rangle} P}, \\ & \overline{\langle \tau \leftarrow t \rangle.P \xrightarrow{\langle t \rangle} P[t/\tau]}, \quad \overline{a(x).P \xrightarrow{a(k)} P[k/x]} \\ & \text{и } \overline{\bar{a}(x).P \xrightarrow{\bar{a}(m)} P} \text{ не имеют числителя, поскольку} \end{aligned}$$

существует единственный вариант их выполнения, заключающийся в продолжении текущего процесса после выполнения активности. Отличия между данными операциями обусловлены различными состояниями переменных процесса и способами их замещения.

Количество вариантов выполнения недетерминированной операции выбора равно количеству процессов, принимающих участие в данной операции:

$$\begin{aligned} & \frac{P_1 \xrightarrow{\alpha} P'_1, \dots, P_k \xrightarrow{\beta} P'_k, \dots,}{\sum_{i \in I} P_i \xrightarrow{\alpha} P'_i} \quad \frac{P_1 \xrightarrow{\beta} P'_1, \dots, P_k \xrightarrow{\alpha} P'_k, \dots,}{\sum_{i \in I} P_i \xrightarrow{\beta} P'_i} \\ & \frac{P_l \xrightarrow{\gamma} P'_l}{\sum_{i \in I} P_i \xrightarrow{\gamma} P'_i} \end{aligned}$$

Операция управляемого выбора в нотации алгебры процессов в знаменателе содержит общую запись возможного выбора. Числитель указывает на то, что осуществляется положительный выбор того процесса, для

которого соответствующая логическая переменная v_i принимает истинное значение:

$$\frac{\{P_i | v_i = true\} \xrightarrow{\alpha} P'}{\sum_{i \in I} v_i! P_i \xrightarrow{\alpha} P'}$$

Вероятностный выбор структурно напоминает операцию управляемого выбора, за исключением того, что в данном случае задаются вероятности для всех процессов, принимающих участие в выборе. Акт выбора всегда осуществляется как результат выполнения вероятности:

$$\frac{\{P_i | w_i\} \xrightarrow{\alpha} P'}{\sum_{i \in I} w_i \# P_i \xrightarrow{\alpha} P'}$$

Нотация параллельной композиции задает общую формулу взаимодействия между двумя процессами, входящими в группу параллельных процессов:

$$\frac{P \xrightarrow{a(x)} P' \quad Q \xrightarrow{\bar{a}(x)} Q'}{P | Q \xrightarrow{\tau} P' | Q'}$$

Конкретизация допустимых видов взаимодействия реализована с помощью операций взаимодействия. Динамика взаимодействия в алгебре процессов включает операции управляемой передачи, синхронного взаимодействия и асинхронного взаимодействия. Отсутствие взаимодействия рассматривается как параллельное выполнение процессов. Нотации этих операций имеют вид:

$$\frac{P \xrightarrow{\bar{a}(x)} P' \quad Q \xrightarrow{a(x)} Q' \dots C \xrightarrow{a(x)} C'}{P \triangleright_L Q | \dots | C \xrightarrow{\tau} P' \triangleright_L Q' | \dots | C'} \quad \text{уп-}$$

равляемая передача,

$$\frac{P \xrightarrow{\bar{a}(x)} P' \quad Q \xrightarrow{a(x)} Q' \dots C \xrightarrow{a(x)} C'}{P \leftrightarrow_L Q | \dots | C \xrightarrow{\tau} P' \leftrightarrow_L Q' | \dots | C'} \quad \text{син-}$$

хронное взаимодействие,

$$\frac{P \xrightarrow{\bar{a}(x)} P' \quad Q \xrightarrow{a(x)} Q' \dots C \xrightarrow{a(x)} C'}{P \triangleleft_L Q | \dots | C \xrightarrow{\tau} P' \triangleleft_L Q' | \dots | C'} \quad \text{асин-}$$

хронное взаимодействие.

При управляемой передаче инициатором действия является префиксный процесс $\bar{a}(x).P$. Процессы параллельной композиции $Q | \dots | C$ постоянно готовы к приему информации благодаря префиксации $a(x).Q | \dots | C$. Передача происходит путем одновременного свершения активности $\bar{a}(x)$ в процессе-передатчике и активностей $a(x)$ в процессах-приемниках. Иначе композиция процессов $Q | \dots | C$ будет находиться в состоянии ожидания приема данных.

В случае синхронного взаимодействия оба процесса могут инициировать взаимодействие. Первый из процессов, прибывших в точку взаимодействия, отправляет

процессу-партнеру активность и останавливается до момента свершения сопряженной активности.

Операция асинхронного взаимодействия предполагает наличие внутренних активностей, обеспечивающих возможность продолжения развития процесса-инициатора взаимодействия и выдачу информации процессу-приемнику в момент активации в нем активности $a(x)$.

Операции переименования и сжатия представлены в алгебре процессов нотациями:

$$\frac{\frac{P \xrightarrow{\alpha} P'}{P[K] \xrightarrow{K(\alpha)} P'[K]}}{P \setminus A \xrightarrow{\alpha} P' \setminus A} \quad (\{\alpha, \bar{\alpha}\} \not\subset A).$$

Эти операции, в отличие от операций выбора и взаимодействия, работают с одним процессом, но предназначены для организации групповых модификаций его активностей. Источником модификаций для операции переименования есть функция K , а для операции сжатия – множество активностей A .

Динамика операции клонирования процессов представлена выражением:

$$\frac{P \xrightarrow{\alpha} P'_1 \quad P \xrightarrow{\alpha} P'_2 \dots P \xrightarrow{\alpha} P'_k}{P \xrightarrow{\alpha} P'_1 | P'_2 | \dots | P'_k}$$

Суть операции $k \times P$ сводится к генерации группы из k процессов на основе шаблона P .

Операции рекурсии и цикла представлены нотациями, которые отображают динамику их выполнения:

$$\frac{P \{ \mathfrak{R}(X = P) / X \} \xrightarrow{\alpha} P'}{\mathfrak{R}(X = P) \xrightarrow{\alpha} P'} \quad \text{рекурсия,}$$

$$\frac{P \{ cycle[x-1](X = P) / X \} \xrightarrow{\alpha} P'}{cycle[x](X = P) \xrightarrow{\alpha} P'} \quad (x \geq 0) \quad \text{цикл.}$$

Операции состоят из трех основных этапов: инициализации элемента цикла путем присвоения значений процесса переменной X , выполнения созданного экземпляра процесса и проверки условий завершения. Проверка условий завершения рекурсивной операции всегда зависит от внутреннего состояния процесса P , а операция цикла использует значение счетчика x для завершения своей работы.

Операция вставки $C \{ P \}$ обеспечивает возможность иерархического построения описания системы путем объединения процессов, входящих в разные уровни стратификации модели. Динамика выполнения данной операции соответствует нотации:

$$\frac{C \xrightarrow{a(x)} C'}{C \{ X \} \xrightarrow{a(P)} C' \{ P \}}, \quad \text{где } X \in \text{PVar}.$$

Выполнение активности, инициирующей операцию вставки, сводится к присвоению переменной процесса X значения некоторого процесса P , ссылку на который содержит упомянутая активность. В ходе развития процесс C использует процесс P как субпроцесс.

Операция сокрытия тоже используется при построении и исследовании стратифицированных моделей. С помощью этой операции становится возможным скрывать детализацию функционирования некоторых компонент системы с целью упрощения процесса моделирования. Операция сокрытия представлена нотацией:

$$\frac{C \xrightarrow{a} C'}{\varphi_P(C) \xrightarrow{\varphi} \varphi_P(C')}, \text{ где } a \in P.$$

Выполнение этой операции сводится к замене некоторой последовательности активностей, составляющих процесс P , одной внутренней активностью φ .

4 ЭКВИВАЛЕНТНОЕ ПОВЕДЕНИЕ

Синтаксис и семантика алгебры процессов задают способ представления объектов модели. Но для того, чтобы модель была адекватной, необходимо придерживаться соответствующих принципов эквивалентности поведения, базирующихся на эквивалентном отношении. Пусть P – процесс, тогда эквивалентным отношением R называется множество $R \subseteq P \times P$ с элементами $(p, q) \in R$ или pRq , удовлетворяющее условиям рефлексивности, симметричности и транзитивности. Рефлексивность обуславливает эквивалентность тождественного отношения, то есть pRq при $p \in P$; симметричность выражает тот факт, что из pRq следует qRp при $p, q \in P$; а транзитивностью называют свойство, для которого из pRq и qRz вытекает, что pRz при $p, q, z \in P$.

Для определения эквивалентности поведения применяют также конгруэнтность, то есть сохранение эквивалентного отношения на разных уровнях стратифицированных моделей. Пусть P и Q – процессы, связанные эквивалентным отношением PRQ . Такие процессы будем называть конгруэнтными, если эквивалентное отношение сохраняется при $C\{P\}RC\{Q\}$, где $C\{\}$ – процесс высшего уровня, использующий в качестве субпроцесса процесс P или Q . Благодаря конгруэнтности эквивалентность поведения может определяться как видимое поведение процессов, а не через их структуру, выраженную количеством допустимых состояний и переходов между ними. При таком подходе эквивалентные процессы распознаются наблюдателем как одинаково взаимодействующие с окружающей средой, а их взаимодействия приводят к одинаковому результату.

Строгое обоснование эквивалентности поведения базируется на использовании понятия взаимного подобия. Преимущество применения такого подхода состоит в возможности ссылки на определение подобия, исходя из общей теории систем и построения доказательств на основе этого определения. В зависимости от того, принимаются ли во внимание скрытые активности, взаимное подобие разделяют на строгое и слабое.

Поскольку алгебра процессов может отображать бесконечное количество систем, отвечающих ее синтаксическим конструкциям, то для конкретного определения взаимного подобия рассмотрим маркированную систему с переходами, заданную кортежем $(\Pi, A, \{ \xrightarrow{a} \mid a \in A \})$, где Π – множество процессов; A – множество активностей; \xrightarrow{a} – множество переходов: $\xrightarrow{a} \subseteq \Pi \times A \times \Pi$. Пусть в данной маркированной системе с переходами существует отношение $R \subseteq \Pi \times \Pi$ и процессы $P, Q \subseteq \Pi$ с допустимыми состояниями $p, p' \in P, q, q' \in Q$ при $(p, q) \in R$. Тогда если для произвольной активности a для каждого перехода $p \xrightarrow{a} p'$ существует состояние q' , такое, что $q \xrightarrow{a} q'$ при $(p', q') \in R$, и для каждого перехода $q \xrightarrow{a} q'$ существует состояние p' , такое, что $p \xrightarrow{a} p'$ при $(p', q') \in R$, то отношение R будем называть строгим взаимным подобием, если $a \in Act$. Состояния p и q строго взаимно подобны $p \sim q$, если существует отношение R , содержащее (p, q) при условии $a \in Act$.

Рассмотренное строгое взаимное подобие – это основной вид эквивалентности, позволяющий разрабатывать сценарии сравнения сложных дискретных систем. Оно объединяет все допустимые отношения взаимного подобия и потому является наибольшим взаимным подобием. Все операторы рассматриваемой алгебры процессов имеют те или иные свойства строгого взаимного подобия. Но очевидно, что применение данного вида эквивалентности не является исчерпывающим для исследования сложных систем. Отношение строгого взаимного подобия строится с учетом всех без исключения активностей, входящих в алфавит системы. Поэтому затруднительно выполнять обобщающие оценки, учитывающие только некоторое заданное подмножество свойств. Известный механизм обобщения базируется на применении методик вертикальной и горизонтальной стратификации. В рамках рассмотренных подходов упомянутые методики используют операции вставки и сокрытия для стратификации на уровне процессов. С целью абстрагирования от внутреннего поведения системы на уровне отдельных активностей, например, в классической алгебре процессов CCS [1], используют специальный тип активности – внутренняя активность. Однако такой подход нуждается в усовершенствовании принципов эквивалентности с целью

распространения их на процессы, в которых существует скрытое внутреннее поведение. Проблема состоит в том, что простое исключение внутренних активностей некорректно, поскольку они могут инициировать видимые активности. При этом создается тупиковая ситуация, нарушающая информационную целостность системы. Для решения этой проблемы вводят новый тип перехода $\xRightarrow{a} \subseteq \Pi \times \Pi$ для произвольных активностей $a \in A$ в соответствии с правилом:

$$\xRightarrow{a} := \begin{cases} \xrightarrow{\varphi} \circ \xrightarrow{a} \circ \xrightarrow{\varphi} & \text{при } a \neq \varphi, \\ \xrightarrow{\varphi} & \text{при } a = \varphi, \end{cases}$$

где φ – внутренняя активность, \circ – символ, указывающий на допустимость вставки в этом месте композиции бинарных отношений.

Пусть в маркированной системе с переходами $(\Pi, A, \{\xRightarrow{a} | a \in A\})$ существует отношение $R \subseteq \Pi \times \Pi$ и процессы $P, Q \subseteq \Pi$ с допустимыми состояниями $p, p' \in P, q, q' \in Q$ при $(p, q) \in R$. Тогда если для произвольной активности a для каждого перехода $p \xrightarrow{a} p'$ существует состояние q' такое, что $q \xRightarrow{a} q'$ и $(p', q') \in R$, и для каждого перехода $q \xrightarrow{a} q'$ существует состояние p' такое, что $p \xRightarrow{a} p'$ и $(p', q') \in R$, то отношение R будем называть слабым взаимным подобием или наблюдаемой эквивалентностью. Состояния p и q слабо взаимно подобны $p \approx q$, если существует наблюдаемая эквивалентность R , содержащая (p, q) .

Строгое и слабое взаимные подобия состояний лежат в основе понимания взаимного подобия процессов. Два процесса P и Q называют строго взаимно подобными $P \sim Q$ или слабо взаимно подобными $P \approx Q$ тогда и только тогда, когда для произвольной активности a существует отношение соответственно строгого или слабого взаимного подобия R такое, что PRQ .

Пусть P, Q, G – процессы. Если $P \sim Q$, то:

- $a.P \sim a.Q$ для произвольной активности a ;
- $\langle \tau \rangle.P \sim \langle \tau \rangle.Q$ для произвольной префиксной задержки $\langle \tau \rangle$;
- $P + G \sim Q + G$ для произвольного процесса G ;
- $P|G \sim Q|G$ для произвольного процесса G ;
- $P[K] \sim Q[K]$ для произвольной функции переименования K ;
- $P \setminus A \sim Q \setminus A$ для произвольного множества активностей A ;
- $k \times P \sim k \times Q$ для произвольного коэффициента клонирования k ;
- $\mathfrak{R}(X = P) \sim \mathfrak{R}(X = Q)$, где X переменная процесса;

– $\text{cycle}[x](X = P) \sim \text{cycle}[x](X = Q)$, где x – индекс цикла.

Для P, Q и активности a справедливы такие наблюдаемые эквивалентности при $P \approx Q$:

- $\langle \tau \rangle.P \approx \langle \tau \rangle.Q$;
- $a.P \approx a.Q$;
- $P + \varphi.P \approx \varphi.P$;
- $a.(P + \varphi.Q) \approx a.(P + \varphi.Q) + a.Q$.

ВЫВОДЫ

Большой интерес к использованию алгебр процессов для формального описания сложных систем вызван, прежде всего, возможностью представления правил функционирования и взаимодействия компонент системы в виде набора математических выражений. Такой подход позволяет оценить адекватность модели на различных этапах ее жизненного цикла. Большинство известных алгебр процессов [1], [3] ориентировано на описание и исследование коммуникаций между структурными элементами систем. Однако для представления моделей с реальной рабочей нагрузкой необходимо иметь возможность задавать алгоритмы функционирования компонент, обрабатывающие внешние потоки данных. С этой целью в предлагаемой алгебре процессов введено множество внутренних активностей Ψ , каждая активность которого представляет определенные действия процесса по преобразованию информации. Расширение операций выбора и взаимодействия, а также введение операций клонирования и сокрытия продиктовано стремлением расширить сферу использования данной алгебры и гибкость представления с ее помощью сложных систем. Строгое обоснование адекватности базируется на использовании понятия эквивалентного отношения в модели, заданной маркированной системой с переходами. Свойство конгруэнтности введено с целью обеспечения возможности вертикальной и горизонтальной стратификации модели. Определены строгое и слабое взаимные подобия, а также правила их сохранения при выполнении базовых операций.

ПЕРЕЧЕНЬ ССЫЛОК

1. Milner R. Calculus of Communicating Systems // Lecture Notes in Computer Sciences. – 1980. – Vol. 92. – P. 260.
2. Milner R. Communication and Concurrency. – Prentice Hall. – 1995. – P. 272.
3. Bergstra J. A., Middelburg C. A. Process algebra for hybrid systems // Theoretical Computer Science. – 2005. – Vol. 335, No. 2/3. – PP. 215–208.
4. Plotkin G. D. A Structured Approach to Operational Semantics. Technical Report DAIMI FM-19, Computer Science Department, Aarhus University. – 1981. – P. 133.

Надійшла 31.01.07

In the paper the process algebra as the tool for the formal description of complex discrete systems is submitted. The structure of sets on which variables and expressions of process algebra are determined is considered. Syntactic structures and semantic rules of performance of the operations allowing to set the formal description for simulations with real working load are described. For definition of adequacy of model it is offered to use the principles of equivalence basing on congruence, strong and weak bisimilarity.

У роботі представлена алгебра процесів як інструмент для формального опису складних дискретних систем. Розглянуто структуру множин, на яких визначені змінні і вирази алгебри процесів. Описано синтаксичні конструкції і семантичні правила виконання операцій, що дозволяють задавати формальний опис імітаційних моделей з реальним робочим навантаженням. Для визначення адекватності моделі запропоновано використовувати принципи еквівалентності, що базуються на конгруентності, строгій і слабкій взаємній подібності.

УДК 62-52.087

А. Л. Становский, Т. В. Лысенко, Т. И. Носенко, Е. В. Лысенко

СИНХРОНИЗИРУЮЩЕЕ УПРАВЛЕНИЕ ЭКСТРЕМУМАМИ ОТДАЛЕННЫХ СОСТОЯНИЙ

Определено понятие «событие» в теории синхронизирующего управления как поименованное состояние. Выполнена классификация событий. Предложен метод нейросеточной экстраполяции дискретных числовых временных трендов, основанный на замене значений тренда в узловых точках на значения производной интерполяционного полинома в этих же точках.

ВВЕДЕНИЕ

Теория синхронизирующего управления предполагает поиск и реализацию такого внешнего воздействия на объект, которое обеспечивает в будущем (иногда, весьма отдаленном) совпадение некоторых событий, ожидаемых в его подсистемах [1].

ПОСТАНОВКА ЗАДАЧИ

Понятие «событие» определено в теории вероятностей как «всякий результат или исход испытания» [2] и является аксиоматической основой для построения булевой алгебры событий.

Пусть некоторая система Ω определена в n_{Ω} -мерном пространстве состояний $\Omega(\tau)$, где τ – время. Разобьем $\Omega(\tau)$ на k множеств:

$$Y_1(\tau) = \{y_{11}(\tau), y_{12}(\tau), \dots, y_{1n_1}(\tau)\},$$

$$Y_2(\tau) = \{y_{21}(\tau), y_{22}(\tau), \dots, y_{2n_2}(\tau)\},$$

$$\dots$$

$$Y_k(\tau) = \{y_{k1}(\tau), y_{k2}(\tau), \dots, y_{kn_k}(\tau)\}$$

с размерностями n_1, n_2, \dots, n_k , соответственно, которые могут частично пересекаться между собой. Назовем части общей системы Ω , определенные на $Y_1(\tau), Y_2(\tau),$

$\dots, Y_k(\tau)$, соответственно подсистемами Y_1, Y_2, \dots, Y_k .

Выделим в подсистемах Y_1, Y_2, \dots, Y_k поименованные состояния:

$$Y_1^{S_{11}}(\tau) = \{y_{11}^{S_{11}}(\tau), y_{12}^{S_{11}}(\tau), \dots, y_{1n_1}^{S_{11}}(\tau)\},$$

$$Y_1^{S_{12}}(\tau) = \{y_{11}^{S_{12}}(\tau), y_{12}^{S_{12}}(\tau), \dots, y_{1n_1}^{S_{12}}(\tau)\},$$

...

$$Y_1^{S_{1q_1}}(\tau) = \{y_{11}^{S_{1q_1}}(\tau), y_{12}^{S_{1q_1}}(\tau), \dots, y_{1n_1}^{S_{1q_1}}(\tau)\},$$

$$Y_2^{S_{21}}(\tau) = \{y_{21}^{S_{21}}(\tau), y_{22}^{S_{21}}(\tau), \dots, y_{2n_2}^{S_{21}}(\tau)\},$$

$$Y_2^{S_{22}}(\tau) = \{y_{21}^{S_{22}}(\tau), y_{22}^{S_{22}}(\tau), \dots, y_{2n_2}^{S_{22}}(\tau)\},$$

...

$$Y_2^{S_{2q_2}}(\tau) = \{y_{21}^{S_{2q_2}}(\tau), y_{22}^{S_{2q_2}}(\tau), \dots, y_{2n_2}^{S_{2q_2}}(\tau)\},$$

...

$$Y_k^{S_{k1}}(\tau) = \{y_{k1}^{S_{k1}}(\tau), y_{k2}^{S_{k1}}(\tau), \dots, y_{kn_k}^{S_{k1}}(\tau)\},$$

$$Y_k^{S_{k2}}(\tau) = \{y_{k1}^{S_{k2}}(\tau), y_{k2}^{S_{k2}}(\tau), \dots, y_{kn_k}^{S_{k2}}(\tau)\},$$

...

$$Y_k^{S_{kq_k}}(\tau) = \{y_{k1}^{S_{kq_k}}(\tau), y_{k2}^{S_{kq_k}}(\tau), \dots, y_{kn_k}^{S_{kq_k}}(\tau)\},$$

которые могут (или не могут) быть достигнуты на интервалах времени $\tau_{\min 1} \leq \tau \leq \tau_{\max 1}$; $\tau_{\min 2} \leq \tau \leq \tau_{\max 2}$; ...; $\tau_{\min k} \leq \tau \leq \tau_{\max k}$ соответственно.

© Становский А. Л., Лысенко Т. В., Носенко Т. И., Лысенко Е. В., 2007

Назовем эти поименованные состояния системы Ω событиями $\Sigma\{S_{11}, S_{12}, \dots, S_{1q_1}; S_{21}, S_{22}, \dots, S_{2q_2}; \dots; S_{k1}, S_{k2}, \dots, S_{kq_k}\}$, причем для каждой из подсистем Y_1, Y_2, \dots, Y_k в отдельности будем считать события, происходящие в них, несовместными:

$$\begin{aligned} S_{11} \cap S_{12} \cap \dots \cap S_{1q_1}; \\ S_{21} \cap S_{22} \cap \dots \cap S_{2q_2}; \\ \dots \\ S_{k1} \cap S_{k2} \cap \dots \cap S_{kq_k}. \end{aligned}$$

Поскольку в процессе движения системы Ω во времени и пространстве своих состояний на время наступления событий, кроме учитываемых, влияет множество других факторов и помех, события $S_{11}, S_{12}, \dots, S_{1q_1}; S_{21}, S_{22}, \dots, S_{2q_2}; \dots; S_{k1}, S_{k2}, \dots, S_{kq_k}$ наступают с вероятностями $p_{11}, p_{12}, \dots, p_{1q_1}; p_{21}, p_{22}, \dots, p_{2q_2}; \dots; p_{k1}, p_{k2}, \dots, p_{kq_k}$.

Дополним перечень событий в каждой из подсистем Y_1, Y_2, \dots, Y_k «нулевыми» событиями $S_{01}, S_{02}, \dots, S_{0k}$, означающими, что в соответствующей подсистеме не произошло ни одного из перечисленных событий $S_{11}, S_{12}, \dots, S_{1q_1}; S_{21}, S_{22}, \dots, S_{2q_2}; \dots; S_{k1}, S_{k2}, \dots, S_{kq_k}$. Вероятности нулевых событий обозначим через $p_{10}, p_{20}, \dots, p_{k0}$. Тогда

$$\begin{aligned} p_{11} + p_{12} + \dots + p_{1q_1} + p_{10} &= 1; \\ p_{21} + p_{22} + \dots + p_{2q_2} + p_{20} &= 1; \\ \dots & \\ p_{k1} + p_{k2} + \dots + p_{kq_k} + p_{k0} &= 1. \end{aligned}$$

Отметим также то, что события в пределах одной подсистемы считаются несовместными, а в пределах разных подсистем – независимыми.

На основе изложенного, а также, принимая за аксиому определение вероятности, приведенное, например, в [3], определим событие S как идеализированный элемент некоторого класса Σ , на котором определена однозначная действительная функция $P\{S\}$, удовлетворяющая условиям:

- 1) $P\{S\} > 0$ для любого S из Σ ;
- 2) $P\{I\} = 1$ для достоверного события;
- 3) $P\{S_1 \cup S_2 \cup \dots\} = P\{S_1\} + P\{S_2\} + \dots$ для любой (конечной или бесконечной) последовательности попарно несовместных событий S_1, S_2, \dots

ОПРЕДЕЛЕНИЕ СОБЫТИЯ

Определение события как «поименованного состояния» позволяет классифицировать их в зависимости от содержания поименования.

В простейшем случае такое поименование представляет собой назначение конкретных числовых значений переменным вектора $\mathbf{Y}^S(\tau)$: $y_1^S(\tau) = a_1, y_2^S(\tau) = a_2, \dots, y_n^S(\tau) = a_n$. При этом набор чисел a_1, a_2, \dots, a_n обуславливает какое-либо отношение. Например, событие S «концентрация угарного газа СО в рабочей зоне равна 50 мг/м^3 » определяет отношение «концентрация СО не опасна для здоровья работников цеха». Назовем такое состояние *событием первого рода* [4].

Иногда количественные изменения переменных вектора $\mathbf{Y}^S(\tau)$ приводят к качественным переходам в системе. К таким качественным преобразованиям могут быть отнесены, например, фазовые превращения (плавление, кристаллизация, кипение, конденсация и т. п.), исчезновение или возникновение фаз (выпадение осадка, растворение и т. п.), прорывы газовых пузырей, скачкообразное изменение цвета и многое другое. Набор значений чисел a_1, a_2, \dots, a_n , обуславливающих качественное изменение, назовем *событием второго рода* [5].

Событиями третьего рода будем считать экстремумы функций $y_1^S(\tau), y_2^S(\tau), \dots, y_n^S(\tau)$.

СИНХРОНИЗАЦИЯ ЭКСТРЕМУМОВ

Очевидно, для достижения цели синхронизирующего управления в акте синхронизации может одновременно участвовать произвольное количество событий *разного* рода.

Во всех случаях синхронизирующего управления соответствующая АСУ должна иметь в своем составе прогнозирующий компонент, экстраполирующий во времени значения переменных вектора $\mathbf{Y}^S(\tau)$. Для этого функции $y_1^S(\tau), y_2^S(\tau), \dots, y_n^S(\tau)$ должны быть определены на интервале времени, априори включающем будущее время синхронизации. Решение такой задачи аналитически не всегда возможно, поэтому в некоторых случаях (например, для временных трендов) экстраполяция может быть выполнена с помощью нейронной сети (НС).

В последнем случае наибольшие проблемы возникают при прогнозировании событий третьего рода, поскольку НС испытывают затруднения при прогнозировании экстремумов, так как для идентификации экстремума нужны, как минимум, три точки экстраполяции. Так, например, в «прямом» эксперименте НС, обученной методом «скользящего окна», не удалось предсказать максимум тестового тренда, представленного на рис. 1, а. Тренд содержит 13 «предыдущих» точек (на рисунке показаны последние 5), одну «текущую» и три «будущие», причем, среди последних – подлежащий прогнозированию максимум в точке

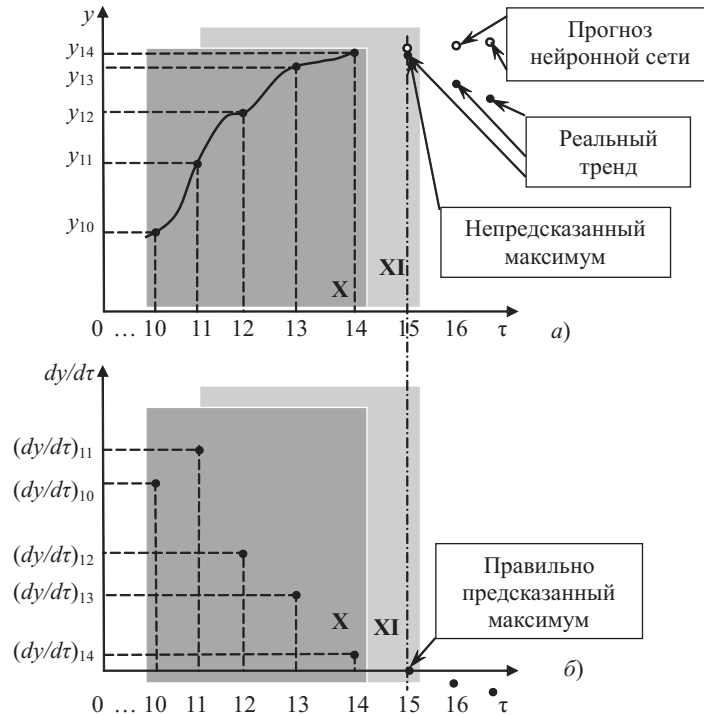


Рисунок 1 – Нейросеточное прогнозирование максимума тренда в точке 15 методом «скользящего окна» по значениям производных интерполяционного полинома в узловых точках 0...14:

X – десятое (последнее) «окно» обучающей выборки; XI – одиннадцатое «окно», на котором должен быть предсказан максимум

$\tau = 15$. Четырнадцать имеющихся точек позволили построить десять скользящих окон (по пять точек в окне) и обучить по ним нейронную сеть прямого распространения структуры $5 \times 20 \times 1$. Однако предсказанные значения тренда в «будущих» точках 15, 16 и 17, к сожалению, не позволили выявить максимум (пунктирная линия).

Тогда был предложен «косвенный» метод поиска экстремума дискретного тренда, алгоритм которого выглядит следующим образом.

1. Подлежащий экстраполяции тренд аппроксимируется в области своего существования функцией $y(\tau)$; $0 \leq \tau \leq 17$ степенным полиномом вида (рис. 1, а):

$$g(\tau) = \sum_{k=0}^{13} a_k \tau^k, \quad (1)$$

проходящего через точки исследуемого тренда:

$$g(\tau_k) = y_k; \quad k = 0, 1, \dots, 13. \quad (2)$$

Величина k принята равной 13, т. к. количество узлов интерполяционного полинома (они же – суть точки исследуемого тренда) должно быть на единицу больше его степени.

2. Определяется производная интерполяционного полинома $dg(\tau)/d\tau$ и рассчитываются ее значения в точках 0...14 (рис. 1, б).

3. По этим значениям формируются десять «скользящих окон», обучается та же НС и осуществляется прогнозирование.

4. Выбирается значение производной интерполяционного полинома, наиболее близкое к оси τ , и соответствующая ему точка тренда признается экстремумом последнего.

ПРАКТИЧЕСКОЕ ПРИМЕНЕНИЕ

Практическое применение данного метода было рассмотрено на примере управления литьем в песчано-смоляные формы. В качестве объекта были выбраны процессы теплообмена в системе «отливка – форма» после заливки металла. В объекте были выделены термическая и массовая подсистемы, событиями в этих подсистемах, соответственно, были приняты: достижение температурой отливки значения начала кристаллизации (событие второго рода) и достижение давлением газов в порах формы максимума (событие третьего рода).

Управление обеими подсистемами процесса тепло-массообмена осуществлялось с помощью внешнего воздействия на систему «отливка – форма» путем регулирования интенсивности принудительной аспирации газов из формы.

Применение синхронизирующего управления процессами литья позволило выявить технологические методы повышения качества отливок, а их практическая реализация на нескольких предприятиях машиностроения – получить положительный технико-экономический эффект.

ПЕРЕЧЕНЬ ССЫЛОК

1. Становский А. Л., Лысенко Т. В. Использование муар-эффекта при синхронизации событий // Труды Одесского политехнического университета. – 2006. – Вып. 1 (25). – С. 114–118.
2. Маркович Э. С. Курс высшей математики с элементами теории вероятностей и математической статистики. – М.: Высшая школа, 1972. – 480 с.
3. Корн Г., Корн Т. Справочник по математике. – М.: Наука, 1974. – 831 с.
4. Найдек В. Л., Становский А. Л., Лысенко Т. В. Оптимизация процессов в системе «отливка – форма» за счет синхронизации событий // Материалы XIII семинара «Моделирование в прикладных научных исследованиях». – Одесса: ОНПУ, 2006. – С. 3–5.
5. Становский А. Л., Лысенко Т. В., Носенко Т. И. Синхронизация событий при работе систем автоматизированного управления // Материалы XIV семинара «Моделирование в прикладных научных исследованиях». – Одесса: ОНПУ, 2007. – С. 43–45.

Надійшла 28.03.07
Після доробки 15.06.07

Визначено поняття «подія» у теорії синхронізуючого керування, як поіменованій стан. Виконано класифікацію подій. Запропоновано метод нейросеточної екстраполяції дискретних числових тимчасових трендів, заснований на заміні значень тренда у вузлових крапках на значення похідної інтерполяційного полінома в цих же крапках.

A concept of «event» is defined in synchronizing control theory as nominal condition. Classification of events is made. A method of a neuron grid extrapolation of discrete numerical temporal trends based on trend value replacement in nodal points by interpolation polynom derivative in the same points is suggested.

НЕЙРОІНФОРМАТИКА ТА ІНТЕЛЕКТУАЛЬНІ СИСТЕМИ

НЕЙРОИНФОРМАТИКА И ИНТЕЛЛЕКТУАЛЬНЫЕ СИСТЕМЫ

NEUROINFORMATICS AND INTELLIGENT SYSTEMS

УДК 519. 7:004.8

Е. В. Бодянский, Н. Е. Кулишова

ИНТЕРПОЛИРУЮЩАЯ НЕЙРО-ФАЗЗИ-СЕТЬ ДЛЯ МОДЕЛИРОВАНИЯ ЦВЕТОПЕРЕДАЧИ ПЕЧАТНОЙ СИСТЕМЫ

Предложены архитектура и алгоритм обучения нейро-фаззи системы, предназначенной для решения задачи интерполяции двумерных функций, заданных на произвольно расположенных на плоскости узлах.

ВВЕДЕНИЕ

Высококачественная передача цвета растровой печатью в значительной степени зависит от стабильности размеров растровых точек, так как отклонения этих размеров приводят к искажениям формируемых цветов. Кроме того, известно, что свойства бумаги и особенности динамики работы печатного оборудования также существенно влияют на цвет печатного оттиска. Поэтому принято рассматривать передачу цвета не только в контексте применяемых красителей, но и в совокупности свойств системы «бумага – краска – печатная машина». Наиболее простым и доступным средством исследования цветопередающих свойств подобной системы является печать цветных контрольных шкал и измерение относительной площади растровых точек на их полях. В полиграфической практике информация о площади растровых точек может быть по-

лучена с помощью оценки относительных коэффициентов отражения красочных участков в зависимости от длины волны отраженного света. Ограничением подобного подхода является то, что формирование контрольных шкал с полями, в которых площадь точек отличается на 1 %, будет излишне затратным. Сложившаяся методика сводится к тому, что поля контрольных шкал имеют разницу в площадях растровых точек 5–10 %, а в диапазонах, особенно важных для восприятия цветов, такая разница может составлять 2 %. Однако, контрольные шкалы каждый производитель формирует индивидуально под собственные условия и поэтому возможны ситуации, когда для каких-либо красок в некоторых диапазонах вообще отсутствуют контрольные поля, а в некоторых они расположены с разной плотностью. Тем не менее, по оттискам таких шкал нужно получать как можно более полную информацию о формировании различных красочных оттенков без разрывов в разных тоновых диапазонах.

С математической точки зрения эта задача сводится к интерполяции двумерной нелинейной функции достаточно сложного вида на множестве произвольно

© Бодянский Е. В., Кулишова Н. Е., 2007

заданных на ограниченном участке плоскости узлов, число которых может быть достаточно велико. В общем случае для решения подобной задачи могут быть использованы искусственные нейронные сети, обладающие интерполирующими свойствами, такие, как радиально-базисные сети [1, 2]. Опыт, однако, показал [3, 4], что качество получаемого решения существенным образом зависит от выбора параметров рецепторных полей радиально-базисных функций, который осуществляется методом проб и ошибок. Неудачный выбор приводит к возникновению либо колебаний между узлами интерполяции, либо к появлению «дыр» между соседними радиально-базисными нейронами.

Альтернативой искусственным нейронным сетям являются нейро-фаззи системы [5, 6], сочетающие возможности обучаемости нейронных сетей с прозрачностью и интерпретируемостью систем нечеткого вывода. В общем случае нейро-фаззи системы представляют собой нечеткие модели, обучаемые с помощью алгоритмов, принятых в теории нейронных сетей, и использующие вместо многомерных радиально-базисных функций одномерные функции принадлежности, что значительно проще с вычислительной точки зрения.

Среди нейро-фаззи систем наиболее популярна адаптивная нейро-система нечеткого вывода ANFIS [7], имеющая многослойную архитектуру с прямой передачей информации, синаптические веса и функции принадлежности которой настраиваются путем последовательной оптимизации некоторого наперед заданного критерия обучения. Основным недостатком ANFIS и ей подобных систем [5–8] является низкая скорость обучения.

Преодолеть этот недостаток возможно, используя так называемые системы, основанные на памяти [9–11], обучение которых может быть сведено к одноразовой установке центров колоколообразных функций активации-принадлежности в узлах интерполяции, т. е. фактически за один проход алгоритма обучения. Вместе с тем, вопрос выбора параметров рецепторных полей в этих системах остается открытым.

Целью настоящей работы является синтез нейро-фаззи системы, основанной на памяти, предназначенной для решения задачи двумерной интерполяции по информации, заданной в узлах, произвольным образом расположенных на плоскости, и некритичной к выбору параметров рецепторных полей функций принадлежности.

1 МАТЕМАТИЧЕСКАЯ ПОСТАНОВКА ЗАДАЧИ

Пусть в узлах, произвольно расположенных на плоскости с координатами $\{x_j, y_j\}$, $j = 1, 2, \dots, l$; заданы значения функции $f(x_j, y_j) \equiv f_j(x, y)$. Необходимо по-

строить достаточно гладкое приближение $F(x, y)$ исходной функции, строго совпадающее в узлах $\{x_j, y_j\}$ со значениями функции $f_j(x, y) = F(x_j, y_j)$.

2 АРХИТЕКТУРА ИНТЕРПОЛИРУЮЩЕЙ НЕЙРО-ФАЗЗИ СИСТЕМЫ

Архитектура интерполирующей нейро-фаззи системы приведена на рис. 1 и состоит из пяти последовательно соединенных слоев.

Нулевой слой системы является рецепторным и на него подаются преобразованные согласно соотношениям

$$\tilde{x} = \frac{x - x_{\min}}{x_{\max} - x_{\min}}, \quad \tilde{y} = \frac{y - y_{\min}}{y_{\max} - y_{\min}} \quad (1)$$

входные сигналы. Здесь x_{\min} , y_{\min} , x_{\max} , y_{\max} являются нижними и верхними границами изменения соответствующих координат. Обратное преобразование при этом имеет вид

$$x = x_{\max}\tilde{x} + x_{\min}(1 - \tilde{x}), \quad y = y_{\max}\tilde{y} + y_{\min}(1 - \tilde{y}).$$

Первый скрытый слой состоит из $2l$ блоков, реализующих функции принадлежности, и выполняет фаззификацию входных сигналов. Второй скрытый слой осуществляет агрегирование уровней принадлежности, вычисленных в первом скрытом слое, и состоит из l блоков умножения. Третий скрытый слой – это слой синаптических весов, определяемых в процессе обучения системы. Четвертый слой образован двумя сумматорами и вычисляет суммы выходных сигналов второго и третьего слоев. И, наконец, в пятом (выходном) слое производится нормализация, в результате которой вычисляется выходной сигнал нейро-фаззи системы $F(\tilde{x}, \tilde{y})$.

Как видно, данная архитектура по сути совпадает с архитектурой системы Ванга-Менделя [12, 13], одна-

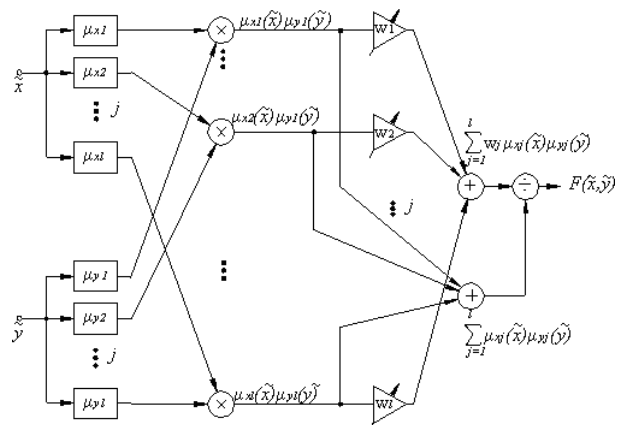


Рисунок 1 – Интерполирующая нейро-фаззи система

ко принцип ее обучения, рассматриваемый ниже, существенно отличается от обратного распространения ошибок, принятого Дженгом [5, 7, 8], Такаги-Сугено [14], Вангом-Менделем [12, 13] и ставших уже традиционными нейро-фаззи системах.

3 ОБУЧЕНИЕ ИНТЕРПОЛИРУЮЩЕЙ НЕЙРО-ФАЗЗИ СИСТЕМЫ

Поскольку рассматриваемая конструкция относится к системам, основанным на памяти, ее настройка производится с помощью одного прохода алгоритма обучения, что делает ее крайне простой и быстрой.

Предварительно преобразованные на единичный квадрат с помощью соотношений (1) данные обучающей выборки $\{\tilde{x}_j, \tilde{y}_j\}$ формируют в первом скрытом слое набор функций принадлежности μ_{xi}, μ_{yi} ($i = 1, 2, \dots, l$ и в общем случае не совпадает с j) так, что их центры определяются точками \tilde{x}_j, \tilde{y}_j по принципу «нейроны в точках данных» [15]. Процесс формирования функций принадлежности по координате \tilde{x} иллюстрируется рис. 2.

Здесь необходимо отметить, что если индексы точек j задаются произвольным образом, то индексы функций принадлежности и их центров в процессе формирования упорядочены так, что $0 = c_{x1} < c_{x2} < \dots < c_{xi} < \dots < c_{xl} = 1$. Функции принадлежности по координате \tilde{y} формируются аналогичным образом.

В простейшем случае функции принадлежности могут быть приняты треугольными, при этом

$$\mu_{x1} = \frac{c_{x2} - \tilde{x}}{c_{x2}}, \quad \tilde{x} \in [0, c_{x2}], \quad (2)$$

$$\mu_{xi} = \begin{cases} \frac{\tilde{x} - c_{x,i-1}}{c_{x,i} - c_{x,i-1}}, \tilde{x} \in [c_{x,i-1}, c_{x,i}]; \\ \frac{c_{x,i+1} - \tilde{x}}{c_{x,i+1} - c_{x,i}}, \tilde{x} \in [c_{x,i}, c_{x,i+1}], i = 2, \dots, l-1, \end{cases} \quad (3)$$

$$\mu_{xn} = \frac{\tilde{x} - c_{x,l-1}}{1 - c_{x,l-1}}, \quad \tilde{x} \in [c_{x,l-1}, 1], \quad (4)$$

$$\mu_{y1} = \frac{c_{y2} - \tilde{y}}{c_{y2}}, \quad \tilde{y} \in [0, c_{y2}], \quad (5)$$

$$\mu_{yi} = \begin{cases} \frac{\tilde{y} - c_{y,i-1}}{c_{y,i} - c_{y,i-1}}, \tilde{y} \in [c_{y,i-1}, c_{y,i}]; \\ \frac{c_{y,i+1} - \tilde{y}}{c_{y,i+1} - c_{y,i}}, \tilde{y} \in [c_{y,i}, c_{y,i+1}], i = 2, \dots, l-1, \end{cases} \quad (6)$$

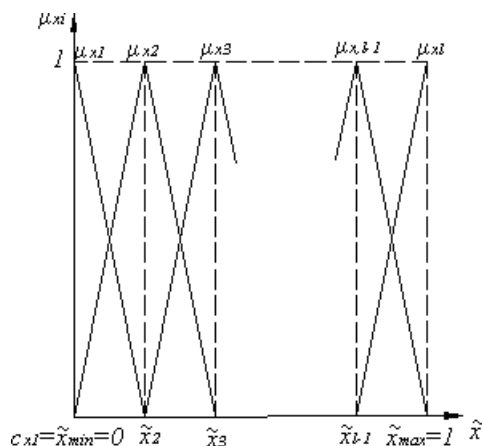


Рисунок 2 – Функции принадлежности интерполирующей нейро-фаззи системы

$$\mu_{yl} = \frac{\tilde{y} - c_{y,l-1}}{1 - c_{y,l-1}}, \quad \tilde{y} \in [c_{y,l-1}, 1], \quad (7)$$

хотя возможно использование и иных колоколообразных функций.

После формирования $2l$ функций принадлежности несложно вновь перейти от упорядоченных индексов i к индексам j , задающим номер конкретного узла интерполяции, что необходимо для последующих расчетов.

После формирования функций принадлежности μ_{xj}, μ_{yj} первого скрытого слоя, в третьем слое происходит установка синаптических весов w_j , которые полагаются равными значениям интерполируемой функции $f_j(x, y)$. Заметим, что такой принцип определения весов использован в обобщенной регрессионной нейронной сети Шпехта [10], в основе которой лежат идеи Парзеневских окон [16] и ядерных оценок Надарая-Ватсона [17, 18].

Таким образом, при подаче на вход системы произвольного сигнала \tilde{x}, \tilde{y} в первом скрытом слое вычисляются уровни принадлежности $\mu_{xj}(\tilde{x}), \mu_{yj}(\tilde{y}), j = 1, 2, \dots, l$, во втором слое производится их агрегирование путем формирования двумерных фаззи-базисных функций

$$\mu_j(\tilde{x}, \tilde{y}) = \mu_{xj}(\tilde{x})\mu_{yj}(\tilde{y}), \quad (8)$$

в третьем слое вычисляются произведения $f_j(x, y)\mu_{xj}(\tilde{x})\mu_{yj}(\tilde{y})$, четвертый слой вычисляет значения сигналов

$$\sum_{j=1}^l f_j(x, y)\mu_{xj}(\tilde{x})\mu_{yj}(\tilde{y}) = \sum_{j=1}^l w_j\mu_{xj}(\tilde{x})\mu_{yj}(\tilde{y}), \quad (9)$$

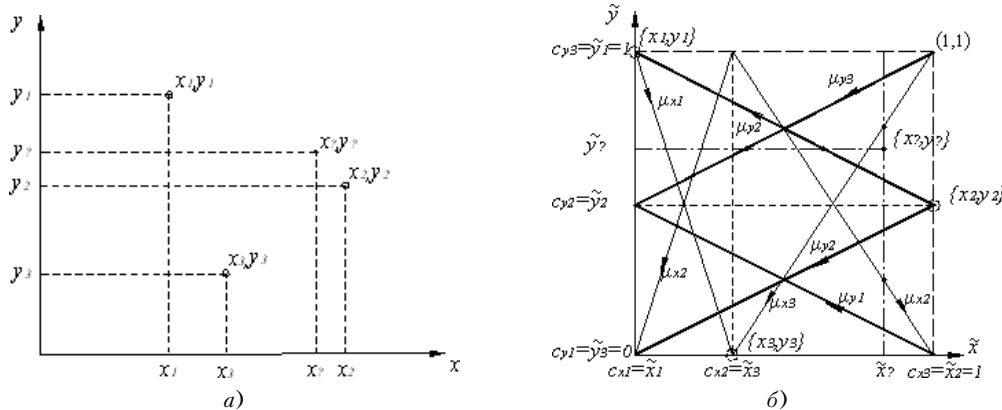


Рисунок 3 – Пример интерполяции

$$\sum_{j=1}^l \mu_j(\tilde{x}, \tilde{y}) = \sum_{j=1}^l \mu_{x_{ij}}(\tilde{x})\mu_{y_{ij}}(\tilde{y}), \quad (10)$$

и, наконец, в выходном слое рассчитывается интерполированное значение функции

$$F(\tilde{x}, \tilde{y}) = \frac{\sum_{j=1}^l f_j(x, y)\mu_{x_j}(\tilde{x})\mu_{y_j}(\tilde{y})}{\sum_{j=1}^l \mu_{x_j}(\tilde{x})\mu_{y_j}(\tilde{y})}. \quad (11)$$

Процесс функционирования интерполирующей нейро-фаззи системы иллюстрируется примером на рис. 3, где в верхней части показаны три узла интерполяции $\{x_1, y_1\}$, $\{x_2, y_2\}$, $\{x_3, y_3\}$ с известными значениями $f(x_1, y_1)$, $f(x_2, y_2)$, $f(x_3, y_3)$ и точка с координатами $\{x_2, y_2\}$, в которой необходимо провести интерполяцию.

Проекции точек $\{x_j, y_j\}$ на оси x и y расположены так, что $x_1 < x_3 < x_2$ и $y_3 < y_2 < y_1$. В нижней части рисунка эти же узлы, подвергнутые преобразованию (1), обозначены $\{\tilde{x}_1, \tilde{y}_1\}$, $\{\tilde{x}_2, \tilde{y}_2\}$, $\{\tilde{x}_3, \tilde{y}_3\}$, $\{\tilde{x}_2, \tilde{y}_2\}$ и могут быть расположены либо в вершинах, либо на сторонах единичного квадрата. В данной ситуации архитектура, приведенная на рис. 1, имеет 6 функций принадлежности. С проекциями узлов $\{\tilde{x}_j, \tilde{y}_j\}$ на оси \tilde{x} , \tilde{y} связываются центры функций принадлежности так, что $c_{x1} = \tilde{x}_1 = 0$, $c_{x2} = \tilde{x}_3$, $c_{x3} = \tilde{x}_2 = 1$ и $c_{y1} = \tilde{y}_3 = 0$, $c_{y2} = \tilde{y}_2$, $c_{y3} = \tilde{y}_1 = 1$. С каждым из центров, в свою очередь, связываются свои функции принадлежности μ_{x_i}, μ_{y_i} ($\mu_{x1}, \mu_{x2}, \mu_{x3}, \mu_{y1}, \mu_{y2}, \mu_{y3}$), также приведенные на рис. 3 и описываемые соотношениями (2)–(7). После формирования всех функций принадлежности они могут быть переиндексированы к виду

μ_{x_j}, μ_{y_j} ($\mu_{x1}, \mu_{x3}, \mu_{x2}, \mu_{y3}, \mu_{y2}, \mu_{y1}$). В результате работы системы в точке с координатами $\{\tilde{x}_2, \tilde{y}_2\}$ будет вычислено значение

$$F(\tilde{x}_2, \tilde{y}_2) = \frac{\sum_{j=1}^3 f(x_j, y_j)\mu_{x_j}(\tilde{x}_2)\mu_{y_j}(\tilde{y}_2)}{\sum_{j=1}^3 \mu_{x_j}(\tilde{x}_2)\mu_{y_j}(\tilde{y}_2)}, \quad (12)$$

определяемое ближайшими к точке $\{\tilde{x}_2, \tilde{y}_2\}$ узлами интерполяции.

4 РЕЗУЛЬТАТЫ ЭКСПЕРИМЕНТАЛЬНОГО ИССЛЕДОВАНИЯ

Для исследования интерполирующих свойств нейро-фаззи системы в задаче интерполяции спектральных коэффициентов отражения красок использованы результаты спектрофотометрических измерений коэффициента отражения полиграфической тестовой шкалы. Эта шкала содержит четыре ряда полей, отпечатанных четырьмя красками синтеза, каждое из полей запечатано растровыми точками разных размеров – от 1 до 100 % заполнения относительной площади ячейки. Полям соответствуют следующие значения относительной запечатанной площади: 1 %, 2 %, 3 %, 4 %, 5 %, 10 %, 20 %, 30 %, 40 %, 50 %, 60 %, 70 %, 80 %, 90 %, 92 %, 94 %, 96 %, 98 %, 100 %. Измерения проводятся прибором автоматически в видимом диапазоне излучений, что соответствует длинам волн от 380 до 730 нм. Из результатов измерений случайным образом отброшены 30 % данных.

Проведена интерполяция коэффициента отражения с помощью нейро-фаззи системы. В качестве функций принадлежности выбраны треугольные функции (2)–(7). На рис. 4–7 показаны результаты интерполяции коэффициента отражения для разных красок синте

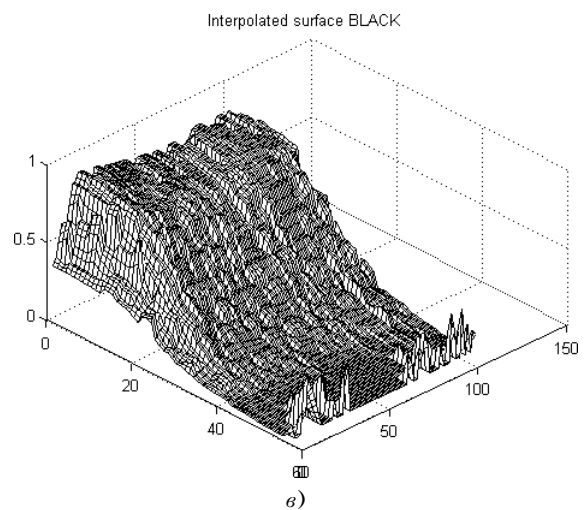
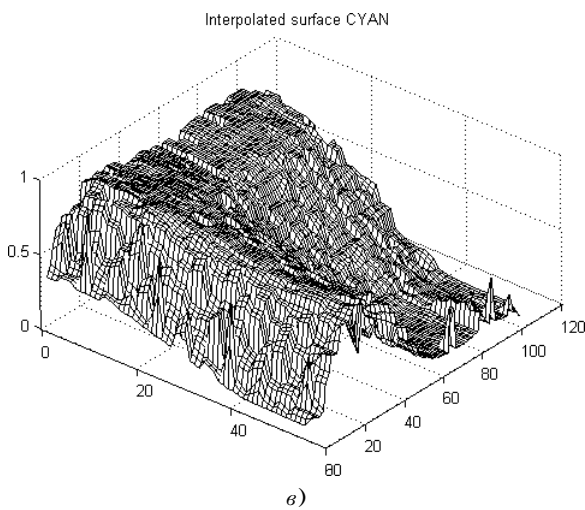
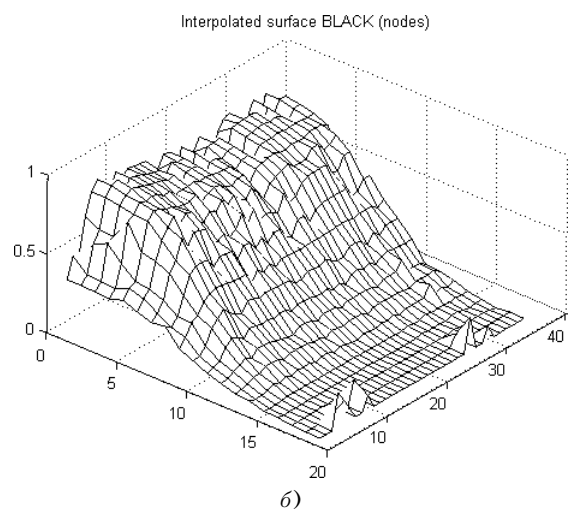
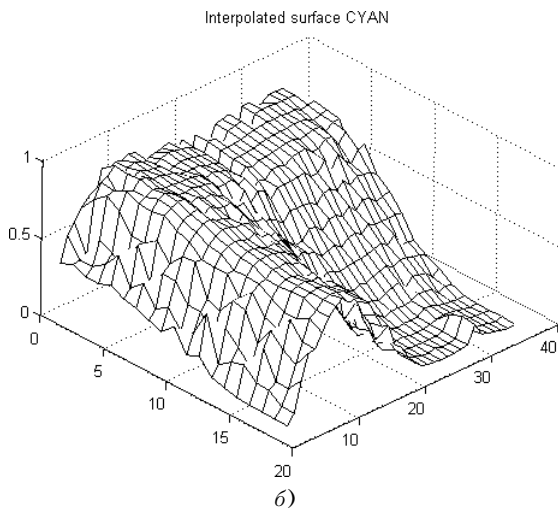
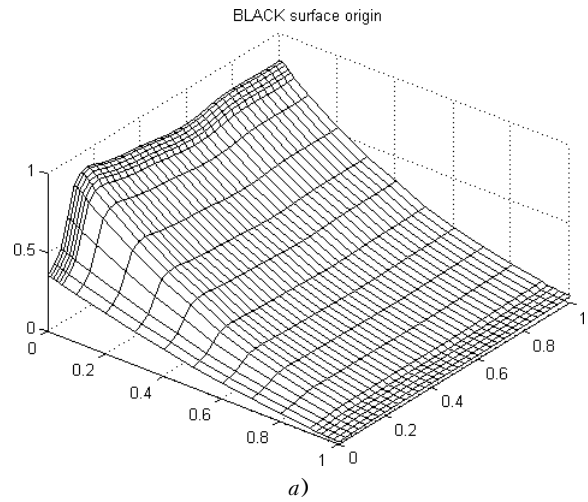
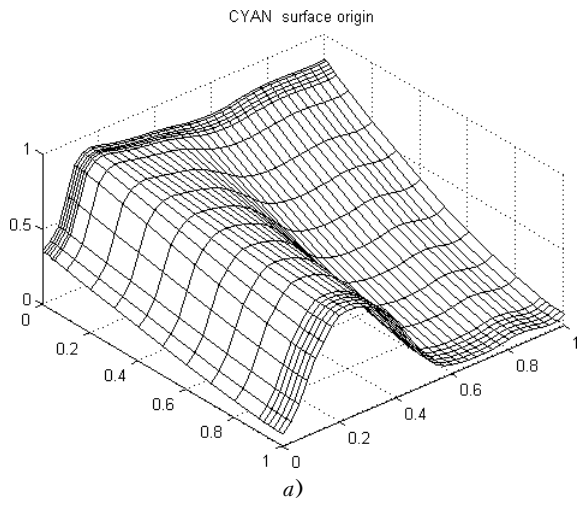


Рисунок 4 – Коэффициент отражения голубой краски:

a – исходные данные; б – узловые точки интерполированной поверхности; в – поверхность, интерполированная между узловыми точками

Рисунок 5 – Коэффициент отражения черной краски:

a – исходные данные; б – узловые точки интерполированной поверхности; в – поверхность, интерполированная между узловыми точками

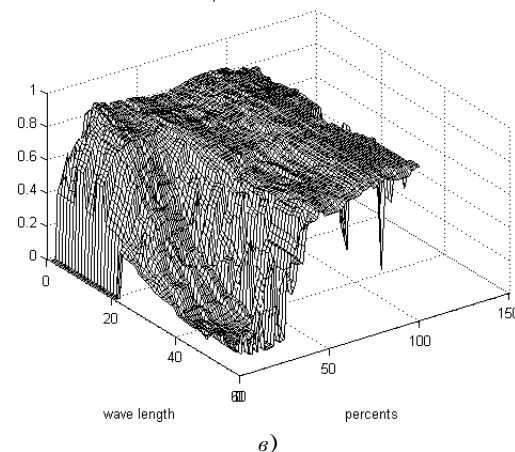
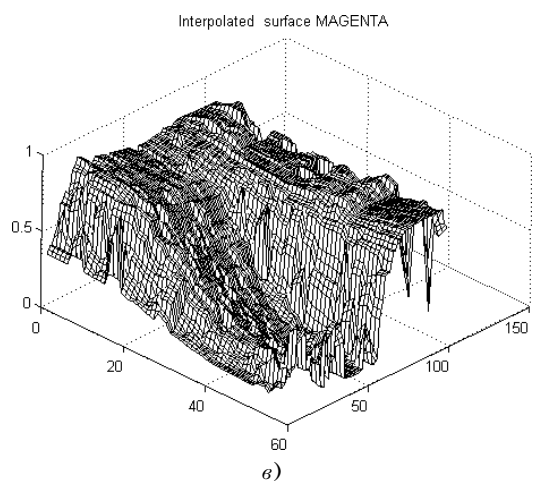
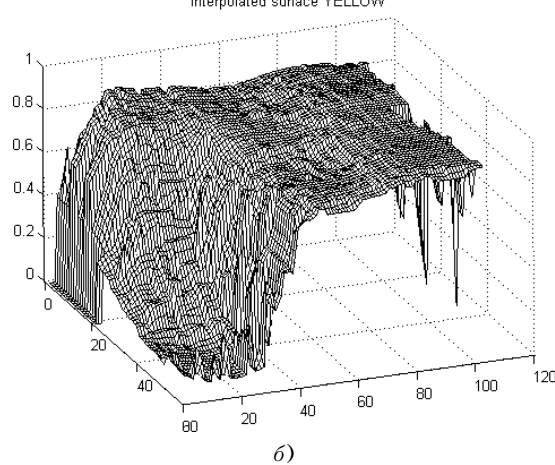
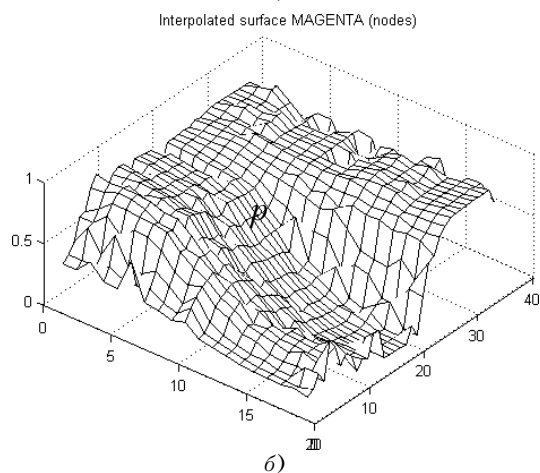
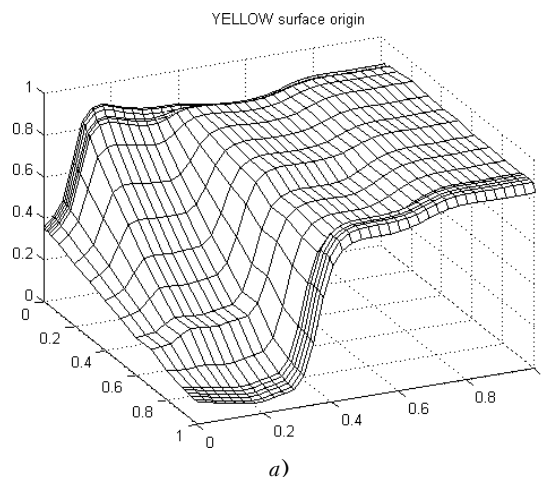
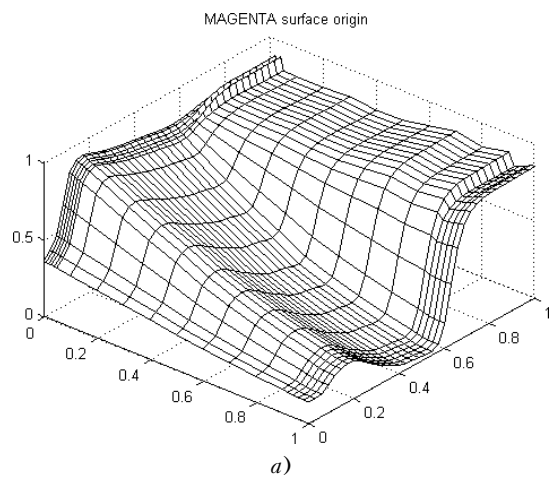


Рисунок 6 – Коэффициент отражения пурпурной краски:

а – исходные данные; б – узловые точки интерполированной поверхности; в – поверхность, интерполированная между узловыми точками

Рисунок 7 – Коэффициент отражения желтой краски:

а – исходные данные; б – узловые точки интерполированной поверхности; в – поверхность, интерполированная между узловыми точками

Таблица 1 – Точность интерполяции коэффициентов отражения разных красок в узловых точках

Краска	Голубая	Черная	Пурпурная	Желтая
Среднеквадратичное отклонение E	0,0385	0,0475	0,0463	0,0413

была оценена с помощью среднеквадратичного отклонения и приведена в табл. 1.

Поскольку максимальное значение коэффициента отражения составляет 1, точность интерполяции с помощью нейро-фаззи системы не превышает 5 %, что является удовлетворительным для большинства практических задач.

ЗАКЛЮЧЕНИЕ

Предложенная нейро-фаззи система обеспечивает высокое качество интерполяции двумерных функций, заданных на произвольно расположенных на плоскости узлах, проста в численной реализации, характеризуется высокой скоростью обучения и может найти применение в решении широкого круга задач. В частности, вычислительная простота алгоритма позволяет применять его при подготовке цветных изображений к отображению с использованием профилей печатных устройств в реальном времени.

ПЕРЕЧЕНЬ ССЫЛОК

1. Bishop C. M. *Neural Networks for Pattern Recognition*. – Oxford: Clarendon Press, 1995. – 482 p.
2. Hristev R. M. *The ANN Book*. – 1998. – 374 p.
3. Sigitani T., Iigumi G., Maeda H. Image interpolation for progressive transmission by using radial-basis functions networks // *IEEE Trans. on Neural Networks*. – 1999. – 10. – Pp. 381–390.
4. Кулишова Н. Е. Интерполяция коэффициентов отражения красок с помощью радиально-базисной искусственной нейронной сети // *Бионика интеллекта*. – 2006. – № 1(64). – С. 105–109.
5. Jang J. – S.R., Sun C. – T., Mizutani E. *Neuro-Fuzzy and Soft Computing*. – Upper Saddle River, NJ: Prentice Hall, 1997. – 614 p.

6. Cios K. J., Pedrycz W. *Neuro-fuzzy algorithms* / In: «Handbook on Neural Computation». – Oxford: IOP Publishing and Oxford University Press, 1997. – D1.3:1–D1.3: 7.
7. Jang J. – S.R. ANFIS: Adaptive-Network-based Fuzzy Inference System // *IEEE Trans. on Systems, Man and Cybernetics*. – 1993. – 23. – Pp. 665–685.
8. Jang J. – S.R., Sun C. – T., Mizutani E. Neuro-fuzzy modeling and control // *Proc. IEEE*. – 1995. – 83. – Pp. 378–406.
9. Brown M., Harris C. J. *Neural networks for modeling and control* / In: Ed. by C. J. Harris «Advances in Intellectual Control». – London: Taylor and Francis, 1994. – Pp. 17–55.
10. Specht D. A general regression neural network // *IEEE Trans. on Neural Networks*. – 1991. – 2. – Pp. 568–576.
11. Wang H., Liu G. P., Harris C. J., Brown M. *Advanced Adaptive Control*. – Oxford: Pergamon, 1995. – 262 p.
12. Wang L., Mendel J. M. Fuzzy basis functions, universal approximation and orthogonal least squares learning // *IEEE Trans. on Neural Networks*. – 1992. – 3. – Pp. 807–814.
13. Wang L. – X., Mendel J. M. Generating fuzzy rules by learning from examples // *IEEE Trans. on Systems, Man and Cybernetics*. – 1992. – 22. – Pp. 1414–1427.
14. Takagi T., Sugeno M. Fuzzy identification of systems and its application to modeling and control // *IEEE Trans. on Systems, Man and Cybernetics*. – 1985. – 15. – Pp. 116–132.
15. Zahirniak D. R., Chapman R., Rogers S. K., Suter B. W., Kabrisky M., Pyati V. Pattern recognition using radial basis function network // *Proc. 6-th Ann. Aerospace Application of AI Conf.* – Dayton, OH, 1990. – Pp. 249–260.
16. Parzen E. On the estimation of a probability density function and the mode // *Ann. Math. Statist.* – 1962. – 38. – Pp. 1065–1076.
17. Надарая Э. А. О непараметрических оценках плотности вероятности и регрессии // *Теория вероятностей и ее применение*. – 1965. – 10. – № 1. – С. 199–203.
18. Friedman J., Hastie J., Tibshirani R. *The Elements of Statistical Learning, Data Mining, Inference, and Prediction*. – Berlin: Springer, 2003. – 552 p.

Надійшла 22.10.07

Запропоновано архітектуру та алгоритм навчання нейро-фаззи системи, яка призначена для розв'язання задачі інтерполяції двовимірних функцій, що задані на довільно розташованих на площині вузлах.

The architecture and learning algorithm for the neuro-fuzzy system is proposed. This system is aimed for decision of the interpolation task of two-variable functions, that are known in nodes, which are arbitrary placed on the plane.

УДК 004.032.26

Т. В. Киприч, В. И. Дубровин

АНАЛИЗ МОДИФИКАЦИЙ САМООРГАНИЗУЮЩИХСЯ КАРТ КОХОНЕНА ПО КРИТЕРИЯМ РЕГУЛЯРНОСТИ И ТОЧНОСТИ АППРОКСИМАЦИИ

Рассматривается проблема взаимовлияния регулярности и точности аппроксимации сети среди различных модификаций самоорганизующихся карт Кохонена. Приводится сравнительный анализ данных модификаций по количеству циклов обучения карты, ошибкам квантования и топографического упорядочивания нейронов.

© Киприч Т. В., Дубровин В. И., 2007

ВВЕДЕНИЕ

Построение самоорганизующейся карты Кохонена (СОК), адекватно описывающей исходное многообразие данных, качественно связано со значениями топографической ошибки и погрешности аппроксимации.

Однако минимизация одного из параметров вызывает определенное противоречие, поскольку сохранение топологических особенностей выборки приводит к ухудшению качества векторного квантования и наоборот [1]. Подобный эффект наблюдается независимо от соответствия структуры карты распределению обучаемого множества. При этом точность аппроксимации зависит от векторного квантования, а регулярность определяет меру сохранения особенностей топологии выборки.

Основной целью векторного квантования является извлечение векторов прототипов из множества данных, которые представляют и аппроксимируют многообразие данных наилучшим образом. При поступлении на вход устройства квантования нового вектора, в первую очередь определяется область, к которой он принадлежит, а после этого создается представление через вектор воспроизведения (прототип) данной области [2]. Таким образом, входное пространство подразделяется на множество четких областей, для каждой из которых определяется свой вектор воспроизведения.

Топологическая упорядоченность в терминах СОК означает, что соседние компоненты в латентном пространстве являются также ближайшими в исходном многообразии [3]. Данное свойство является прямым следствием уравнения перемещения весового вектора победившего нейрона и его окружения в сторону входных данных. С помощью топографической ошибки оценивается величина локального искажения карты (степень регулярности) по расположению нейрона относительно его окружения.

Проведенные ранее исследования дилеммы «регулярность – точность» показали, что такая закономерность не имеет отношения к реальному конфликту в выборке и возникает вследствие свойств обучения СОК [1].

ПОСТАНОВКА ЗАДАЧИ

Существующие модификации СОК по-разному представляют акценты при решении задачи оптимального расположения нейронов сети. Рассмотрев алгоритмы основных модификаций, проведем сравнительный анализ качества полученных результатов по соответствующим показателям карты.

АЛГОРИТМЫ ОБУЧЕНИЯ НЕКОТОРЫХ МОДИФИКАЦИЙ СОК

Исходный алгоритм построения СОК (Sequential SOM либо SEQ SOM) можно описать следующим образом [4]:

- 1) инициализация весов нейронов случайными величинами w_j ;
- 2) предвыборка. Выбор вектора x из исходного множества X с определенной вероятностью;

- 3) определение нейрона победителя, используя минимум Евклидова расстояния:

$$c = \operatorname{argmin} \|x - w_j\|; \quad (1)$$

- 4) коррекция векторов синаптических весов всех нейронов в соответствии со следующей формулой:

$$w_j(t+1) = w_j(t) + \alpha h_{j,c}(t)(x - w_j(t)), \quad (2)$$

где $\alpha(t)$ – параметр скорости обучения; $h_{j,c}(t)$ – функция окрестности с центром в победившем нейроне;

- 5) повторение шагов 2–4 до тех пор, пока на карте признаков не перестанут происходить заметные изменения.

Обучение сети обычно выполняется в два этапа [5]:

1. Этап 1. Ordering / Rough-Tuning. На первом этапе обычно используются относительно большие значения начальной скорости обучения и радиуса настройки соседей.

2. Этап 2. Fine-tuning. На втором этапе скорость обучения $\alpha(t) \approx 0,01$, а радиус обучения – выбирается так, чтобы в движении участвовало 2–3 узла.

Рассмотрим несколько характерных модификаций данного алгоритма.

Пакетный алгоритм (batch SOM), также как и SEQ SOM является итеративным, но смещение весовых векторов происходит в конце каждой эпохи обучения, т. е. после того, как карте будут представлены все записи [6]:

$$w_i(t+1) = \frac{\sum_{j=1}^n h_{i,c}(t)x_j}{\sum_{j=1}^n h_{i,c}(t)}. \quad (3)$$

В результате, порядок представления входных образцов не оказывает влияния на окончательную форму карты признаков и пропадает необходимость в зависимости скорости обучения от времени [2]. В последнем случае преимуществом является устранение потенциального источника плохой сходимости, когда коэффициент $\alpha(t)$ подобран не должным образом.

Основной целью Regularized SOM является улучшение гладкости сгенерированной карты. Для этого вводится понятие «идеального положения» нейрона. Поскольку под «хорошей» регулярностью понимается такое местоположение весового вектора, при котором он располагается на прямой линии L , между левым и правым соседями, идеальное положение определяется с помощью ортогональной проекции данного вектора на прямую L . Соответственно, для сетки правило движения нейрона модифицируется таким образом, чтобы помимо смещения в сторону входных данных, весовой

вектор испытывал смещение относительно своего «идеального положения». Таким образом, достигается частичное спрямление линии, соединяющей три соседних узла, и сетка становится более гладкой [1].

Модификация AdSOM сводит величину топографической ошибки практически к нулю за счет локальной настройки радиуса захвата соседей δ_k . В результате карта оказывается менее «изогнутой», что приводит к несколько худшей точности аппроксимации [5].

Алгоритм Density Tracking SOM аналогичен Batch SOM, за исключением того, что процедура смещения узла происходит в сторону более «весомого» соседа, в окрестности которого большее количество точек данных [5]:

$$w_{i+1} = w_i + \varepsilon w_j, \quad (4)$$

где w_j – центр одного из соседних таксонов, в котором содержится максимальное (среди всех соседей) количество точек; ε – некоторый параметр порядка десятых единицы.

В результате применения Density Tracking SOM в областях скопления данных оказывается большее количество узлов, чем в разреженных областях [5].

Таким образом, перечисленные алгоритмы СОК стремятся сделать сетку или более гладкой, или лучше соответствующей локальной структуре данных.

Проведем численное сравнение данных модификаций.

СРАВНИТЕЛЬНЫЙ АНАЛИЗ МОДИФИКАЦИЙ СОК

При исследовании свойств алгоритмов СОК параметрами качества сети были выбраны: погрешность аппроксимации MSE, топографическая ошибка TPE и число итераций.

MSE рассчитывалась, как среднеквадратичное расстояние от точки данных до ближайшего к ней образца [5]:

$$MSE = \sqrt{\frac{1}{N} \sum_{i=1}^N \|X_i - y_{bmu}(X_i)\|^2}, \quad (5)$$

где $y_{bmu}(X_i)$ – ближайший к точке данных X_i образец.

При определении топографической ошибки исходили из того, что для некоторых точек данных ближайший узел сетки и второй по близости не являются соседями. Такие точки были названы неустойчивыми [5]. В результате топографическая ошибка TPE рассчитывалась как отношение числа неустойчивых точек к общему количеству точек исходного многообразия.

Число итераций указывало, сколько раз для каждого метода происходит полное обучение карты выборкой.

В качестве анализируемой информации были выбраны данные из файла 'iris.dat' [7]. Файл содержал 150 строк. Измерения проводились по четырем показателям. Карта, с помощью которой осуществлялось обобщение информации, имела следующие параметры:

- 1) тип – прямоугольный;
- 2) размер 10×10;
- 3) процесс обучения заканчивался при достижении определенного уровня точности настройки карты.

По результатам обработки информации из 'iris.dat' была составлена табл. 1. Значения параметров алгоритмов в данной таблице были зафиксированы, после этапа тонкой настройки, без учета «Rough-tune».

Таблица 1 – Сравнение различных алгоритмов СОК

№	Алгоритм СОК	MSE	TPE	Число итераций
1	SEQ SOM	0,32	0,25	230
2	Batch SOM	0,18	0,28	40
3	Density Tracking SOM	0,12	0,3	22
4	Regularity SOM	0,27	0,074	50
5	Adaptive SOM	0,48	0,061	212
Минимальные значения		0,12	0,061	22

Информация в табл. 1 показывает, что с помощью Density Tracking SOM была получена самая высокая точность аппроксимации, и понадобилось наименьшее количество циклов обучения карты выборкой, а Adaptive SOM определил минимум топографической ошибки. Однако для данных алгоритмов, наряду с оптимальными значениями исследуемых параметров, наблюдается дилемма «регулярность – точность». Расчет соотношений погрешностей MSE и TPE, как $\sqrt{\Delta R}$, приводится в табл. 2.

Таблица 2 – Нахождение оптимального соотношения MSE-TPE

№	Алгоритм СОК	MSE	TPE	$\frac{mse^2+tpe^2}{\Delta R}$	$\sqrt{\Delta R}$
1	SEQ SOM	0,32	0,25	0,16	0,4
2	Batch SOM	0,18	0,28	0,11	0,33
3	Density Tracking SOM	0,12	0,3	0,10	0,32
4	Regularity SOM	0,27	0,074	0,08	0,28
5	Adaptive SOM	0,48	0,061	0,23	0,48
Минимальные значения		0,12	0,061	0,08	0,28

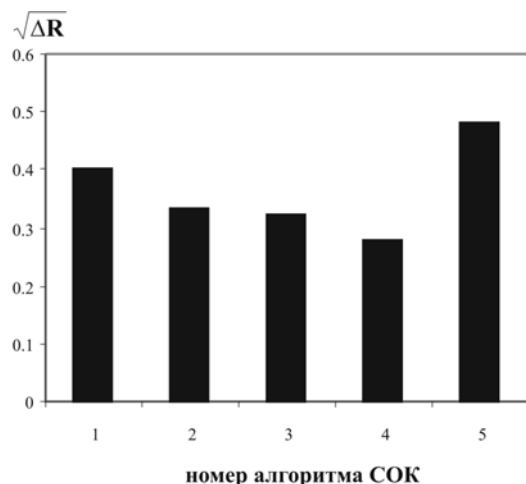


Рисунок 1 – Диаграмма соотношения MSE-TPE для данного эксперимента

На рис. 1 представлена диаграмма изменения $\sqrt{\Delta R}$ для перечисленных модификаций. В результате алгоритм Regularity SOM имеет оптимальное соотношение данных показателей, поскольку обладает минимальным значением $\sqrt{\Delta R}$.

ВЫВОДЫ

Сравнительный анализ величин погрешности аппроксимации и топографической ошибки упорядочивания нейронов среди различных алгоритмов СОК показал что, с помощью Density Tracking SOM была получена самая высокая точность аппроксимации, и понадобилось наименьшее количество циклов обучения карты выборкой, а Adaptive SOM определил минимум топографической ошибки. Однако для перечисленных модификаций, наряду с оптимальными значениями исследуемых параметров, характерна дилемма «регу-

лярность – точность». Результаты расчета наилучшего соотношения погрешности аппроксимации и топографической ошибки упорядочивания нейронов, показали, что карта, построенная с помощью Regularity SOM, обладает оптимальными суммарными показателями.

ПЕРЕЧЕНЬ ССЫЛОК

1. Josef Goppert. Wolfgang Rosenstiel Regularized SOM-Traning: A solution to the Topology-Approximation Dilemma? Режим доступа: <http://citeseer.ist.psu.edu/422534.html>, свободный. – Загл. с экрана. – Англ.
2. Саймон Хайкин. Нейронные сети: полный курс, 2-е издание // Пер. с англ. М.: Издательский дом «Вильямс», 2006. – С. 1104.
3. J. J. Verbeek, N. Vlassis. The generative self-organizing map: a probabilistic generalization of Kohonen's SOM // Technical Report IAS-UVA-02-03 on European Symposium on Artificial Neural Networks 2002, Amsterdam. Доступный режим: <http://citeseer.ist.psu.edu/verbeek02generative.html>, свободный. – Загл. с экрана. – Англ.
4. Juha Vesanto, Johan Himberg. SOM Toolbox for Matlab5. Доступный режим: <http://www.cis.hut.fi/projects/somtoolbox/package/papers/techrep.zip>, свободный. – Загл. с экрана. – Англ.
5. Зиновьев А. Ю. Визуализация многомерных данных: Монография. Красноярск: ИПЦ КГТУ, 2000. – 168 с.
6. Amalendu Roy. A survey on data clustering using self-organizing maps. Доступный режим: <http://www.cs.ndsu.nodak.edu/~amroy/courses.html>, свободный. – Загл. с экрана. – Англ.
7. <http://www.ihes.fr/~zinovyev/vida/vidaexpert.htm>.

Надійшла 5.11.07

Розглядається проблема взаємовпливу регулярності та точності апроксимації сітки серед різних модифікацій власнеорганізованих карт Кохонена. Приводиться порівняльний аналіз даних модифікацій по кількості циклів навчання карти, похибкам квантування й топографічного впорядкування нейронів.

The interference problem of accuracy and regularity of the net among various self-organizing Kohonen's map modifications is considered. The comparative analysis of these modifications over the cycles amount of training map, errors of quantization and topographical neurons ordering is given.

УДК 004.3:004.6

О. Г. Руденко, А. В. Заец, Ю. Э. Ткач

УПРАВЛЕНИЕ НАГРУЗКОЙ В РАСПРЕДЕЛЕННОЙ КОМПЬЮТЕРНОЙ СИСТЕМЕ ПРИ ПОМОЩИ НЕЙРОННЫХ СЕТЕЙ

Разработка диспетчера распределенной компьютерной системы на основе нейронных сетей перцептронного типа разрешила упростить процесс управления ресурсами распределенной компьютерной системы благодаря предложенной системе оценивания задач по критерию необходимых ресурсов. Использование нейронных сетей позволяет сократить время оценивания и повысить точность

© Руденко О. Г., Заец А. В., Ткач Ю. Э., 2007

оценки, которая разрешит более точно выбрать вычислительные ресурсы, на которые будет выполняться задача.

ВВЕДЕНИЕ

При построении распределенных компьютерных систем возникает проблема эффективного управления

ресурсами системи. Далі запропоновано спосіб управління ресурсами при допомозі штучних нейронних мереж.

Метою роботи було створення ефективного методу управління навантаженням в розподілених комп'ютерних системах.

1 РАСПРЕДЕЛЕННЫЕ КОМПЬЮТЕРНЫЕ СИСТЕМЫ И ЗАДАЧИ БОЛЬШОЙ РАЗМЕРНОСТИ

Створення розподілених комп'ютерних систем дозволяє створювати високопродуктивні системи, ефективність яких порівнянна зі спеціалізованими суперкомп'ютерами, що дає можливість вирішувати задачі великої розмірності з великими об'ємами вхідних даних. Існують декілька методів створення розподілених комп'ютерних систем паралельних обчислень (РКСПВ): створення Grid-систем, створення кластерів, а також об'єднання обчислювальних вузлів, які тимчасово не використовуються господарями. Останній підхід вважається більш ефективним, оскільки не вимагає особливих матеріальних витрат і не вимагає повного відчуження ресурсів обчислювального вузла від власника, як це відбувається при створенні кластерів.

При останньому підході РКСПВ є об'єднанням певної кількості гетерогенних обчислювальних вузлів (ВУ), об'єднаних між собою каналами зв'язу, а також одним головним обчислювальним вузлом, який виконує функції управління рештими обчислювальними вузлами системи.

Для вирішення задачі великої розмірності необхідно вибрати спосіб її рішення. Найбільш ефективним вважається спосіб декомпозиції основної задачі на певну кількість атомарних завдань, кожне з яких є певним математичним алгоритмом, який можна реалізувати і виконувати паралельно.

Кожне атомарне завдання має свої характеристики, вхідними і вихідними даними.

2 ПРИМЕНЕНИЕ ИСКУССТВЕННЫХ НЕЙРОННЫХ СЕТЕЙ В ЗАДАЧЕ УПРАВЛЕНИЯ РЕСУРСАМИ

Для вирішення кожного атомарного завдання необхідно вибрати з усієї множини ВУ один або декілька ВУ, на яких виконання завдання займе мінімальний час.

Існують декілька способів управління завданнями в розподілених комп'ютерних системах: використання економічної моделі, попереднє планування, використання першого вільного ресурса. Усі ці способи є неефективними, по-

скільки завдання оцінювання покладається на користувача, або оцінювання не проводиться взагалі, що не дозволяє ефективно керувати навантаженням системи і знижує її ефективність.

Процес вибору обчислювальних вузлів було зведено до рішення задачі оцінювання атомарного завдання за критерієм необхідних ресурсів, а також оцінювання ВУ за критерієм надаваних ресурсів.

Пропонується використовувати штучні нейронні мережі для рішення задачі оцінювання атомарних завдань за критерієм необхідних ресурсів, для оцінювання обчислювальних вузлів за критерієм надаваних ресурсів також було вирішено застосовувати штучні нейронні мережі.

Для рішення задачі оцінювання атомарних завдань за критерієм необхідних ресурсів всі завдання, розв'язані РКСПВ, були розбиті на певну кількість класів, в кожному з яких були виділені еталонні алгоритми.

Для оцінювання кількості ресурсів були виділені наступні параметри завдань, які найбільш сильно впливають на кількість необхідних ресурсів: розмірність вхідного і вихідного вектора завдання, порядок вхідних і вихідних даних, пріоритет завдання, клас завдання. Значення цих параметрів подаються на вхід нейронної мережі. Оцінювання проводиться за 4 параметрами: коефіцієнт потужності процесора, коефіцієнт оперативної пам'яті, коефіцієнт каналу зв'язу, коефіцієнт надійності. Значення цих параметрів отримуємо на виході нейронної мережі.

Для кожного еталонного алгоритму було визначено кількість ресурсів необхідних для виконання за мінімальний час. На основі аналізу потреби в ресурсах еталонних алгоритмів формуються навчальні вибірки для нейронної мережі підсистеми оцінювання завдань.

Схема нейронної мережі оцінювання підзадач представлена на рис. 1.

В результаті навчання була отримана нейронна мережа, що дозволяє оцінювати кількість ресурсів необхідних для рішення кожного атомарного завдання в залежності від значення параметрів завдання.

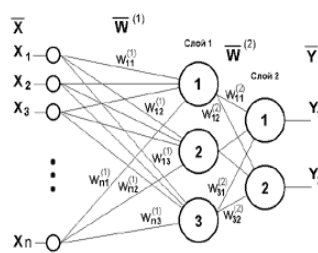


Рисунок 1 – Схема нейронної мережі оцінювання підзадач

Оценивание ресурсов предоставляемых вычислительными узлами системы проводится по тем же 4 параметрам, по которым проводится оценка ресурсов необходимых для решения задачи. Оценивание проводится с помощью искусственных нейронных сетей, в результате чего каждый вычислительный узел получает коэффициент соответствия данному атомарному заданию. Диапазон коэффициентов от 0 до 1.

После оценивания каждого атомарного задания и получения коэффициентов соответствия вычислительных узлов производится выбор минимального количества вычислительных узлов суммарный коэффициент соответствия, которых наиболее близок к 1. После чего задание разбивается на определенное количество частей, которые рассылаются на выполнение на вычислительные узлы.

ЗАКЛЮЧЕНИЕ

С целью повышения эффективности использования вычислительных ресурсов в распределенных компьютерных системах в работе решена актуальная задача управления загруженностью вычислительных узлов при помощи нейронных сетей.

Применение нейронных сетей позволяет избежать сложной процедуры оценивания необходимого количества ресурсов алгоритмическими способами и макси-

мально уменьшить время, необходимое для оценивания задач.

Используя такой способ оценивания, представляется возможным получить оценку количества ресурсов, необходимых для выполнения заданий каждого класса, основываясь только на их описании и анализе их данных.

Научная новизна работы заключается в том, что впервые предложен метод оценивания заданий по критерию необходимых ресурсов, что позволяет избежать сложной процедуры оценивания ресурсоемкости.

Надійшла 2.03.07

Розробка диспетчера розподіленої комп'ютерної системи на основі нейронних мереж перцептронного типу дозволила спростити процес управління ресурсами розподіленої комп'ютерної системи завдяки запропонованій системі оцінювання задач по критерію необхідних ресурсів. Використання нейронних мереж дозволяє скоротити час оцінювання і підвищити точність оцінки, що дозволить більш точно вибрати обчислювальні ресурси, на яких буде виконуватися задача.

Development of the manager of the distributed computer system on a basis of perceptron-like neural networks allowed to simplify managerial process of resources of the distributed computer system due to the offered system of estimating tasks by criterion of necessary resources. Using neural networks allows to reduce time and to increase accuracy of a rating which will allow to choose more precisely computing resources on which tasks will be solved.

УДК [007:572. 788]. 001.57

А. В. Савельев

ОБЗОР ИЗОБРЕТЕНИЙ В ОБЛАСТИ НЕЙРОКИБЕРНЕТИКИ И НЕЙРОМОДЕЛИРОВАНИЯ¹

В статье кратко представлены результаты некоторых работ по направлениям нейрокибернетики и нейрокомпьютеров, проводимых автором с 1977 г., с позиций воплощения их в ряд изобретений, защищенных авторскими свидетельствами и патентами. К этим направлениям относятся: исследование волнового проведения по дендритам, дендропроцессоры; динамическое моделирование нервной системы; изучение циклов в нервной системе и реверберационных моделей; исследование посттетанической потенциации и механизмов памяти; исследование хаотических режимов в нейронах и нейросетях; изучение нейронной логики; исследование механизмов кодирования информации в нервной системе; моделирование явлений пессимума и трансформации ритма; разработка и применение нейросетевой концепции нейрона.

ВВЕДЕНИЕ

Весьма отрадным событием последнего времени можно считать появление аналитических обзоров изобретений в области нейрокибернетики таких, например, как обзор г. С. А. Субботина, опубликованный в сборнике докладов на IV Всероссийской НТК «Нейроинформатика-2002», т. 1, стр. 48–54. Патентные исследования в области отечественной нейрокибернетики очень нужны и чрезвычайно полезны, поскольку открывают доступ научному сообществу к ценнейшей кладовой достижений нейрокибернетики, не только

1. Работа выполнена при финансовой поддержке РФФИ (грант № 04-06-80460) и РГНФ (грант № 04-03-00066а).

стоящей особняком в мире научной информации и имеющей определенную специфику изложения, но и, по большей части, не востребованную современниками. Скепсис по поводу ценности идей, заключенных в изобретениях, по причине этой невостребованности вряд ли уместен, поскольку достаточно вспомнить всем известную историю, ставшую легендой, о японском миллионере, сделавшим свое состояние на разделах полезных советов из российских научно-популярных журналов или попытку японцев заключить договор о покупке всех отказных заявок на изобретения, из числа поступающих в Федеральный Институт Промышленной Собственности (ФИПС).

Проведение подобных исследований и составление аналитических обзоров является очень трудоемким делом, ввиду разнородности и разнонаправленности идей, методов и подходов различных авторов, не полного раскрытия «ноу-хау» с целью защиты интеллектуальной собственности, специфической формы изложения в материалах заявок по стандартам патентования, трудности поиска в патентной литературе и отсутствия этой литературы в последнее время практически во всех городах, за исключением Москвы.

В связи с этим настоящий обзор имеет цель, в том числе, помочь в проведении патентных исследований, поскольку автору в отношении своей работы это сделать проще с максимальной точностью. Изобретения можно классифицировать по группам научных направлений исследований, проводимых нами в Уфимском государственном авиационном техническом университете (УГАТУ) в течение 30 лет с 1977 г. по настоящее время. Ввиду ограниченного объема публикации, просьба сравнение с мировым уровнем читать в соответствующих опубликованных описаниях изобретений, официально обладающих мировым уровнем новизны, и публикациях автора из перечня ссылок.

Существующие нейроинформационные технологии, нейросети и нейрокомпьютеры (hard, soft) во всех случаях, когда с большей эффективностью, когда с меньшей, решают те же задачи, которые решают непрерывные автоматные модели и также не решают задач, недоступных последним. В конечном счете, нейросетевое моделирование с использованием в качестве элементов сетей формальных нейронов Мак-Каллока – Питса сводится в вычислительном отношении во всем своем кажущемся разнообразии к решению все той же задачи наименьших квадратов. В области обработки изображений, хотя и был получен определенный прогресс, тем не менее, результаты его несравнимы с биологическими объектами ни по скорости обработки, ни по удельным затратам вычислительных ресурсов, ни по качеству работы с изображениями. Стоит ли говорить о множестве неформализуемых и трудно формализуемых задач, которые вообще остаются за предела-

ми досягаемости конечно-автоматных и нейросетевых моделей.

Достаточно взглянуть на электронные микрофотографии, чтобы убедиться, насколько наши представления о нейроне как об элементе нейросетей или о самих нейросетях не имеют практически ничего общего с реальным биологическим объектом. Часто разработчики нейрокомпьютеров и технические специалисты, пытающиеся применять нейросетевые концепции для решения специальных задач, априорно принимают как постулат давно устаревшую схему нейрона как элемента Мак-Каллока – Питса почти 60-летней давности. Такая скудная функциональность элементов нейросетей заведомо закладывает жесткие ограничения в решении специальных задач, в результате чего, как правило, дальше довольно примитивных эвристик дело не идет. Справедливости ради необходимо отметить, что в последнее время появились тенденции поворота к обращению внимания технических специалистов на природу самих биологических объектов, правда, это относится пока, в основном, к западным школам. Однако, эти попытки находятся только в начальной стадии своего развития, поскольку захватывают пока еще недостаточно глубокие парадигмальные слои.

1 ВОЛНОВОЕ ПРОВЕДЕНИЕ ПО ДЕНДРИТУ. ДЕНДРОПРОЦЕССОРЫ

Нами проводилось исследование волновых явлений распространения пассивных локальных постсинаптических потенциалов (ЛПСП) в объемной структуре дендрита с реальной конфигурацией [1], которая реконструировалась по электронным микрофотографиям продольных срезов дендритных волокон. Был разработан ряд оригинальных, нашедших отражение в изобретениях [2–8], моделирующих алгоритмов, основанных на групповом анализе уравнений проведения по объемному проводнику сложной конфигурации:

$$\begin{cases} v''_{xx} = F_1[v'_x, v'_t, r(x), C(x), L(x)]; \\ i''_{xx} = F_2[i'_x, i'_t, r(x), C(x), L(x)]. \end{cases}$$

Поскольку такая задача в математическом, так и в вычислительно-математическом отношении представляется чрезвычайно трудной, нами был разработан ряд методов, нашедших отражение в изобретениях [2–8]. Используя группы преобразований Ли – Беклунда и теоремы Э. Нетер, Н. Х. Ибрагимова [9], а также ряд теорем о продолжении преобразований [10], были построены соответствующие законы сохранения, что использовалось для синтеза моделирующих схем по а. с. № № 1501101, 1585811, 1815658. Были определены условия возникновения дифракции и интерференции волн ЛПСП в дендритных структурах и было

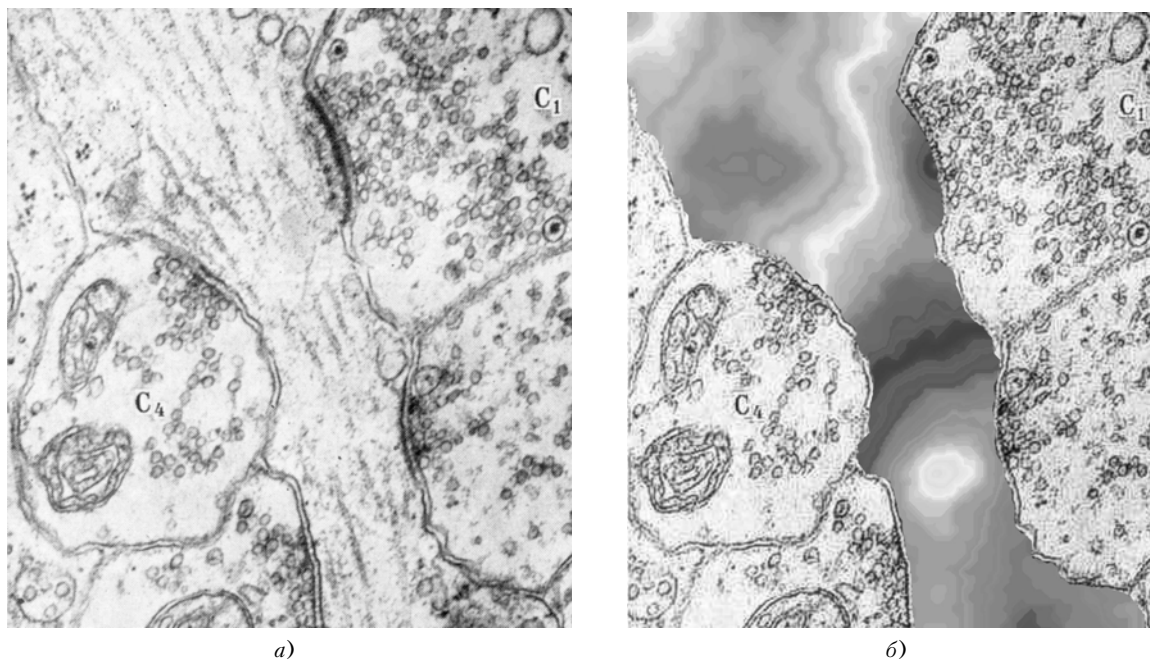


Рисунок 1:

а – электронно-микроскопическая фотография продольного среза дендрита. Ретикулярная формация агамы 1:33000. С1, С4 – синапсы (1,5–3 мкм). С1 – синапс I типа по Грэй, С4 – синапс II типа по Грэй (Е. G. Gray). Фото Н. С. Косицина; б – реконструкция интерференционной картины потенциально-токовых артефактов на том же образце

выполнено их моделирование (рис. 1). Воспроизведение волновых процессов, протекающих в дендритах реальных нейронов показано на рис. 1, б.

Проводился также бифуркационный анализ уравнений проведения по пространственным параметрам конфигурации, в том числе, с использованием методов проекций [11] и теоремы о центральном многообразии [12] для сведения исходной бифуркационной задачи к соответствующим задачам для обыкновенных эволюционных дифференциальных уравнений. На основании такого анализа была показана возможность возникновения во внутридендритном интерференционном поле потенциально-токовых артефактов, в том числе, рождение циклов. Исходя из принципа стремления системы к минимальным энергетическим состояниям, устранение этих артефактов может осуществляться только при вариации пространственного параметра, описывающего конфигурацию дендритного ствола. Таким образом, нами еще в 1983–1984 г. г. была предсказана возможность механических микроперемещений дендритных ветвей нейрона в связи с его функциональной активностью, что может отражаться в моделирующих схемах по а. с. 1501101 и 1585811 при соответствующей настройке параметров, а также в моделирующей системе по а. с. 1815658, отражающей компенсацию потенциально-токовых артефактов при механическом перемещении синаптических контактов по поверхности дендрита. Таким образом, была доказана принципиальная *электро-механическая* природа ней-

рона как элемента нейросетей, что впоследствии в 1987 г. было подтверждено экспериментально японскими учеными, которые в остром эксперименте вводили световод в трепанированное отверстие в черепе и, подсвечивая неокортекс, производили микровидеосъемку, подтвердившую наше теоретически сделанное открытие. Далее, нами рассматривались более сложные модели взаимодействия волнового поля с внутридендритными спайками (а. с. 1394975), предлагающие, в том числе, разрешение проблемы эффективности дистальных синапсов (а. с. 1439631). Кроме того, воспроизводились некоторые нелинейные явления (а. с. 1561076), в том числе, потенциалозависимость гиперполяризации дендритной мембраны от активности тормозящих синапсов и влияние этого на интерференционную картину (пат. 2028669). Каждая моделирующая схема представляла собой, фактически, специализированный дендропроектор с набором частично-эвристических процедур, решающих определенную модифицированную систему уравнений в частных производных и бифуркационные явления, возникающие в ней при перенастройке параметров.

2 ДИНАМИЧЕСКИЕ МОДЕЛИ НЕРВНОЙ СИСТЕМЫ

Методология точных наук открыла (по словам И. М. Сеченова) и продолжает открывать новый

взгляд на природу и функционирование биообъектов, дополняющий и расширяющий классические биологические подходы. Нами была поставлена и решена в классической постановке задача идентификации динамических свойств моносинаптического возбуждения и торможения различных типов живых нейронов млекопитающих, в том числе человека. Необходимо отметить, что не только такая задача ранее не решалась, но также отсутствовали какие-либо сравнительные идентификационные исследования моносинаптических вызванных потенциалов одного нейрона и, тем более, различных нейронов. В классических трудах Дж. Экклза [13], В. Катца [14], К. Учиозо [15], как правило, фиксировали воздействия на нейрон только полисинаптического характера. В результате исследований, проведенных автором частично в Ростовском институте нейрокибернетики им. А. Б. Когана в 1985–1987 гг., были определены динамические характеристики синаптической передачи α -мотонейронов поясничного отдела спинного мозга кошки, нейронов Пуркиньи мозжечка крысы и пирамидных нейронов неокортекса крысы и человека. Были определены области вариаций структурно-коррелированных динамических параметров синаптической передачи, что позволило проводить сравнительные исследования. На основе этого нами было предложено существенно модифицировать модель формального нейрона Мак-Каллока – Питса, используемую повсеместно для организации нейросетей и в нейровычислениях. Предложенная модель [16] (а. с. 1329449) не только воспроизводила реальные динамические свойства синаптической передачи, но и вносила обнаруженную нами существенную асимметрию в динамических свойствах возбуждения и торможения, присутствующую в любом реальном нейроне. Это могло бы дать возможность при небольшом увеличении аппаратных затрат значительно повысить интеллекту-

альные свойства искусственных нейронных сетей (ИНС).

Нами была осуществлена динамическая композиция популяции α -мотонейронов на основании однородности динамических свойств эффекторных приводов по трем группам различного уровня быстродействия и депрессирования. Таким образом, в устройстве-нейросети [17] по а. с. 1515938 реализован новый принцип построения нейросетей, заключающийся в моделировании ансамбля нейронов со сходными динамическими свойствами одним звеном, охватывающим область динамических реакций ансамбля, идентификация структуры и параметров которого проводится на основании ответов реальной нервной ткани. Соответственно этому нейросеть клеток Реншоу в устройстве [17] (здесь и далее термин «устройство» условен, т. к. может быть реализовано в виде алгоритма, концептуальной схемы и т. д.) также редуцируется к сети 3-го порядка с тремя обобщенными группами. Такой подход может быть одним из перспективных направлений снижения сложности искусственных нейронных сетей путем предложенного понижения их размерности. Использование его в устройстве управления эффекторами может значительно повысить гибкость, плавность и точность движений робота-манипулятора путем приближения логики его работы к биологическому прототипу при приемлемой затрате вычислительных ресурсов. Моделирование клеток Реншоу само по себе представляет значительный интерес в силу их уникально акцентированных свойств, характерных для биоткани вообще. В а. с. 1292494 [18] реализовано устройство управления активностью мотонейрона, осуществляемое клетками Реншоу, образующими т. н. гибкие обратные связи (ОС), изменяющие свой знак в зависимости от положительную и наоборот, дискретно или градуаль-

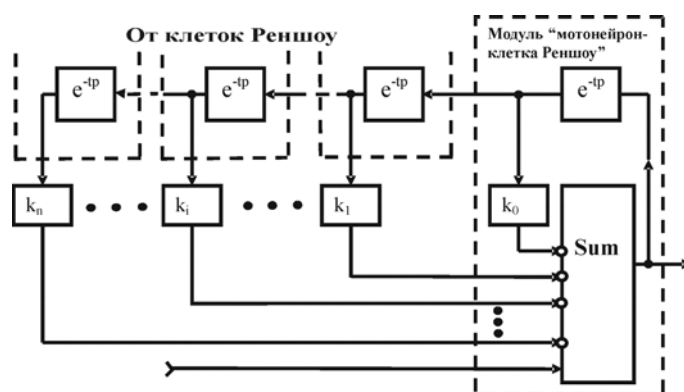


Рисунок 2 – Элементарная процессорная схема нейросетевого процессора комплексов «мотонейрон – клетка Реншоу» по а. с. 1292494 [18]:

k_i – весовые коэффициенты; e^{-tp} – звенья запаздывания; пунктиром обозначен собственный элементарный комплекс «мотонейрон – клетка Реншоу»

и на системном уровнях. Нами был предложен ряд решений улучшающих простым образом, удобным для реализации, формальный нейрон Мак-Каллока – Питса. Используемые в ИНС, они допускают лишь настройку весовых коэффициентов синаптических контактов, что ограничивает функциональные возможности таких ИНС. В а. с. 1238583 и 1406613 [23–24] нами реализована возможность управления порогом нейрона и алгоритмы такого управления, включая самоорганизующуюся настройку. Алгоритмы управления и самоуправления порогом были получены на основании изучения посттетанической потенциации при длительных воздействиях, приводящих к высокочастотным разрядам нейронов. Использование таких моделей как элементов ИНС значительно увеличивает их функциональные возможности, расширяя число решаемых вычислительных задач с одновременным снижением требуемых ресурсов (уменьшение количества нейронов и упрощение связей между ними), а также позволяет улучшить и ускорить сходимость вычислительных алгоритмов за счет использования усиленных адаптивных свойств предложенных схем. Предложенные алгоритмы управления индивидуальными порогами синапсов в а. с. 1340421 и 1477140 [25–26] позволяют организовать процессы обработки информации в нейроне как в существенно нелинейном элементе, обладающем нелинейными весовыми синаптическими коэффициентами, а также реализовать кодирование рисунка активированных синапсов в индивидуальные состояния самого нейрона с сохранением его во времени и в состоянии самих синапсов при учете их внутридендритных взаимодействий (а. с. 1477140).

Механизмы функциональной памяти, основанной на ассоциативных явлениях изучались также при моделировании условного рефлекса. В устройстве по а. с. 1424575 [27] реализован поведенческий нейровычислитель ассоциаций клеточного уровня, позволяющий воспроизводить ассоциирование поступающих стимулов при определенном их качестве и условиях поступления. Обучение устройства осуществляется подачей пар стимулов и выражается в эффекте повышения порога возбуждения. Предложенная схема позволяет повысить интеллектуальные свойства элементов нейросети, расширить их функциональность и организовать ассоциативную память в интегральном виде для всего нейрона в целом. В устройстве по пат. 2093889 [28] информационная обработка ассоциаций происходит на более тонком уровне с элементами их семантического анализа с учетом значения не только межимпульсного кода, но также и межстимульного поведенческого кода. В этом случае предложенный процессор ассоциаций может также выступать в роли семантического процессора.

5 ХАОТИЧЕСКИЕ РЕЖИМЫ В НЕЙРОНАХ И НЕЙРОСЕТЯХ

Согласно теории нейронного ансамбля А. Б. Когана [29] и вероятностного характера связей между нейронами в ансамбле до обучения, робастный характер этих связей понимается относительно целенаправленной активности ансамбля. Учитывая нашу концепцию автофрактальности нейрона, т. е. возможности построения самого нейрона в виде нейросети [30], мы распространили эти положения, весьма справедливые для нейронного ансамбля, и на нервную клетку. Устройство по а. с. 1564654, [31] позволяет реализовать хаотическую логику работы нейрона на макроуровне. Хаотические режимы может иметь результат функционирования нервной клетки (как нейросети) – частота спайковых разрядов, проявляющиеся при определенных качественных состояниях нейрона, распознаваемых предложенной схемой. Устройство может служить нейроном-детектором в задачах распознавания образов, принятия решений, управления, кодирования информации, а также в биомедицинской кибернетике в качестве электростимулятора или в гибридных технобиологических системах. С теоретической точки зрения предложенное решение может выражать новый взгляд в понимании механизмов кодирования информации в нервной системе.

В устройстве по а. с. 1585810, [32] реализован механизм продуцирования эндогенных спонтанных разрядов, присутствие которого характерно для многих нейронов нервной системы, в особенности, мотонейронов, обеспечивающих поддержание некоторого тонуса иннервируемой мышечной системы. Участие таких нейронов в нейросетях позволяет организовать фоновое возбуждение молчащей нейросети в отсутствие внешних стимулов, что дает возможность повысить ее линейризуемость, осуществить поддержание активированного состояния оперативной памяти, а также снизить пороги и гистерезис эффекторов.

6 НЕЙРОННАЯ ЛОГИКА

Свойства нейрона как логического элемента чрезвычайно многообразны. Логические свойства нейрона могут проявляться в избирательном реагировании его, как, например, в конвергентных нейронах спинальных ганглиев, на активизацию синапсов, связанных с низкочастотными или высокочастотными рецепторами, спайковыми ответами обычной или увеличенной длительности (а. с. 1645975, [33]). Или же в мотонейронах, реагирующих быстрыми Na^+ -спайками, соответствующих быстрым движениям эффекторов, или медленными Ca^{2+} -спайками, обеспечивающими медленные движения эффекторов (а. с. 1324044, [34]). И в том, и в другом случаях нейрон обнаруживает более бога-

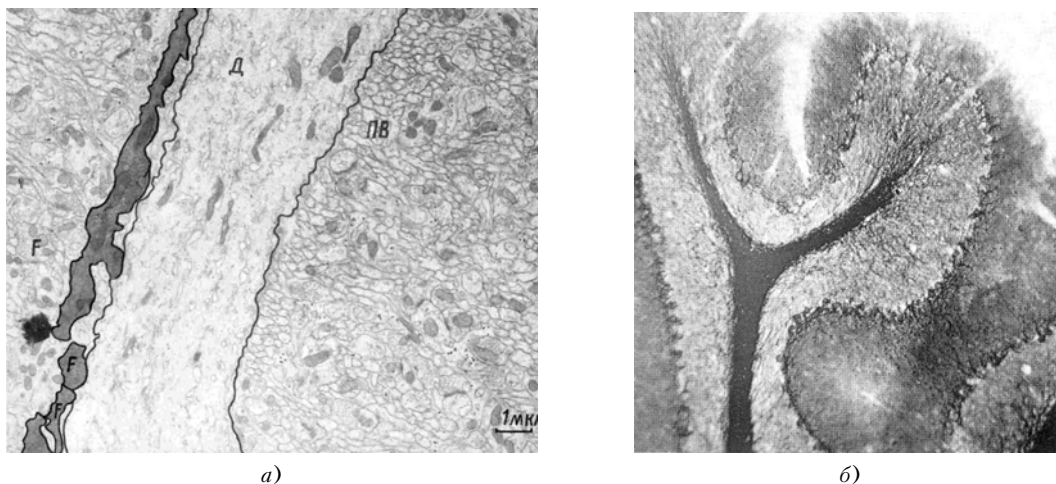


Рисунок 4:

а – взаимосвязь между лиановидным волокном (темные участки) и дендритом (Д) клетки Пуркинье. Д – дендрит клетки Пуркинье мозжечка кошки. Параллельно ему следует лиановидное волокно (темное) на большом протяжении дендрита, осуществляя с ним «протяженные» синаптические контакты. Буквой F отмечены единичные синапсы F-типа. ПВ – параллельные волокна гранулярных клеток. Фото К. Uchizono;

б – строение мозжечка млекопитающего (кошка, метод управляемой импрегнации В. В. Коротченко, микрофото). На снимке показана цито- и миеоархитектоника извилин коры мозжечка: хорошо видны слои трех основных типов нейронов мозжечка – молекулярный слой, слой клеток Пуркинье (темный, тонкий), слой гранулярных клеток

тую логику, кроме традиционно учитываемой тривиально пороговой, а следовательно, обладает более богатой и гибкой функциональностью, что позволяет значительно сокращать размерность нейросети, требуемой для решения поставленных задач.

Свойства нейрона как нетривиального логического элемента могут проявляться также при переключении его на различные виды экзоцитоза – использующего холинергический тип и нехолинергические типы медиатора. При этом такая логика работает «по выходу», т. е. качественное состояние нейрона определяется характером потока его же выходной информации – определенным видом спайковой последовательности, как в вычислителе экзоцитоза по а. с. 1645973, [35], являющимся анализатором степени «пачечности» своих разрядов. Устройство может использоваться также в качестве анализатора импульсных кодовых последовательностей.

В то же время, количественные показатели пачечности могут служить логическими состояниями нейрона, причем, с градацией, превышающей обычную двузначную логику и могут реализовывать многозначную, бесконечнозначную, непрерывную логику, логику «неисключенного четвертого» и т. д. Такое управление характером логических свойств нейрона может осуществляться, например, модуляцией со стороны лиановидного волокна, конвергирующего с нейроном Пуркинье мозжечка (а. с. 1497626, [36]) и определяется топологией их контактирования (рис. 4).

Логические свойства нейрона проявляются также в эмоциональных реакциях на клеточном уровне, например, в нейронах неокортекса при действии на их рецепторы диффузионных процессов распространения моноаминов, выделяемых нейронами эмоциональных центров. В устройстве – процессоре эмоций по а. с. 1464181, [37], как и комплексе мотонейрон – клетка Реншоу по а. с. 1292494, [18], может происходить дискретная или непрерывная трансформация обратной связи с положительной на отрицательную, и наоборот. При этом процессор эмоций по а. с. 1464181 реализует реальные алгоритмы динамики изменения состояний популяции нейронов (нейросети) при действии на нее моноаминов, выделяемых при отрицательных, так и положительных эмоциях. Кроме того, здесь можно выделить особый тип логики – логику с учетом предистории.

В [31] реализуется совершенно новый тип логики для технических устройств, предложенный авторами. Бинарные отношения могут кодироваться аналогично явлениям, наблюдаемым во многих нейронах отделов нервной системы, если учитывать реальные процессы, происходящие в их функционирующих аксонах, степень детерминированности или хаотичности сигнала.

Логические функции в нервной системе может выполнять также внеклеточное электромагнитное поле. В устройстве по а. с. 1439632, [38] реализована нейристорная сеть, моделирующая нервный пучок. Дополнительная логика задается взаимной топологией отдельных аксонов в нервном пучке, которая определяет

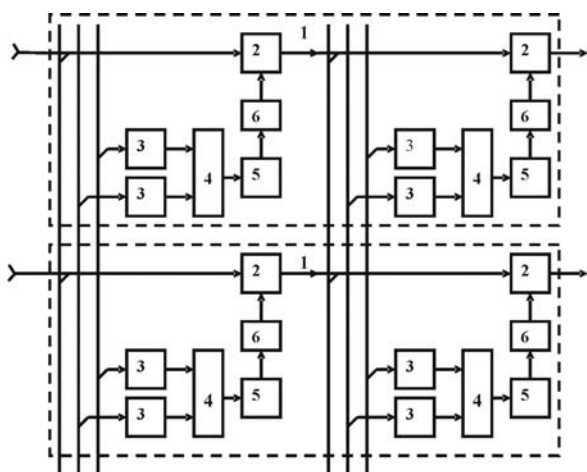


Рисунок 5:

Нейросетевой процессор нервного пучка по а. с. 1439632, [38] содержит M нейристоров 1, состоящих из N последовательно соединенных нейристорных ячеек 2, выходы последних ячеек в нейристорах 1 являются выходами нервного пучка; масштабирующие блоки 3, сумматоры 4, блоки 5 дифференцирования, блоки 6 запаздывания; входы первых нейристорных ячеек являются входами устройства

характер и степень влияния их друг на друга в момент прохождения спайков (рис. 5).

Это влияние может выражаться в динамическом синхронном понижении или повышении порогов соответствующих (соответственно расположению волокон и рисунку их активации) отдельных нейристорных ячеек нейронной сети пучка. При этом поле, связывая отдельные волокна, и образует эту аксонную нейронную сеть, причем, связи носят логический характер,

реализуя собой динамическую синхронную непрерывнозначную логику.

7 ИЗУЧЕНИЕ МЕХАНИЗМОВ КОДИРОВАНИЯ ИНФОРМАЦИИ В НЕРВНОЙ СИСТЕМЕ

Одно из направлений – исследование механизмов кодирования информации формой спайка (а. с. 1324044, 1436720 [39]) или его длительностью (а. с. 1645975). Устройства используются в цифровом варианте в качестве детекторов и модуляторов (см. также п. 6). В устройстве по а. с. 1306368, [40] реализован нейросетевой соматический процессор, генерирующий сложный спайк, вклад в который осуществляется множеством компонент – локальных спайков от триггерных зон соматической активной мембраны. Устройство предназначено для воспроизведения бионической передачи множества данных по одному каналу без какого-либо частотного, временного или другого разделения.

Устройство по а. с. 1463030, [41] является нейропроцессором средовых воздействий, кодирующим информацию о воздействиях на нейрон (или нейросеть) в виде гуморальных влияний, механических воздействий на аксон (сгиб, нажатие), что имеет место, например, при иннервации движущихся конечностей и т. д. Эта информация кодируется в виде фазы задержанной постспайковой деполяризации, возникающей на фоне рефрактерного периода. При этом информация передается не только величиной времени задержки, но также выраженностью и формой волны деполяризации, что значительно повышает информацион-

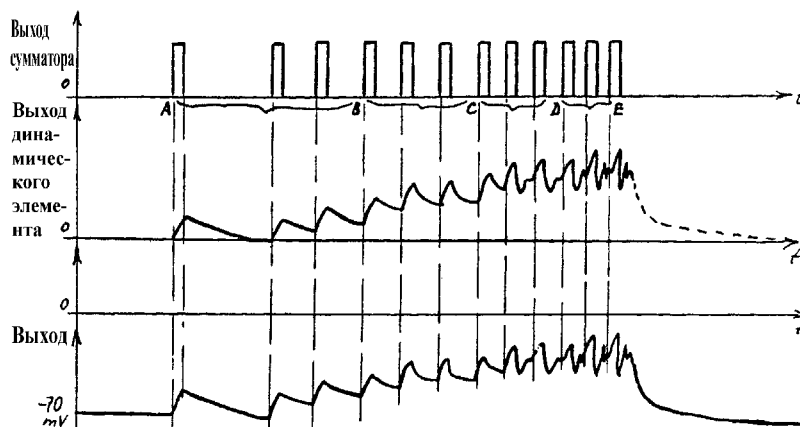


Рисунок 6:

На выходе управляемого динамического звена наблюдаются осцилляции мембранного потенциала, вызванные интенсивной деполяризацией мембраны (участки С-Д, D-E), аналогично реальному комиссуральному мотонейрону. Во всех случаях при любых уровнях возбуждения сигнал на выходе устройства остается градуальным аналогично реальному комиссуральному мотонейрону и не представляет собой обычную спайковую последовательность, однако, тем не менее, способен передавать информацию о различных своих состояниях и уровнях возбуждения с большей информативностью, чем при спайковой передаче, т. к. может передавать информацию не только в точке возбуждения, но и во всем диапазоне функционирования

ную емкость канала связи с применением данного устройства.

Другое направление – кодирование информации в нейросетях на подпороговых уровнях без спайковой активности нейронов до сих пор вообще не рассматривается в нейрокомпьютинге.

Тем не менее, такой способ кодирования и обработки информации может дать параллельную и независимо существующую нейросеть, т. е. снизить затрачиваемые ресурсы и сложность вдвое (а.с. 1672482, [42]). В устройстве по а. с. 1807503, [43] – модели комиссурального мотонейрона высших видов воспроизводится динамика экзоцитоза медиатора под действием градуальных активных ответов нейрона без генерации ими потенциалов действия (рис. 6). Устройство используется в качестве передатчика информации в нейропроцессорных сетях с повышенной информационной плотностью за счет возможности передачи информации не только в точке возбуждения, но и во всем динамическом диапазоне функционирования устройства.

Одно из перспективных направлений – кодирование информации хаотическими сигналами (а. с. 1564654, см. п. 5) позволяет в значительной степени повысить как информационную емкость канала передачи, так и помехозащищенность, поскольку использует шум в качестве источника полезного сигнала. Все эти устройства могут выполнять также роль нейроэлектростимуляторов, оптимально согласующих параметры воздействия с внутренней средой и текущей функциональной активностью биообъектов.

8 ЯВЛЕНИЕ ПЕССИМУМА И ТРАНСФОРМАЦИИ РИТМА

На основе исследования сверхмедленных механизмов волновой памяти, проводимых нами в УГАТУ с 1984 г. и в Ростовском НИИ нейрокибернетики им. А. Б. Когана в 1985–1987 г.г. была разработана теория памяти, основанная не на возбуждении (что являлось основой практически всех известных теорий) или контрастировании процессов возбуждения и торможения, а на явлении пессимума, т. е. на абсолютном торможении безотносительно к возбуждению. Предложенная теория позволяла, в том числе, объяснить природу двух видов торможения по А. А. Ухтомскому: высший тип торможения (пост-возбуждение) является переходом в долговременную память процессов, протекающих в оперативной памяти (возбуждение). Развитие этих принципов используется в серии моделирующих схем по а. с. 1387023, 1691858, пат. 2024059 [44–46]. Кроме того, реакция торможения нейрона и нейросети, сама по себе, имеет гораздо более сложную структуру и заключается не только в соответствующих вариациях порога возбуждения, но выражается также в изменении формы, длительности спайков, величине зоны ги-

перполяризации, флуктуациях потенциала покоя и т. д. При этом зона следовой гиперполяризации также может участвовать в хранении следов памяти, что наиболее выражено в т. н. «молчащих» нейронах ц. н. с. (а. с. 1436720, [39]). Явления памяти и кодирования информации обнаруживаются и в пессимальных реакциях трансформации ритма разрядов. Благодаря этим процессам спайковая последовательность помимо аналоговых параметров, отражающих состояние нейрона, может представлять собой последовательный код, что реализовано в устройстве по а. с. 1458879 [47]. В связи с этим появляется дополнительный канал передачи данных.

9 НЕЙРОСЕТЕВАЯ КОНЦЕПЦИЯ НЕЙРОНА

Кроме отмеченных изобретений по а. с. 1513938 – представляющему собой модель нейросети и а. с. 1292494 – включающему принцип построения нейросетей из комплексов мотонейрон – клетка Реншоу, некоторые отмеченные устройства для моделирования нейронов являются фактически нейросетями, согласно разработанному нами ряду положений [48]. Так устройство по а. с. 1306368, [40] представляет собой аксо-соматический нейросетевой процессор; устройство для моделирования нейронного пучка по а. с. 1439632 – многослойная нейросеть в явном виде (разве что, только это не отражено в названии). Реверберационные модели (а. с. 1425731, 1642485) являются минимальными нейросетями из двух (как минимум) нейронов, а устройство для морфодинамического моделирования нейрона (1815658) содержит N однослойных нейросетей из блоков моделирования морфодинамики (где N – число дендритов), как и устройство по а. с. 1561076. В виде нейросети можно рассматривать структуру дендрона, где отдельный дендрит представляет собой подобие нейрона (или является субклеточной единицей [49]), как в а. с. 1394975, 1439631, 14771140, 1501101, 1585811, пат. 2028669. Нейросеть из 3-х нейронов содержит устройство для моделирования нейрона неокортекса [50] по а. с. 1464181 и а. с. 1645975 – логический нейроподобный элемент, а. с. 1691858, пат. 2024059; двухнейронные соединения используются также в устройствах по а. с. 1324044, 1463030, 1585810, 1672482.

ЗАКЛЮЧЕНИЕ

Считаем, что рассмотрение разработанных нами концепций моделирования нейронов и нейросетей может составить фундаментальные основания нейрокомпьютинга наряду с существующими на сегодняшний день подходами [49, 51].

Учитывая, что основные усилия в нейрокомпьютеринге сосредоточены на нейросетях, т. е. на принципах и алгоритмах соединений, по сути дела, простейших элементов, нам хотелось бы показать какое большое значение может иметь характер представления самого нейрона, более детальное рассмотрение которого может повлечь за собой воспроизведение совершенно иных принципов организации и работы нейросетей в целом [52].

На основе этого, в частности, можно сделать выводы, что в центральной нервной системе не всегда соблюдается принцип сложности – элемент (нейрон) может быть и часто бывает сложнее сети (точно также, как вопрос, что сложнее – человек или группа людей?).

Работа также имеет цель стимулировать дальнейшие исследования возможности концептуально-парадигмального пересмотра устоявшихся взглядов на нейрон как на элемент нейросетей и самих нейросетей в нейрокомпьютеринге, что могло бы дать существенное дальнейшее продвижение вперед.

ПЕРЕЧЕНЬ ССЫЛОК

1. Косицин Н. С., Федоренко Г. М., Свинов М. М. Пространственная организация внутриклеточных органелл механо-рецепторных нейронов как показатель их различного функционального состояния // ДАН. – 2002. – № 1(385). – С. 123–125.
2. А. с. № 1501101. Устройство для моделирования нейрона / Савельев А. В., Савельева Н. А., Колесников А. А., Жуков А. Г. – 1989, Бюл. № 30.
3. А. с. № 1585811. Устройство для моделирования нейрона / Жуков А. Г., Колесников А. А., Савельева Н. А., Савельев А. В. – 1990, Бюл. № 30.
4. А. с. № 1815658. Устройство для морфодинамического моделирования нейрона / Жуков А. Г., Лаврова Т. С., Савельев А. В. – 1993, Бюл. № 18.
5. А. с. № 1394975. Устройство для моделирования нейрона / Савельева Н. А., Савельев А. В. – 1988.
6. А. с. № 1439631. Устройство для моделирования нейрона / Савельев А. В. – 1988. Бюл. № 43.
7. А. с. № 1561076. Устройство для моделирования нейрона высших отделов / Газутдинов И. Ф., Лакомкин И. М., Савельев А. В., Сергеев Н. А. – 1990, Бюл. № 16.
8. Пат. № 2028669 РФ. Устройство для моделирования нейрона / Жуков А. Г., Лаврова Т. С., Савельев А. В. – 1995, Бюл. № 4.
9. Ибрагимов Н. Х. Группы преобразований в математической физике. – М.: Наука, 1983. – 278 с.
10. Помаре Ж. Системы уравнений с частными производными и псевдогруппы Ли. – М.: Мир, 1983. – 398 с.
11. Йосс Ж., Джозеф Д. Элементарная теория устойчивости и бифуркаций. – М.: Мир, 1983. – 300 с.
12. Carr J. Applications of Centre Manifold Theory // Appl. Math. Science, NY: Springer-Verlag. Vol. 35. – 1981.
13. Экклз Дж. Физиология синапсов. – М.: Мир, 1966. – 395 с.
14. Katz B. The release of neural transmitter substance. – Liverpool, 1969. – 463 p.
15. Учизоно К. Возбуждение и торможение. Морфология синапсов. – Киев: Наукова думка, 1980. – 213 с.
16. А. с. № 1329449. Устройство для моделирования нейрона / Межецкая Т. А., Савельева Н. А., Савельев А. В., Колесников А. А. – 1987.
17. А. с. № 1515938. Устройство для моделирования системы возвратного торможения мотонейронов клетками Реншоу / Ильясов Б. Г., Кабальнов Ю. С., Савельев А. В., Валиева Н. Э. – 1989.
18. А. с. № 1292494. Устройство для моделирования нейрона / Савельев А. В. – 1987.
19. А. с. N 1642485. Устройство для моделирования нейрона / Жуков А. Г., Колесников А. А., Савельева Н. А., Савельев А. В. – 1991, Бюл. № 14.
20. А. с. N 1425731. Устройство для моделирования нейрона / Колесников А. А., Жуков А. Г., Савельева Н. А., Савельев А. В. – 1988, Бюл. N 35.
21. Бакусов Л. М. Компаратментные модели микромеханизмов усвоения закономерностей нейронными сетями // Материалы международной науч. конф. «Теория информации-2000. От теории к практике». – Уфа, 2000. – С. 191–197;
22. Бакусов Л. М. Моделирование механизмов каудальной самоорганизации в сложных динамических системах. Автореферат докт. дисс. по специальности 05.13.18. – Уфа, 2006. – 42 с.
23. А. с. № 1238583. Устройство для моделирования нейрона / Савельев А. В. – 1986.
24. А. с. № 1406613. Устройство для моделирования нейрона / Савельев А. В. – 1988, Бюл. N 24.
25. А. с. № 1340421. Устройство для моделирования нейрона / Савельев А. В. – 1987.
26. А. с. № 1477140. Устройство для моделирования нейрона / Жуков А. Г., Колесников А. А., Савельева Н. А., Савельев А. В. – 1989.
27. А. с. № 1424575. Устройство для моделирования нейрона / Савельев А. В. – 1988.
28. Пат. № 2093889 РФ. Устройство для моделирования нейрона / Ильясов Б. Г., Лаврова Т. С., Савельев А. В. – 1997, Бюл. N 29.
29. Коган А. Б. Функциональная организация нейронных механизмов мозга. М.: Медицина, 1979. – 224 с.
30. Базарова Д. Р., Демочкина Л. В., Савельев А. В. Новая нейробионическая модель онтогенеза // В сб. тр. МИФИ: «Нейроинформатика-2002». – 2002. – Т. I. – С. 97–106.
31. А. с. № 1564654. Устройство для моделирования нейрона / Газутдинов И. Ф., Лакомкин И. М., Савельев А. В., Сергеев Н. А. – 1990, Бюл. № 18.
32. А. с. № 1585810. Устройство для моделирования мотонейронов / Ильясов Б. Г., Савельев А. В., Колесников А. А. – 1990, Бюл. № 30.
33. А. с. N 1645975. Логический нейроподобный элемент / Ильясов Б. Г., Савельев А. В., Колесников А. А. – 1991, Бюл. № 16.
34. А. с. № 1324044. Устройство для моделирования мотонейрона / Савельева Н. А., Савельев А. В. – 1987, Бюл. № 26.
35. А. с. № 1645973. Устройство для моделирования нейрона / Савельев А. В., Савельева Н. А., Колесников А. А., Жуков А. Г. – 1991, Бюл. № 16.
36. А.с. № 1497626. Устройство для моделирования нейрона Пуркинье / Межецкая Т. А., Савельева Н. А., Савельев А. В., Колесников А. А. – 1989, Бюл. № 18.
37. А. с. № 1464181. Устройство для моделирования нейрона неокортекса / Жуков А. Г., Колесников А. А., Савельева Н. А., Савельев А. В. – 1989, Бюл. № 9.
38. А. с. № 1439632. Устройство для моделирования нервного пучка / Межецкая Т. А., Савельев А. В., Колесников А. А. – 1988, Бюл. № 43.
39. А. с. № 1436720. Устройство для моделирования нейрона / Савельева Н. А., Савельев А. В. – 1988.
40. А. с. № 1306368. Устройство для моделирования нейрона / Межецкая Т. А., Савельев А. В., Колесников А. А. – 1986.
41. А. с. № 1463030. Устройство для моделирования нейрона / Савельев А. В. – 1988.
42. А. с. № 1672482. Устройство для моделирования нейрона / Жуков А. Г., Савельева Н. А., Савельев А. В. – 1991, Бюл. № 31.
43. А. с. № 1807503. Устройство для моделирования комиссурального мотонейрона / Ильясов Б. Г., Савельев А. В., Лаврова Т. С. – 1993, Бюл. № 13.
44. А. с. № 1387023. Устройство для моделирования нейрона / Савельев А. В. – 1988, Бюл. № 13.
45. А. с. № 1691858. Устройство для моделирования нейрона / Лаврова Т. С., Савельев А. В., Жуков А. Г. – 1991, Бюл. N 42.
46. Пат. № 2024059 РФ. Устройство для моделирования нейрона / Жуков А. Г., Савельева Н. А., Савельев А. В., Лаврова Т. С. – 1994, Бюл. № 22.

47. А. с. № 1458879. Устройство для моделирования нейрона / Савельев А. В., Савельева Н. А., Жуков А. Г., Колесников А. А. – 1989, Бюл. № 6.
48. Савельев А. В. Источники вариаций динамических свойств нервной системы на синаптическом уровне в нейрокомпьютинге // Искусственный интеллект. Донецк, НАН Украины. – 2006. – № 4. – С. 323–338.
49. Савельев А. В. На пути к общей теории нейросетей. К вопросу о сложности // Нейрокомпьютеры: разработка и применение. М.: Радиотехника. – 2006. – № 4–5. – С. 4–14.
50. Савельев А. В. К вопросу о принципе сложности нейрокомпьютеров // В сб. докладов V Всерос. НТК: «Нейроинформатика-2003». – М.: МИФИ, 2003. – Ч. I. – С. 40–53.
51. Савельев А. В. Информационная гиперсложность нейронов и нейросетей в новой методологии нейромоделирования // В сб. докладов VIII Всерос. НТК: «Нейроинформатика-2007». – М.: МИФИ, 2007. – Ч. I. – С. 49–61.
52. Савельев А. В. Реализм теории модульной самоорганизации мозжечка // Журнал проблем эволюции открытых систем. Казахстан: Алматы, 2007. – № 1 (9). – С. 118–130.

Надійшла 2.03.07

У статті коротко представлені результати деяких робіт з напрямків нейрокибернетики й нейрокомпьютерів, проведених автором з 1977 р. з позицій втілення їх у ряд винаходів, захищених авторськими посвідченнями

й патентами. До цих напрямків ставляться: дослідження хвильового проведення по дендритах, дендропроектори; динамічне моделювання нервової системи; вивчення циклів у нервовій системі й ревербераційних моделях; дослідження посттетаничної потенціації й механізмів пам'яті; дослідження хаотичних режимів у нейронах і нейросетях; вивчення нейронної логіки; дослідження механізмів кодування інформації в нервовій системі; моделювання явищ пессимума й трансформації ритму; розробка й застосування нейросетової концепції нейрона.

In the article the results of same works on neurocybernetics and neural computers directions was implemented by author since 1977 from items of them embodiment in the inventions series protected by the patents are briefly submitted. They are: the research of wave conduction on dendrites, dendral processors; dynamic simulation of neural system; research of cycles in neural system and reverberationary models; research of post-tetany potentiation and memory mechanisms; research of chaotic modes in neurons and neural networks; learning of neuron logic; research of coding mechanisms in neural system, simulation of pessimum phenomenon and rhythm transformation; development and using of neural network concept of neuron.

ПРОГРЕСИВНІ ІНФОРМАЦІЙНІ ТЕХНОЛОГІЇ

ПРОГРЕССИВНЫЕ ИНФОРМАЦИОННЫЕ ТЕХНОЛОГИИ

PROGRESSIVE INFORMATION TECHNOLOGIES

УДК 004.67

Ю. С. Афонин, В. И. Дубровин, Т. В. Манило

ПРИМЕНЕНИЕ МЕТОДОВ ВЕЙВЛЕТ-АНАЛИЗА И ПРЕОБРАЗОВАНИЯ ФУРЬЕ В ЗАДАЧАХ ДИАГНОСТИКИ ОБЪЕКТОВ ПО ВИБРОАКУСТИЧЕСКОМУ СИГНАЛУ

Представлены методы вейвлет-преобразования (быстрое и пакетное) и преобразование Фурье для обработки виброакустического сигнала в контексте их применения в задачах диагностики объектов. Для каждого метода выявлены характеристики (результаты обработки виброакустических сигналов), связанные с состоянием исследуемого объекта. Выполнено сравнение методов по времени обработки сигналов.

ВВЕДЕНИЕ

Диагностика объектов является неотъемлемой частью поддержки необходимого уровня надежности, обеспечения требований безопасности и эффективности использования объектов.

Разработка и применение методов и средств диагностики позволяет повысить эффективность эксплуатации объекта. Важной частью диагностики является изучение и анализ диагностического сигнала.

При выборе диагностического сигнала для определения места возникновения дефекта, вида дефекта, степени его развития и для прогнозирования изменения

технического состояния объекта требуется большой объем диагностической информации.

Такие диагностические сигналы, как температура, давление и т. п., можно характеризовать практически только одним параметром – их величиной. Значительно больший объем диагностической информации содержится в виброакустическом сигнале – это его общий уровень, уровни в определенных полосах частот, соотношения между этими уровнями, амплитуды, частоты и начальные фазы каждой составляющей спектра частот, соотношения между амплитудами и частотами и т. д. Таким образом, именно виброакустические сигналы лучше всего подходят для решения задач диагностики и прогноза состояния объектов.

Еще одним важным обстоятельством в пользу выбора виброакустического сигнала в качестве диагностического является то, что это позволяет диагностировать объект на рабочем месте без остановки и разборки.

Виброакустический сигнал является одномерным и нестационарным. Виброакустический сигнал в диапазоне слышимых частот в дальнейшем будем называть

© Афонин Ю. С., Дубровин В. И., Манило Т. В., 2007

звуковым сигналом. Преимуществом выбора для анализа звукового сигнала является простота его регистрации.

Существует множество методов цифровой обработки одномерных сигналов [1, 2], в каждом из которых можно выделить ряд преимуществ и недостатков. Выбор метода для анализа сигнала зависит от поставленных целей и задач, требований к скорости обработки, точности, устойчивости, возможности использования данного метода в режиме реального времени и т. п.

В работе рассмотрены известные методы обработки сигналов, а именно преобразования Фурье и вейвлет-преобразование, в контексте задачи диагностики износа режущего инструмента по звуковому сигналу.

МЕТОД ПРЕОБРАЗОВАНИЯ ФУРЬЕ

Наиболее широко используемым методом цифровой обработки сигнала является преобразование Фурье (ФП). Звуковой сигнал является функцией времени. В большинстве случаев наиболее значимая информация скрыта в частотной области сигнала, поскольку информация, незаметная во временной области, проявляется в его частотном представлении. В таких случаях необходимо получить оценку составляющих сигнала в частотных диапазонах. В результате ФП сигнала, заданного во временной области, можно получить его частотное представление [2, 3]. Сигнал $s(t)$ может быть декомпозирован с помощью ФП $F(\omega)$, исходя из выражения (1)

$$F(\omega) = \int_{-\infty}^{+\infty} f(t)e^{-i\omega t} dt, \quad (1)$$

где $f(t)$ – исходный сигнал, $F(\omega)$ – спектр функции $f(t)$, ω – частота функции.

Дискретное ФП (ДФП) выполняется по формуле (2)

$$F_d(\omega) = \Delta t \sum_{-\infty}^{+\infty} x_n e^{-i\omega n \Delta t}, \quad (2)$$

где $x_n = f(n\Delta t)$ – выборочные данные функции $f(t)$.

Фурье-преобразование успешно применяется при обработке звукового сигнала процессов различной природы. Для примера рассмотрим звуковые сигналы процесса резания при диагностике износа режущего инструмента. Экспериментальные результаты показывают, что амплитуда звукового сигнала в частотной области чувствительна к изменению степени износа инструмента.

На рис. 1 изображены звуковые сигналы процесса резания, полученные при различной степени износа режущего инструмента ($<0,15$ мм, $0,15-0,3$ мм, $>0,3$ мм). На рис. 2 приведены результаты быстрого ФП (БФП) звуковых сигналов, представленных на рис. 1.

Реализация БФП приводит к существенному уменьшению вычислительной сложности ДФП. При использовании алгоритмов БФП объем вычислений сокращается примерно на два порядка [1]. Как видно из рис. 2, с увеличением износа режущего инструмента увеличивается амплитуда звукового сигнала в некоторых частотных областях.

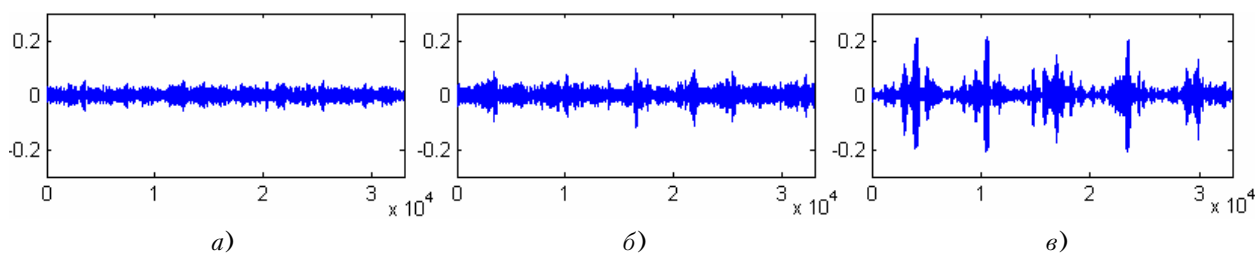


Рисунок 1 – Звуковые сигналы процесса резания при различной степени износа инструмента:

$a - <0,15$ мм, $b - 0,15-0,3$ мм, $v - >0,3$ мм

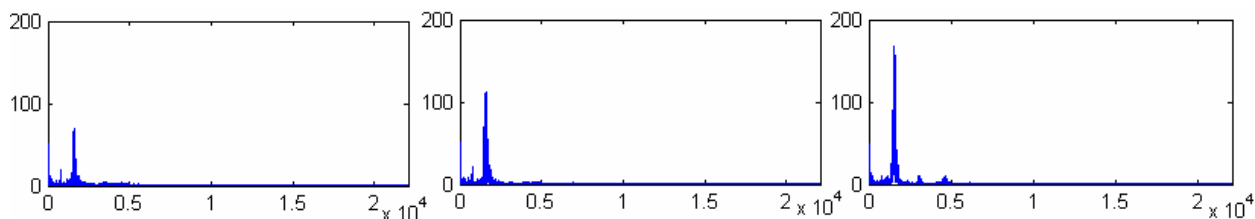


Рисунок 2 – Результаты БФП

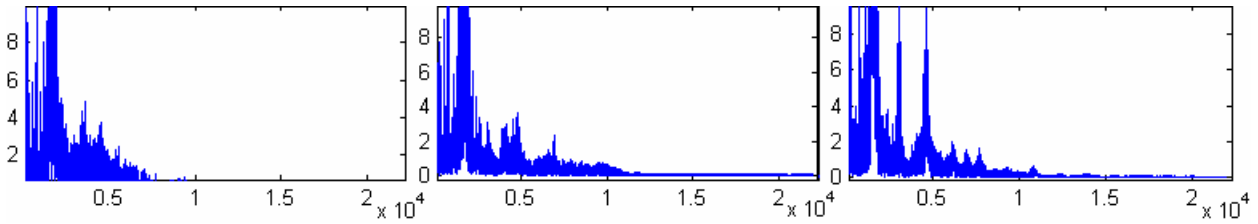


Рисунок 3 – Результати БПФ, масштаб 1:10 по осі ординат

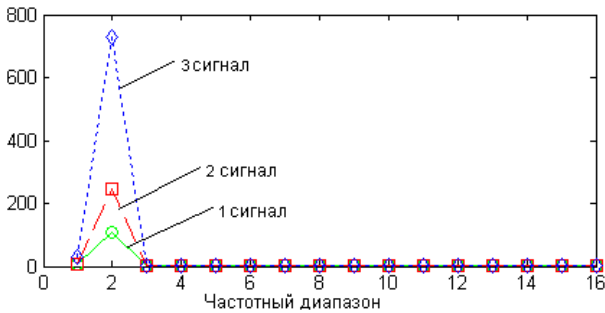


Рисунок 4 – Чувствительность результатов БПФ в 16 диапазонах к изменению износа

На рис. 3 для наглядности показаны результаты БПФ, увеличенные в 10 раз по оси ординат. Ось абсцисс захватывает диапазон частот от 0 до 20 кГц. Разбив область частот на 16 диапазонов, можно более детально определить чувствительность диапазона к изменению износа.

Среднеквадратические значения (СКЗ) результатов БПФ в каждом из 16 диапазонов, определяющие энергетический уровень сигналов в соответствующем диапазоне, показаны на рис. 4.

Как видно из рис. 4, лишь значения в первых диапазонах заметно отличаются для трех сигналов, которым соответствует различная степень износа инструмента. Как наиболее чувствительные диапазоны можно выделить первый и второй, которым соответствуют область частот 0–1250 Гц, 1250–2500 Гц.

СКЗ результатов БПФ сигнала в полученных диапазонах могут служить как информативные характеристики для оценки состояния исследуемого объекта, в данном случае износа режущего инструмента.

МЕТОДЫ ВЕЙВЛЕТ-ПРЕОБРАЗОВАНИЯ: БЫСТРОЕ И ПАКЕТНОЕ

При вейвлет-преобразовании (ВП) сигнал декомпозируется при помощи семейства базовых функций, называемых вейвлетами, которые имеют две независимые переменные: время и частоту. В результате получаем частотно-временное представление сигнала [4, 5]. Результатом ВП являются аппроксимирующие и детализирующие коэффициенты, которые в общем случае можно вычислить по формулам (3) и (4), соответственно.

$$c_{j,k} = \frac{1}{\sqrt{2^j}} \sum_{t=0}^L s(t) \phi\left(\frac{t-2^j k}{2^j}\right), \quad (3)$$

$$d_{j,k} = \frac{1}{\sqrt{2^j}} \sum_{t=0}^L s(t) \psi\left(\frac{t-2^j k}{2^j}\right), \quad (4)$$

где $c_{j,k}$, $d_{j,k}$ – коэффициенты аппроксимации и детализации, j – параметр масштаба или уровень разложения, k – параметр сдвига, $s(t)$ – исходный сигнал, L – длина сигнала, Φ – масштабирующая функция, ψ – вейвлет функция.

Рекурсивные выражения для расчета коэффициентов вейвлет-разложения, позволяющие рассчитывать коэффициенты разложения на основании коэффициентов предыдущего уровня (рис. 5, а), имеют название каскадного алгоритма (алгоритма Маллата), а ВП, выполняемое по данному алгоритму, называют быстрым (БВП) [6, 7].

Аппроксимирующие и детализирующие коэффициенты при БВП могут быть получены из выражений (5) и (6), соответственно.

$$c_{j+1,k} = \sum_n h_n c_{j,2k+n}, \quad (5)$$

$$d_{j+1,k} = \sum_n g_n c_{j,2k+n}, \quad (6)$$

где h_n , g_n – вейвлет-последовательность и масштабирующая последовательность для заданного уровня, n – количество коэффициентов последовательности h и g .

Для цифровых данных в качестве значений аппроксимирующих коэффициентов на нулевом уровне $c_{0,k}$ обычно принимаются исходные значения данных, исходя из выражения (7).

$$c_{0,k} = s(k), \quad (7)$$

где $s(k)$ – значение в k -й точке исходного сигнала.

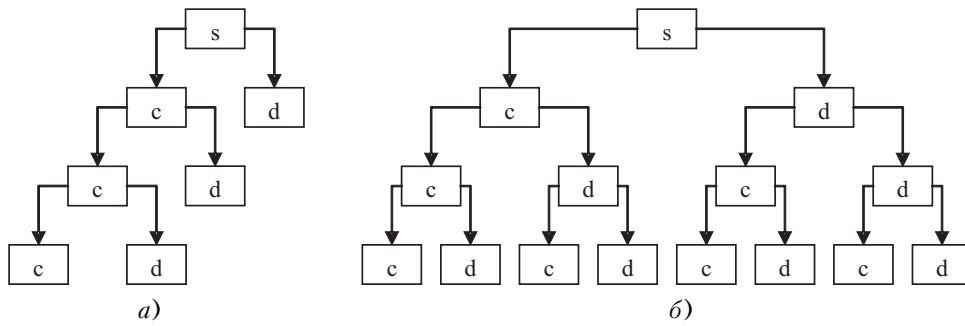


Рисунок 5 – Структура вейвлет-представления сигнала:
а – при БВП, б – при ПВП

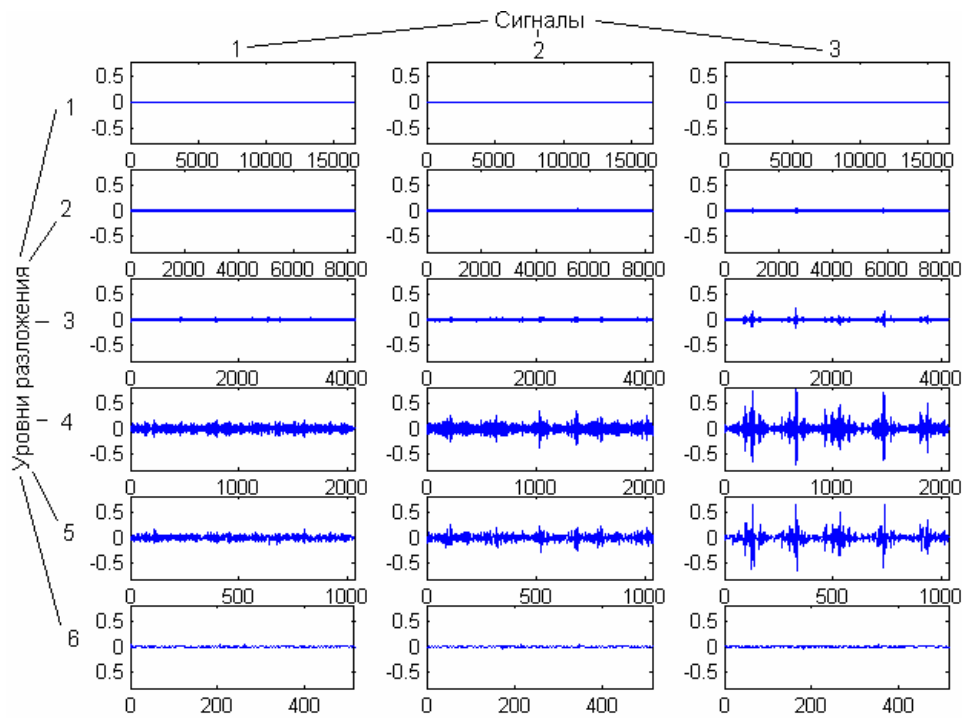


Рисунок 6 – Детализирующие коэффициенты на 6 уровнях при БВП

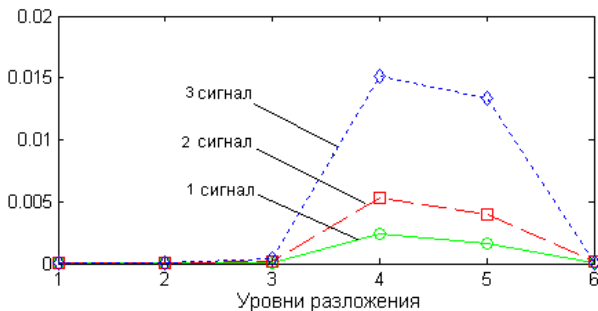


Рисунок 7 – Чувствительность СКЗ детализирующих коэффициентов на 6 уровнях при БВП к изменению износа

Установлено, что информативным параметром, характеризующим состояние режущего инструмента, является энергия детализирующих коэффициентов вейвлет-разложения сигналов. На рис. 6 показаны детализирующие коэффициенты, как результаты декомпозиции трех сигналов, представленных на рис. 1, на шести уровнях.

На рис. 7 изображена чувствительность СКЗ детализирующих коэффициентов на каждом из шести уровней к изменению износа. Из рисунка видно, что наиболее информативными являются 4-й и 5-й уровни разложения. В подобных задачах СКЗ детализирующих коэффициентов в выделенных уровнях могут быть использованы для определения состояния исследуемого объекта.

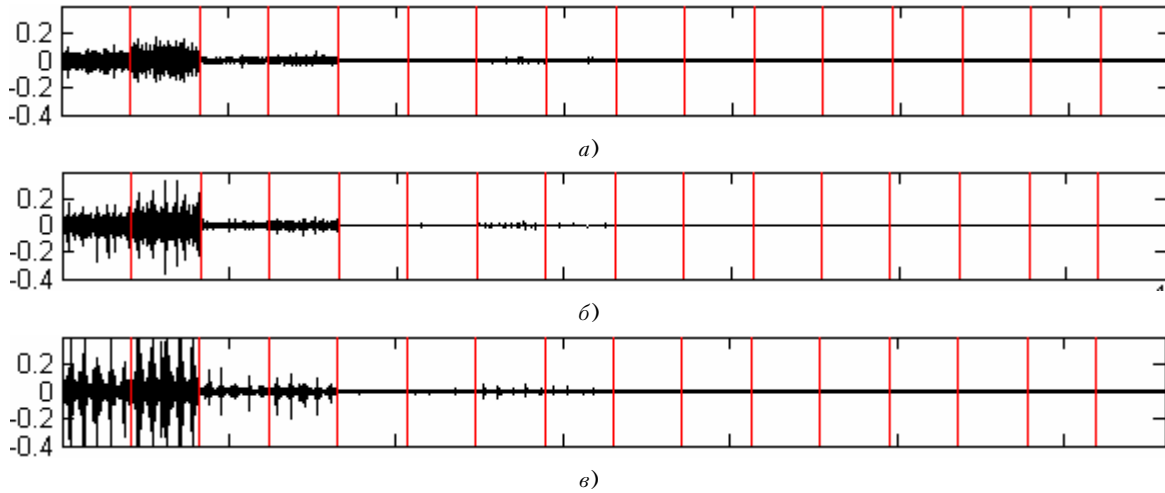


Рисунок 8 – Результати ПВП на четвертому рівні розкладу для сигналів, представлених на рис. 1, а (а), рис. 1, б (б), рис. 1, в (в)

При БВП на кожному шагу декомпозирується лише низькочастотна компонента сигналу. Ефективне використання даного алгоритму виходить з факта про більшу інформативність низькочастотної частини спектру сигналу, що для багатьох випадків не завжди справедливо.

Не виключено, що більш детальну інформацію про характеристики звукового сигналу в деяких випадках можна отримати, розклавши сигнал за допомогою пакетів вейвлетів, тим самим, проаналізувавши як низькочастотні, так і високочастотні складові частини сигналу. Дану процедуру називають пакетним вейвлет-перетворенням (ПВП) [6–9]. В цьому випадку відбувається заміна вейвлета $\psi(t)$ на два нових вейвлета: (8) і (9).

$$\psi_{2j}(t) = \sqrt{2} \sum h_n \psi_j(2t - n), \quad (8)$$

$$\psi_{2j+1}(t) = \sqrt{2} \sum g_n \psi_j(2t - n). \quad (9)$$

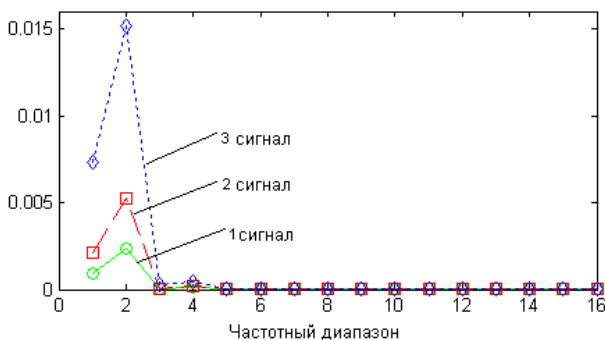


Рисунок 9 – Чувствителість вікон при ПВП до зміни зносу

Схема ПВП представлена на рис. 5, б.

На рис. 8 показані результати ПВП на четвертому рівні. В результаті маємо 16 вікон. СКЗ коефіцієнтів в цих вікнах визначають енергетичний рівень сигналу в відповідних частотних діапазонах. На рис. 9 показана чутливість вікон до змін зносу. Вибрані вікна можуть бути прийняті за значимі параметри в задачі діагностики стану об'єкта.

СРАВНЕНИЕ МЕТОДОВ

По результатам ПВП (рис. 9) чітко виражені відмінності значень енергетичних рівнів сигналів в першому, другому діапазонах і, менш помітно, в четвертому діапазоні, в той час як за результатами перетворення Фур'є (рис. 4) дані відмінності були помітні тільки во другому і, в меншій ступені, в першому діапазонах. Отже, ПВП може в більшій мірі описати стан досліджуваного об'єкта. При цьому застосування методу ПВП вимагає більше часу на обробку звукового сигналу, ніж перетворення Фур'є (БФП), що можна побачити з табл. 1. В даній таблиці наведено порівняння методів БФП, БВП, ПВП за швидкістю обробки сигналу.

Таблиця 1 – Порівняння методів за часом обробки сигналу

Час обробки одного сигналу, с	БФП	БВП	ПВП
мінімальне	0,015	0,015	0,079
максимальне	0,016	0,031	0,141

Данные параметры могут варьироваться в зависимости от длины обрабатываемого сигнала, частоты дискретизации сигнала, программной среды, в которой обрабатывается сигнал, способа реализации алгоритма, быстродействия ЭВМ и т. д. Поэтому приведенные результаты имеют относительный характер.

ЗАКЛЮЧЕНИЕ

Выбор метода обработки сигнала не может ограничиваться лишь значением параметра скорости обработки. Он должен вытекать из контекста поставленных задач и целей. Примером может быть необходимость наблюдения сигнала во времени. В таком случае применение преобразования Фурье будет нецелесообразным, поскольку результаты будут усреднены по всей длительности сигнала, в то время как вейвлет-преобразование успешно справится с такой задачей. Немаловажно учитывать требования к предъявляемой точности обработки сигнала. Обычно достижение высокой точности сопровождается увеличением длительности обработки сигнала и наоборот. Следовательно, главной задачей является подбор рационального способа обработки сигнала, чтобы с приемлемой скоростью и в достаточной мере описать необходимые характеристики акустического сигнала.

ПЕРЕЧЕНЬ ССЫЛОК

1. Гольденберг А. М., Матюшкин Б. Д., Поляк М. Н. Цифровая обработка сигналов: Учеб. пособие для вузов. –

2-изд., перераб. и доп. – М.: Радио и связь, 1990. – 256 с.

2. Оппенгейм А. В., Шафер Р. В. Цифровая обработка сигналов: Пер. с англ./ Под ред. С. Я. Шаца. – М.: Связь, 1979. – 416 с.
3. Поликар Р. Введение в вейвлет-преобразование. – СПб.: АВТЭКС, 2001. – 59 с.
4. Астафьева Н. М. Вейвлет-анализ: основы теории и примеры применения // Успехи физических наук. – 1996. – Т. 166, № 1. – С. 1145–1170.
5. Новиков Л. В. Основы вейвлет-анализа сигналов. Учебное пособие. – СПб., МОДУС+, 1999. – 152 с.
6. Дремин И. М., Иванов О. В., Нечитайло В. А. Вейвлеты и их использование // Успехи физических наук. – 2000. – Т. 171, № 5. – С. 465–501.
7. Дьяконов В. П. Вейвлеты. От теории к практике. Изд. 2-е, перераб. и доп. – М.: Солон-Пресс, 2004. – 400 с.
8. Чуи К. Введение в вейвлеты. М.: Мир, 2001. – 412 с.
9. Jaideva C. Goswami, Andrew K. Chan. Fundamentals of wavelets. Theory, algorithms and applications. – Wiley Interscience, 1999. – 324 p.

Надійшла 11.09.07

Представлені методи вейвлет-перетворення (швидке та пакетне) та перетворення Фур'є для обробки віброакустичного сигналу в контексті їх використання в задачах діагностики об'єктів. Для кожного методу виявлені характеристики (результати обробки віброакустичного сигналу), пов'язані зі станом об'єкта, що досліджується. Виконано порівняння методів за часом обробки сигналів.

The methods of wavelet analysis (fast and packet wavelet transform) and Fourier-transform for vibroacoustic signal processing in the context of their application for the task of object diagnostics are presented. The characteristics (results of vibroacoustic signals processing), related to the object state under consideration are revealed for each method. The methods comparison by signal processing time is performed.

УДК 658.5.011.56:681.3

В. И. Дубровин, Л. Ю. Дейнега

ПОДСИСТЕМА ВИЗУАЛИЗАЦИИ БАЗЫ ЗНАНИЙ ЭКСПЕРТНОЙ СИСТЕМЫ ТЕХНИЧЕСКОЙ ДИАГНОСТИКИ

В статье представлена автоматизированная подсистема визуализации базы знаний, применение которой упрощает работу пользователя с экспертной системой. Приведена структура представления знаний, функциональные возможности подсистемы и пример ее использования.

ВВЕДЕНИЕ

Одним из перспективных направлений систем поддержки принятия решений являются экспертные системы (ЭС). ЭС – особые компьютерные программы, мо-

делирующие действия эксперта-человека при решении задач в какой-либо предметной области на основе накопленных знаний, составляющих базу знаний (БЗ) [1].

Основной особенностью ЭС является ориентация на решение практических задач в неформализованных предметных областях. При этом результаты работы ЭС сравнимы с результатами, получаемыми человеком-экспертом. Накопленные знания определенным образом представлены в памяти ЭВМ в виде БЗ и, в отличие от знаний человека-эксперта, достаточно мобильны [2].

© Дубровин В. И., Дейнега Л. Ю., 2007

Классическая ЭС состоит из механизма логического вывода, БЗ и интерфейса пользователя. При работе с такой системой часто может возникнуть необходимость прибегнуть к помощи разработчика. В данной работе представлена автоматизированная подсистема визуализации знаний, которая устраняет этот недостаток. Эта подсистема может быть как частью ЭС, так и системы, основанной на формализованных знаниях.

**ПОДСИСТЕМА ВИЗУАЛИЗАЦИИ
В СТРУКТУРЕ ЭКСПЕРТНОЙ СИСТЕМЫ**

Структура экспертной системы представлена на рис. 1. Пользователь посылает запрос экспертной системе. Механизм логического вывода, используя знания, накопленные в БЗ, выдает рекомендации системы. Для удобства работы эксперта при пополнении БЗ процесс работы со знаниями должен абстрагироваться от модели их представления. Эти возможности предоставляет подсистема визуализации БЗ [3] (см. рис. 2).

При работе с ЭС такой структуры пользователю предоставляется инструментарий работы с БЗ без участия разработчика, а сам процесс работы со знаниями абстрагируется от модели их представления. Пользователь (или эксперт), работая с графическими объектами и их связями, может также визуальным образом отслеживать процесс принятия решения, проверять все дерево решения [4].

ЭС в современном понимании – прецедентные ЭС. Такие системы в корне отличаются от своих предше-

ственниц тем, что их базы знаний образуют не только логические правила, но и базу прецедентов (БП). Прецедент – это описание проблемы или ситуации в совокупности с подробным указанием действий, предпринимаемых в данной ситуации для решения данной проблемы [5].

В прецедентных экспертных системах чаще всего используется продукционная модель, т. е. модель, основанная на правилах в виде предложения типа «если (условие), то (действие)». Под «условием» (антецедентом) понимают некоторое предложение-образец, за которым осуществляется поиск в базе знаний, а под «действием» (консеквентом) – действия, которые будут выполнены при благоприятном исходе поиска.

СТРУКТУРА ПРЕДСТАВЛЕНИЯ ЗНАНИЙ

Продукционная модель отличается высокой модульностью, легкостью внесения дополнений и изменений, простотой механизма логического вывода. Но представление БЗ в виде предложений-продукций не является достаточно наглядным, что не позволяет изучить представленную предметную область в общесистемном подходе. Вдобавок, пополнение БЗ нуждается в определенных сведениях относительно формата записи правил-продукций. Поэтому стоит вопрос построения легкого в использовании визуального интерфейса для создания и управления БЗ ЭС.

Интуитивно понятной пользователю и возможной к применению в ЭС любого типа является древовидная структура представления знаний. Любой технологический объект можно представить в виде дерева, корнем которого является сам технологический объект, узлы дерева следующего уровня иерархии – некоторые функциональные блоки, на которые можно разбить объект, и так далее. Листы такого дерева представляют собой конечные элементы объекта, которые могут описываться некоторым состоянием из возможного набора.

Возможные результаты работы ЭС также можно представить в виде отдельного дерева. Например, для задачи диагностики можно группировать меры, выполнение которых рекомендуется для устранения неисправности. В простейших случаях дерево результата может вырождаться до одного узла с набором вариантов. В общем случае БЗ представляет собой множество входных параметров и множество результатов, каждое из которых представляется в виде семейства деревьев.

Пример такого представления БЗ приведен на рис. 3. На нем узел s0 может представлять собой устройство, которое подлежит диагностике; узлы s1, s2 – его отдельные блоки; s3, s4, s5 – элементы блока s1, при этом элемент s3 может иметь одно из состояний {o1, o2, o3}, s4 – одно из состояний {o4, o5}, s5 – одно из состояний {o6, o7, o8, o9}; аналогично – узлы s6 и s7.

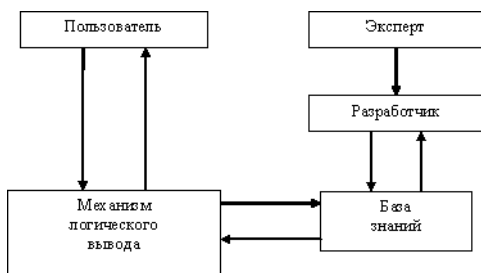


Рисунок 1 – Структура ЭС

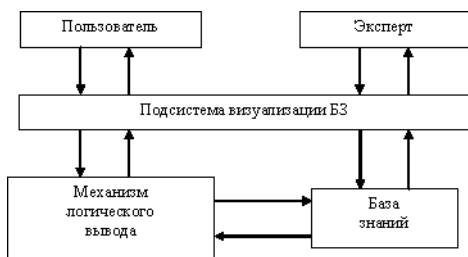


Рисунок 2 – Структура расширенной ЭС

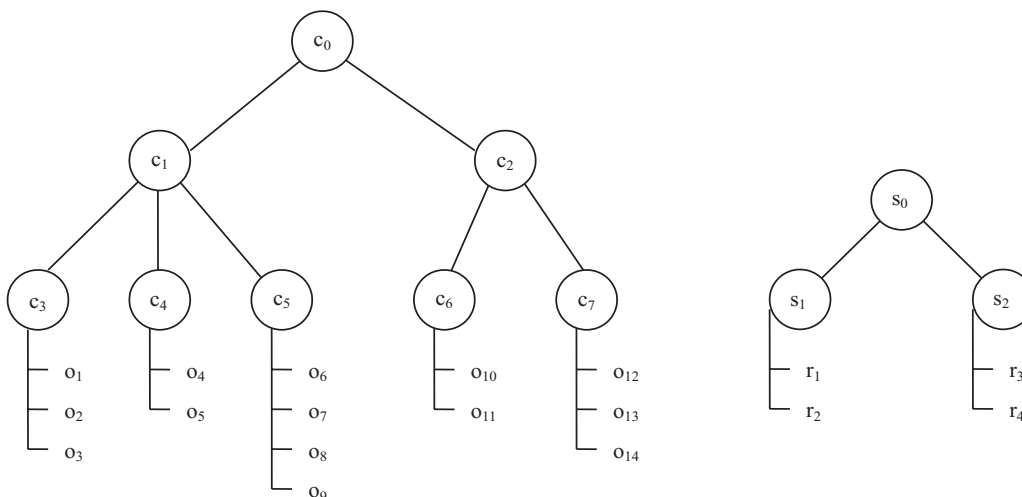


Рисунок 3 – Пример древовидного представления БЗ ЭС

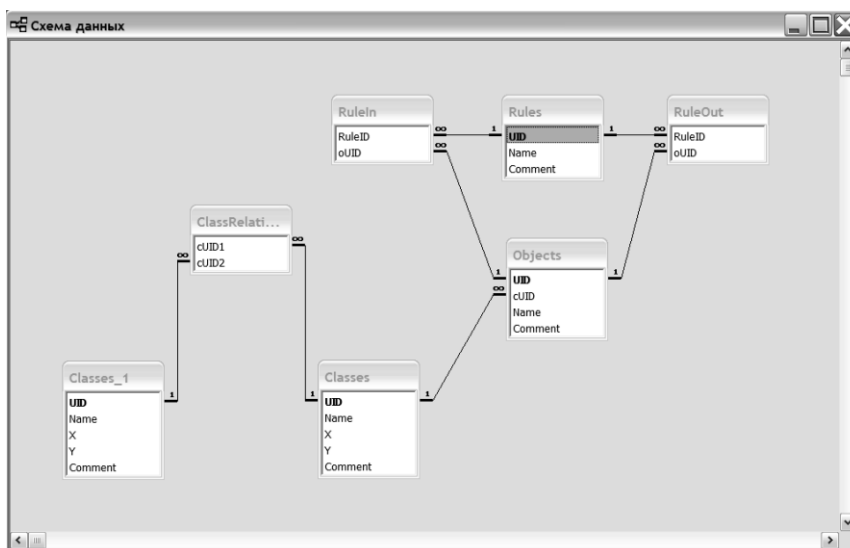


Рисунок 4 – Схема представления данных

Узел s0 является корнем другого дерева и может соответствовать понятию «неисправность»; s1 и s2 – две группы неисправностей с возможными значениями соответственно {r1, r2} и {r3, r4}.

Такое представление является лишь структурным. Семантическая нагрузка в виде правил-продукций, которые задают отношения между листовыми узлами дерева, должны задаваться дополнительно. Типичное правило-продукция, которое соответствует схеме на рис. 3, может иметь вид:

если c3 = o1 и c4 = o5 и c7 = o13, то s2 = r3.

Как видно из этого примера, нелистовые узлы деревьев не используются в правилах, т. к. несут в себе

лишь структурную составляющую. Если предоставить возможность этим узлам также иметь свои множества состояний, это придаст системе большей гибкости.

Схема представления данных подсистемы приведена на рис. 4.

ФУНКЦИОНАЛЬНЫЕ ВОЗМОЖНОСТИ ПОДСИСТЕМЫ ВИЗУАЛИЗАЦИИ БЗ

Разработанная подсистема визуализации предоставляет такие возможности:

- создание новой базы знаний;
- загрузка существующей базы знаний;
- сохранение измененной базы знаний;

- добавление, редактирование, удаление понятия базы знаний;
- добавление, редактирование и удаление вариантов состояний понятий;
- добавление и удаление взаимосвязи между понятиями;
- добавление, редактирование и удаление правила базы знаний.

Рабочее окно подсистемы содержит следующие элементы:

- главное меню, с помощью которого можно выполнить действия, которые предлагает программа;
- панель инструментов, которая дублирует пункты раздела «редактор» главного меню;
- рабочее пространство, которое позволяет работать с текущей базой знаний;
- список правил и два списка, которые хранят antecedent и консеквент текущего правила;
- поле для ввода комментария и панель статуса.

ПРИМЕР РАБОТЫ ПОДСИСТЕМЫ ВИЗУАЛИЗАЦИИ БЗ

Работа подсистемы была успешно протестирована для визуализации БЗ, содержащей сведения о мобильных телефонах, поступающих в сервисный центр. БЗ содержит понятия «Производитель», «Модель», «Номер», «Неисправность», «Причина», «Материалы» и «Описание работ», а также около 26000 правил, которые устанавливают связь между значениями этих понятий, причем понятие «Неисправность» используется в качестве выхода и входа правил (неисправность заявленная и неисправность фактическая).

Пример рабочего окна представлен на рис. 5 (окно разбито на две части для полного просмотра). Левая его часть содержит понятия БЗ (понятие «Номер»), подразумевающее серийный номер телефона, обозначено как «IMEI») и их значения. Правая часть – список правил БЗ (их номера, имена и комментарии к ним) и входные-выходные значения понятий для выделенного правила. В данном примере выделено правило № 24 с именем Rule20.

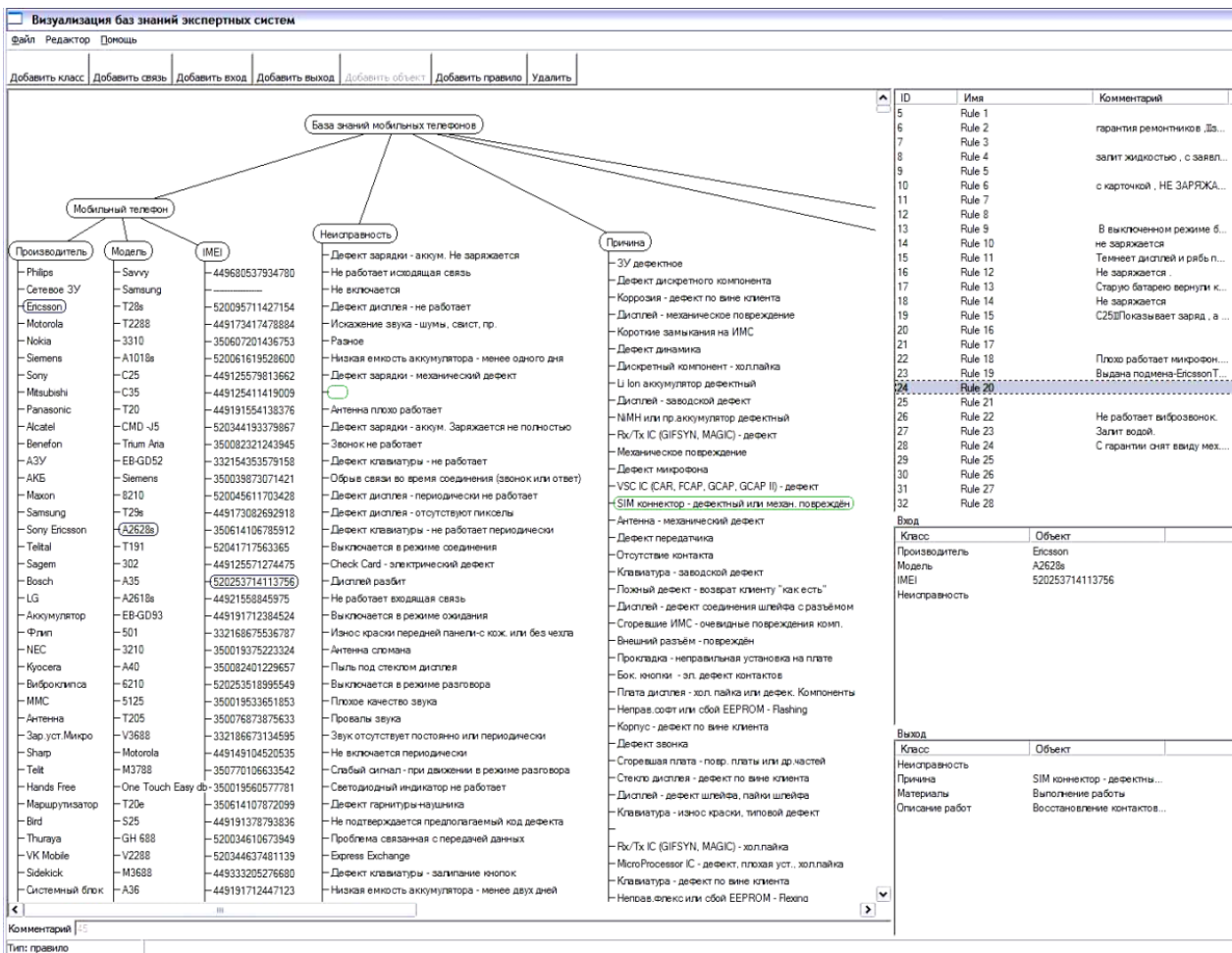


Рисунок 5.1 – Пример рабочего окна программы

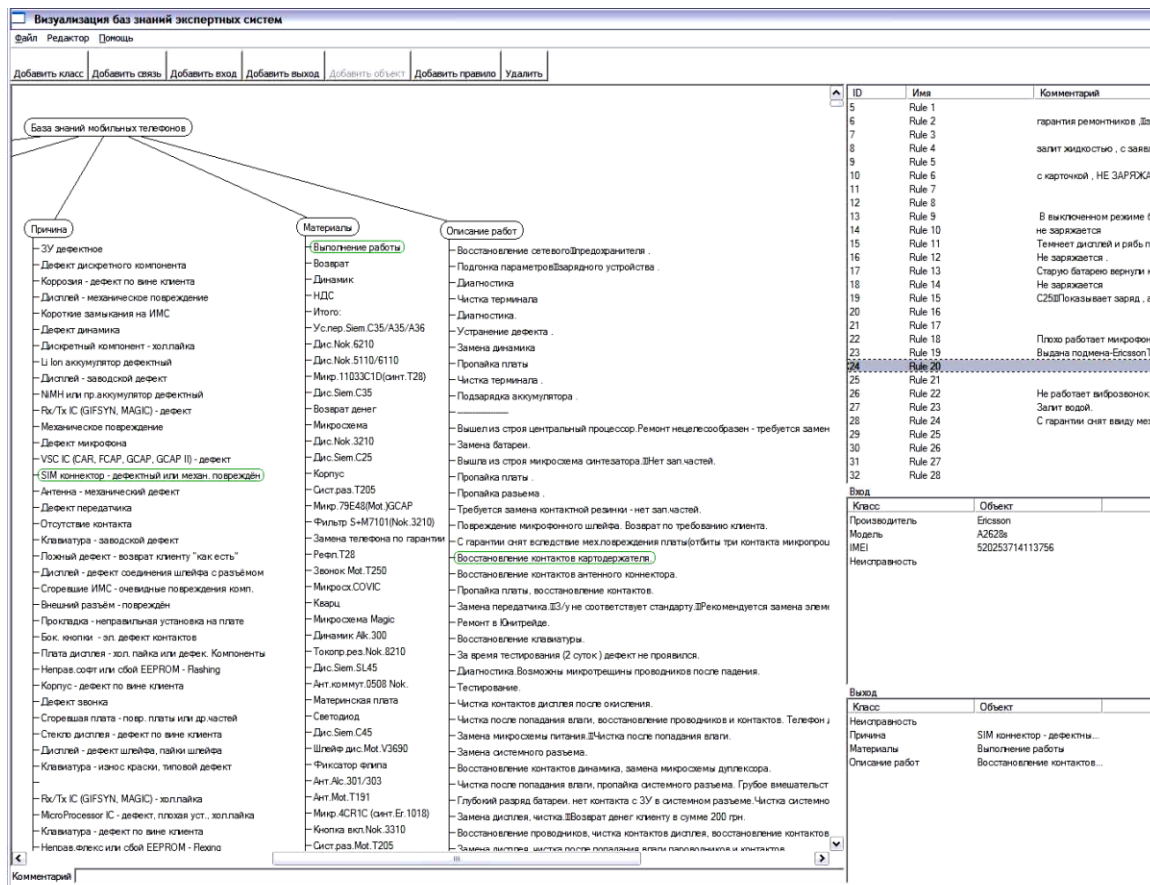


Рисунок 5.2 – Пример рабочего окна программы

ЗАКЛЮЧЕНИЕ

Разработанная подсистема визуализации является полноценной частью ЭС, а при необходимости может быть расширена с помощью следующих действий:

- добавлением новых библиотек для работы с разными форматами баз знаний;
- добавлением библиотеки для поиска необходимых правил в БЗ большого объема;
- добавлением возможности экспорта данных для языков логического программирования;
- использованием технологий лингвистического обеспечения;
- разработкой альтернативных режимов визуализации;
- переориентацией на использование подсистемы в сети Интернет.

ПЕРЕЧЕНЬ ССЫЛОК

1. Сойер Б., Фостер Д. Программирование экспертных систем. – М.: Финансы и статистика, 1998. – 254 с.

2. Дубровин В. И., Дейнега Р. В., Дейнега Л. Ю. Основные подходы к проектированию систем поддержки принятия решений // Проектирование и технология электронных средств, 2003. – № 2. – С. 6–10.

3. Дубровин В. И., Дейнега Л. Ю. Визуализации знаний в системах диагностики авиационных двигателей // Вестник двигателестроения, 2006. – № 1. – С. 88–90.

4. Дубровин В. И., Дейнега Л. Ю., Дейнега Р. В. Подсистема визуализации для работы с базой знаний // Академический вестник, 2004. – № 13. – С. 57–61.

5. Черняховская Л. Р., Никулина Н. О., Халиков Т. А., Федорова Н. И., Водопьянов Р. В. Разработка динамической модели процесса управления в проблемных ситуациях на основе базы знаний прецедентов // Управление в сложных системах, 1999. – № 2. – С. 207–212.

Надійшла 15.10.07

В статті представлена автоматизована підсистема візуалізації бази знань, застосування якої спрощує роботу користувача з експертною системою. Наведена структура представлення знань, функціональні можливості підсистеми і приклад її використання.

The automated subsystem of visualization of knowledges base, application of which simplifies work of user with a expert system, is presented in the article. The structure of knowledges representation, functional possibilities of subsystem and example of its use, is resulted.

Г. Г. Киричек

ЕЛЕКТРОННА БІБЛІОТЕКА ЯК ОБ'ЄКТ ОСВІТНІХ ТЕХНОЛОГІЙ

Мета дослідження – визначення основних засад та напрямків розвитку електронної бібліотеки, технологічне визначення принципів і джерел комплектування для забезпечення створення електронної бібліотеки.

ВСТУП

Зміни, що зазнає бібліотечно-інформаційна діяльність вищого навчального закладу (ВНЗ) в умовах впровадження сучасних інформаційних технологій, є такими, що потребують проведення наукових досліджень на новому рівні. Розповсюдження у науковому товаристві нових засобів і форм електронних ресурсів та надання інформації в електронному вигляді, змушують бібліотечних працівників та науковців вузів надавати все більшого значення електронним бібліотекам (ЕБ). В системі інформаційного забезпечення науки і освіти, питання створення та формування спеціалізованих ЕБ вузів є дійсно актуальним і забезпечується сукупністю державних заходів. Одним з яких є виконання Державної програми «Інформаційні та комунікаційні технології в освіті і науці» на 2006–2010 роки. При відсутності достатнього фінансування спеціалізовані ЕБ вузів дозволяють вирішити ряд питань, пов'язаних з книгозабезпеченням нових дисциплін в умовах відкритого навчання.

У кінці 80-х – початку 90-х років у США та Великобританії з'явилися перші публікації, присвячені різним аспектам створення ЕБ та надання доступу до них [1, 2]. Зарубіжні автори розглядають загальні проблеми формування ЕБ, концепції та стратегії розвитку, результати тестувань щодо комфортності використання ресурсів ЕБ та шляхи надання доступу до наукових електронних публікацій [3–5].

Перші публікації українських авторів із зазначеної проблематики відносяться до кінця 90-х років. Це, головним чином, оцінка сучасної ситуації та перспектив створення ЕБ [6], концептуальних засад [7, 8], опис досвіду формування електронних ресурсів та зберігання великих і надвеликих обсягів інформації [9, 10], проблеми лінгвістичного забезпечення та шляхи їхнього вирішення [11].

Автори більшості публікацій основними завданнями діяльності електронної бібліотеки вважають [9, 12, 10]:

- тематичне упорядкування накопичених масивів інформації для реалізації інформаційного пошуку та задоволення інформаційних потреб науковців;
- аналітико-синтетичне опрацювання електронних документів, створення їхніх ефективних змістових пошукових образів;
- надання комфортного онлайн-доступу до наукової інформації;
- ефективність використання ресурсів електронних бібліотек.

ОСНОВНІ ЗАСАДИ ТА НАПРЯМКИ РОЗВИТКУ ЕЛЕКТРОННОЇ БІБЛІОТЕКИ ВУЗУ

Поняття електронної бібліотеки є широким та емним, уміщує велику кількість різних аспектів та проблем [13]. Електронна бібліотека – не аналог традиційної бібліотеки, а складна інформаційно-бібліотечна система, що дозволяє надійно зберігати й ефективно використовувати різноманітні колекції електронних документів доступних через телекомунікаційні мережі зв'язку [8]. В монографії В. Армса надається таке «інформаційне» визначення: «електронна бібліотека – це керована колекція інформації у сукупності з відповідними сервісами, причому інформація зберігається у цифрових форматах і доступна по мережі [14]». Отже, створення електронних бібліотек пов'язано з формуванням цифрових колекцій, розподілених ресурсів і організацією широкого спектру онлайн-інформаційних сервісів [15].

Таким чином можна виділити коло питань, більшість з яких стосується проблем розробки створення та формування електронних бібліотек:

- організація робочого місця наукового співробітника. Чільне місце в системі інформаційного забезпечення науки займають електронні бібліотеки інститутів та університетів. Колекція ЕБ дозволяє переглядати, вивчати, роздруковувати та копіювати на диски чи дискети повну версію документа: курсу, програми, статті, книжки, навчальну та наукову інформацію. Особливістю названої електронної бібліотеки є поєднання різних форм представлення інформації та високий процент її використання [16];
- забезпечення інтерактивного навчання. Співробітники різних освітніх установ разом з бібліотечними фахівцями ведуть велику роботу зі створення принци-

пово нової технології інтерактивного навчання, заснованої на новій інформаційній інфраструктурі освітнього процесу, центром якої повинні стати електронні бібліотеки. При цьому передбачаються засоби адаптації програм навчання і його інформаційної підтримки до потреб і можливостей конкретного учня без яких-небудь обмежень на його вік, географічне місце розташування, розпорядок дня й т. ін.;

– створення «єдиного бібліотечного фонду (ЄБФ) на основі централізованого комплектування». При цьому до складу єдиного бібліотечного фонду повинні входити фонди видань і інших документів «всіх підрозділів вузу». ЄБФ містить у собі різні види «вітчизняних і закордонних видань». Це наукові і навчальні електронні документи для забезпечення науки та освіти;

– формування фонду «відповідно до тематичного плану комплектування і картотеки книгозабезпеченості навчального процесу». Тематичний план комплектування вузівської бібліотеки повинен відображати «профіль навчальних дисциплін вузу й тематику науково-дослідних робіт», а картотека книгозабезпеченості містити «інформацію про навчальні дисципліни, що читають у вузі; контингент студентів і формах їхнього навчання; видання, що рекомендують до використання в навчальному процесі незалежно від виду документа та коефіцієнта книгозабезпеченості».

Інформатизація вузівських бібліотек може бути визначена як створення оптимальних умов для задоволення інформаційних потреб вищих навчальних закладів за рахунок формування й використання інформаційних ресурсів. Інформаційні ресурси нового типу на сучасному етапі виступають основою успішного впровадження сучасних освітніх технологій у вузі. Асоціацією освітніх комунікацій і технологій США (АОКТ) запропоноване визначення поняття «освітня технологія» – це теорія й практика організації, розвитку, використання, керування й оцінки процесів і ресурсів навчання [17]. Дане визначення поміщає ресурси навчання, і, відповідно, серед їхнього різноманіття й інформаційні ресурси, безпосередньо в об'єкт освітніх технологій.

Основними напрямками роботи вузу є навчально-методична, науково-технічна й освітня діяльність. Тому модернізація, комп'ютеризація й інформатизація процесів вищої освіти дозволять розширити спектр інформаційно-методичних послуг, наблизять викладачів і студентів до актуальної для кожної з категорій інформації. Схема формування й використання електронної інформації в навчальному процесі наведена на рис. 1 [18].

Таким чином, можна сказати, що оптимальна організація й використання інформаційних ресурсів у сучасній системі освіти виступає об'єктом освітніх технологій. При цьому для комп'ютерних і мультимедійних освітніх технологій на перше місце висува-

ються нетрадиційні інформаційні системи нового типу, до яких можна віднести й ЕБ.

ПЕРЕВАГИ І ШЛЯХИ СТВОРЕННЯ ЕЛЕКТРОННОЇ БІБЛІОТЕКИ

Переваги створення електронної колекції:

- створення електронної копії друкованого видання дозволить відновити інформацію у випадку втрати;
- електронна книга ніколи не старіє й не губиться;
- працювати з електронною версією може будь-яка кількість співробітників з різних відділів разом з відвідувачами бібліотеки;
- можливість розширеного пошуку дозволяє шукати інформацію відразу по декількох характеристиках;
- за рахунок швидкого пошуку й можливості одночасного доступу до документів продуктивність праці може збільшитись в кілька разів;
- наявність електронних копій друкованих видань дає можливість користувачам в будь-який час переглядати інформацію.

Першим кроком на шляху до створення електронної бібліотеки є сканування. Це можливість одержання образу паперового документа в електронному вигляді. Сьогодні сканування актуально для бібліотек, освітніх установ і тих, для кого надання інтелектуальних знань завжди стоїть на першому місці. Для бібліотек можна виділити 2 основних типи сканування:

- дбайливе сканування рідких книг, унікальних рукописів і архівних матеріалів, що мають різний стан, надруковані на папері різної товщини, різних форматів (від А5 до А0), у зшитих і розшитих видах;
- потокове сканування часто запитуваних видань різних форматів (від А5 до А3) для одержання універсальних зборів повнотекстових електронних копій.

У зв'язку із цим бібліотекам доводиться вирішувати наступні завдання:

- як провести сканування документів різних форматів і якості, щоб на виході завжди мати документ із чітким зображенням і потрібною якістю;
- як зробити це швидко, надійно і якісно.

Для переведення бібліотечних фондів в електронний вид потрібно:

- проведення попереднього аналізу: вибір і уточнення кількості літератури, якості, форматів і видів документів, які необхідно буде відсканувати;
- сортування по форматах і старості видань із наступним вибором спеціального устаткування;
- вибір сканерів, програмного забезпечення для сканування, для розпізнавання тексту й перегляду зображень, для формування книг з відсканованих зображень;
- налаштування програмного забезпечення для одержання максимальної чіткості зображення;
- проведення самого процесу сканування;

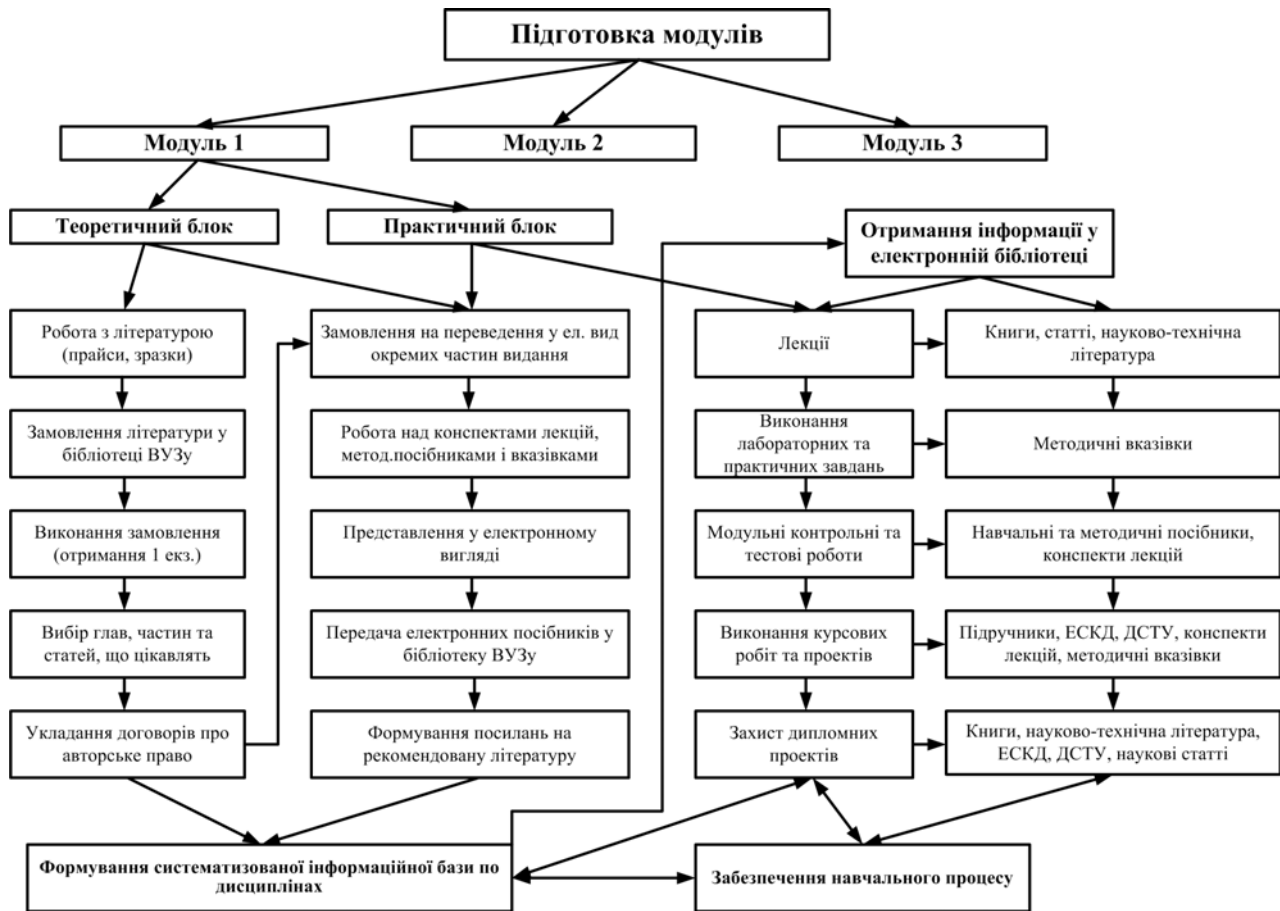


Рисунок 1 – Схема формування й використання електронної інформації в навчальному процесі

– збереження зображень у форматі TIFF, JPEG, PDF;

– перевірка на повну відповідність електронної версії паперовому оригіналу видання;

– коректування отриманих зображень, при необхідності підбор певних настроювань сканера з наступним перескануванням;

– редагування помилок;

– збереження отриманого масиву зображень у необхідних форматах і в заданій структурі;

– запис даних на електронний носій;

– заповнення вже існуючої інформаційної системи бібліотеки з можливістю паралельного перегляду уведеної інформації співробітниками й користувачами:

– внесення атрибутів для кожного електронного документа;

– занесення електронних документів у бібліотечну систему;

– розпізнавання або ручне введення тексту для подальшого здійснення необхідного пошуку по змісту картки або книги.

ФОРМУВАННЯ ЕЛЕКТРОННИХ БІБЛІОТЕК

Формування ЕБ здійснюється такими засобами: мануальними (набір текстової інформації, створення цифрових малюнків тощо); автоматизованими (методи сканування тексту, рисунків, слайд-сканування, імпорт цифрових фото, поповнення ЕБ вже готовими електронними документами: мультимедіа виданнями, відео тощо); автоматичними (використання пошукових робіт індексування та накопичення інформації, система імпорту електронних документів тощо). Робота повинна ґрунтуватись на принципі зацікавленості індивідуальних і колективних авторів (професорсько-викладацького складу та вчених навчального закладу) в тому, щоб їхні педагогічні та наукові досягнення увійшли до бібліотечної інформаційної системи, що надасть їм можливість довести до тих хто навчається інформацію про навчальні та методичні посібники, публікації та результати наукових досліджень. Інтеграційним ядром електронної бібліотеки повинні стати навчальні та методичні вказівки і рекомендації, конспекти лекцій, монографії та статті науковців, норма-

тивні документи, науково-теоретичні збірники, бібліографічні покажчики.

Основними принципами формування електронної бібліотеки є профільність, наукова цінність документів, їх відповідність читацьким вимогам. Формування повинно вестись згідно з навчальними дисциплінами та тематикою наукових досліджень.

Комплектування електронних колекцій здійснюється з 4-х напрямків:

- колекція навчальних документів (основних і необхідних додаткових інформаційних матеріалів);
- колекція наукових документів (монографії, наукові журнали, реферативні видання);
- довідково-інформаційна колекція (інформаційні документи з різноманітних галузей знань);
- фондоутворююча, спрямована на поповнення фонду наукової бібліотеки документами, які є тільки в електронному вигляді і заповнення існуючих у фонді прогалин за рахунок легітимного придбання електронних копій з друкованих документів.

Завдання, що ставляться перед комплекторами по відбору конкретних документів для конкретних бібліотек, як відомо, характеризуються їх значною складністю й багатоаспектністю. Для прийняття обґрунтованих рішень по первинному відборі електронних навчальних видань комплектору вже на рівні вторинної інформації необхідно мати у своєму розпорядженні повноцінний набір їх характеристик, що всебічно відбивають якість кожного конкретного видання. При цьому поряд із загальними, властивими всім навчальним виданням взагалі (науковість, точність і вірогідність наведених відомостей, чіткість і ясність викладу матеріалу, доступність інформації, лаконічність, логічність і послідовність, систематичність і спадкоємність матеріалу, що викладає, чіткість структури, відповідність мови викладення матеріалу нормам літературної мови, естетичність оформлення й ін.) до електронних навчальних видань пред'являються специфічні вимоги, властивій природі електронної інформації.

У силу того, що, політика й стратегія в області комплектування електронних матеріалів ще до кінця не розроблена, комплекторам доводиться керуватись досвідом роботи із друкованими документами [19]. При цьому ймовірність помилки при відборі електронних документів за рахунок їхньої принципової специфіки різко зростає. Тим самим відкривається шлях до включення в бібліотечний фонд електронних навчальних видань, що не відповідають вимогам модернізації освіти й підготовки висококваліфікованих фахівців.

Про необхідність створення нового підходу до комплектування сьогодні заявляють як теоретики, так і практики бібліотечної справи. Підкреслюючи акту-

альність цього питання, Шрайберг Я. Л. відзначає: «Сегодня в условиях тотального наступления электронных изданий и пересмотра многими издательствами своих планов по выпуску продукции в части выбора / сочетания электронных и печатных публикаций, обновление политики комплектования стало одной из заметных проявляющихся тенденций деятельности всех библиотек...» [19]. При цьому, як справедливо зауважують практики, «... документи на електронних носителях требуют собственных критериев отбора...» [20].

Електронні бібліотеки формуються на інформаційній та технологічній базі різних інформаційних установ: бібліотек, науково-дослідних інститутів, учбових закладів, архівів, музеїв, великих крупних видавництв, державних інформаційних органів тощо. Таким чином, об'єктом електронної бібліотеки може бути будь-яка інформація (документи, об'єкти матеріального світу, явища, дані), що зафіксована або відтворена у цифровій формі. Це обумовлює різноманіття електронних ресурсів та багатоаспектність їх типології. Так, за специфікою інформаційного вмісту електронні ресурси можна поділити на дві основні групи:

- текстові, які охоплюють усі види текстових документів (тексти природною мовою, листінги програм, статистичні дані, економічні розрахунки, технічна документація тощо);
- нетекстові, якими є цифрові образи речей та творів історико-культурної спадщини, зображувальні та картографічні матеріали, музичні твори, креслення, програмні продукти, аудіо- та відео матеріали тощо.

Усі колекції, що комплектуються повинні задовольняти вимозі єдності прийнятого формату.

Продумано підібрані колекції матеріалів для оцифрування, технології оцифрування й методи будівництва електронних бібліотек сприяють народженню нових інформаційних ресурсів, які: істотно підвищують ефективність освіти й досліджень; дозволяють створювати унікальні колекції з раніше розрізаних документів, які до того ж можуть знаходитись в різних бібліотеках і архівах, у тому числі іноземних; дозволяють виявити невідомі колекції шляхом об'єднання електронних описів матеріалів по певній тематиці з різних сховищ, які просто не знали про існування відповідних колекцій друг у друга [19].

Колекції електронних навчальних матеріалів, що формуються у вузах потрібно систематизувати; привести (бажано) до стандартного виду; довгочасно зберігати дані на надійних електронних носіях і при необхідності актуалізувати їх переважно силами самих авторів; організувати таким чином, щоб забезпечити зручний і відкритий доступ професорсько-викладацькому складу й студентам, які навчаються, у даному вузі, а при можливості й стороннім користувачам.

ВИМОГИ ДО СКЛАДУ ФОНДІВ ЕЛЕКТРОННОЇ БІБЛІОТЕКИ

В області створення фондів (колекцій) ЕБ визначаються вимоги до видового й тематичного складу фондів ЕБ на основі критеріїв їхньої соціальної значимості, а також ступеня готовності електронних ресурсів:

- формування представницької колекції електронних фондів ЕБ на базі наявних електронних інформаційних ресурсів;
- організація оцифровки особливо коштовних документів;
- організація комплектування фондів ЕБ на основі електронних форм друківаних видань;
- визначення складу й розробка типових варіантів і форматів подання основних типів даних в ЕБ (реляційних БД, повнотекстових, графічних і мультимедійних документів і даних), а також кодувань;
- розробка архітектури системи метаданих ЕБ, принципів і технологій її застосування;
- створення словників, класифікаторів і інших засобів лінгвістичного забезпечення;
- вибір і створення засобів автоматичної обробки даних природною мовою, а також графічних і мультимедійних даних;
- розробка й стандартизація програмно-технологічних засобів підготовки й завантаження в ЕБ документів і даних основних типів;
- розробка технології зберігання більших і надвеликих масивів в ЕБ, їхній архівації й забезпеченню доступу до них;
- вибір або розробка типових програмно-технологічних засобів ЕБ;
- вибір або розробка програмно-технологічних засобів захисту інформації в ЕБ; розробка або визначення вимог до системи навігації й користувальницьких інтерфейсів в ЕБ;
- вибір або розробка протоколів і стандартів для забезпечення сумісності програмно-технологічного забезпечення ЕБ;
- вибір або розробка програмно-технологічних засобів керування в ЕБ, включаючи контроль виконання послуг.

Фонд електронних документів складається з електронних публікацій як документів першого покоління та електронних копій друківаних видань. Останній напрямок може мати модифікації: якщо друківаний аналог відсутній у фонді, то для бібліотеки це буде документ першого покоління, призначений для постійного контингенту читачів; якщо ж копія знята з документа, що перебуває в бібліотеці, то вона буде призначатися й стороннім абонентам. Ці нюанси необхідно брати до уваги при визначенні політики комплектування.

Інформаційні ресурси можуть бути класифіковані по наступних ознаках. До них належить:

- тематика інформації: суспільно-політична, фінансово-економічна, екологічна, медична, природнонаукова, технічна й т. ін.;
- доступність інформації: відкрита, закрита, конфіденційна;
- приналежність до певної інформаційної системи: бібліотечна, архівна, науково-технічна й т. ін.;
- вид інформації: навчальна, наукова, науково-популярна, довідкова й т. ін.;
- форма подання інформації: текстова, цифрова, графічна, мультимедійна;
- вид матеріального носія інформації: друківана, електронна, аудіовізуальна й інша [21].

У структуру інформаційних ресурсів сучасної вузівської бібліотеки входять:

- інформаційний фонд на традиційних друківаних носіях;
- інформаційний фонд на електронних носіях.

Крім того, документи на електронних носіях вимагають власних критеріїв відбору. Це пов'язано з тим, що вони існують як:

- документи на електронних носіях, що мають паралельно копію на паперовому носії;
- документи на електронних носіях, що використовуються як локальні ресурси (компакт-диски, дискети).

ВИСНОВКИ

Необхідність забезпечення навчального процесу й наукових досліджень інформаційними ресурсами визначає доцільність впровадження в найкоротший термін нових методів надання інформації всім учасникам навчального процесу, а саме створенню повнотекстової навчально-методичної бази даних (БД), що включає методичні посібники й вказівки, конспекти лекцій, літературу й рекомендації, наукові статті по всіх дисциплінах і напрямках наукових досліджень.

Варіанти створення й використання електронних бібліотек доцільно розглядати й застосовувати в кожному конкретному вузі, тобто в процесі проведення досліджень його можливостей і потреб, у тому числі можливостей професорсько-викладацького складу, науковців й студентів цих вузів, що дозволяє реалізувати ту або іншу схему організації й експлуатації навчальної електронної бібліотеки.

Подання інформації в електронному вигляді дає незрівнянно більші можливості для роботи з даними, на відміну від друківаних видань. Дослідження показують, що використання електронного змісту або придбання матеріалів через ЕБ не тільки скорочують час на їхній пошук і одержання, але також значно зменшуються витрати на книгозбереження, замовлення й одержання замовленої літератури.

Формування ЕБ може здійснюватись двома шляхами – оцифровкою документів на паперових носіях

з фондів бібліотеки і придбанням електронних документів. Інформаційні ресурси ЕБ повинні складатись з двох частин – навчальної і наукової. До першої входять електронні підручники, навчальні посібники та методичні розробки, що полегшують освоєння студентами навчального матеріалу. Основу другої частини ЕБ складають електронні книги і статті.

«Електронна бібліотека» вузу може стати основою інформаційного забезпечення навчального процесу, науково-дослідної роботи студентів, аспірантів, здобувачів, викладачів і співробітників вузу.

ПЕРЕЛІК ПОСИЛАНЬ

1. *Barker Ph.* Electronic libraries-vision of the future // Electronic Library. – 1994. – Vol. 12. – No 4. – Pp. 221–230.
2. *Millar R. H.* Electronic Resources and Academic Libraries, 1980–2000: A Historical Perspective // Library Trends. – 2000. – Vol. 48. – No 4. – Pp. 645–670.
3. *Moyo L. M.* Electronic libraries and the emergence of new service paradigms // Electronic Library. – 2004. – Vol. 22. – No 3. – Pp. 220–230.
4. *Odendhal S.* World Wide virtual veterinary library-perils and pitfalls, but why not? // Animal Health Information: Structuring and Sharing, Global and Local, July 1–4, 1997. – Frederiksberg. – 1998. – Pp.110–113.
5. *Williams P.* Information for the public about disease: usability issues in the development of the National electronic library for communicable diseases // Aslib proceedings. – 2004. – Vol. 56. – No 2. – Pp. 99–103.
6. *Щепко В.* Електронні бібліотеки в Україні: перспективи розвитку // Бібл. вісн. – 2001. – № 5. – С. 31–34.
7. *Чекмарьов А. О.* Національна система електронних бібліотек / Чекмарьов А. О., Костенко Л. Й., Павлуша Т. П. // НАН України, Нац. б-ка України ім. В. І. Вернадського. – К.: НБУВ, 1998. – 52 с.
8. *Шемаєва Г.* Галузева електронна бібліотека: концептуальні положення / Шемаєва Г., Шаповалова Т., Приходько Т. // Вісн. Кн. палати. – 2002. – № 2. – С. 15–17.
9. *Баркова О. В.* Досвід створення наукової електронної бібліотеки в Національній бібліотеці України імені В. І. Вернадського // Реєстрація, зберігання і обробка даних. – 2001. – Т. 3, № 4. – С. 51–62.
10. *Костишин О., Широков В.* Трирівнева система цифрових бібліотек з технологією «клієнт – сервер» // Бібл. вісн. – 2002. – № 3. – С. 39–42.
11. *Бардієр К. В.* Лінгвістичне забезпечення електронних бібліотек // Реєстрація, зберігання і обробка даних. – 2001. – Т. 3, № 3. – С. 78–87.
12. *Глухов В. А.* Электронные библиотеки. Организация, технология и средства доступа / Глухов В. А., Голицын О. А., Максимов Н. В. // НТИ. – 2000. – № 10. – С. 1–8. – (Сер. 1).
13. *Буль В. А.* Что хранят электронные библиотеки России? // Библ. Дзюло. – 2003. – № 3. – С. 16–19.
14. *Армс В.* Электронные библиотеки: Учебн. пособие / Пер. с англ. Арнаутова С. А. – М. – 2002.
15. *Лапо П.* Виртуальная библиотека и будущее научных коммуникаций // Бібл. світ. – 2001. – № 3. – С. 9–11.
16. *Майстрович Т. В.* Электронный документ как компонент фонда [Электронный ресурс] Дис... д-ра пед. наук: 05.25.03. – М.: РГБ, 2006.
17. *The Field of Educational Technology: Update 2000 A Dozen Frequently Asked Questions / Donald P.* // ERIC Digest. – March 2000. – EDO-IR – 2000. – 01.
18. *Киричек Г. Г.* Управление информационными потоками вуза – как подсистема инновационного образования / Г. Г. Киричек, Д. М. Пиза // Инженерное образование – М., 2007. – № 4. – С. 182–189.
19. *Шрайберг Я. А.* Библиотеки, компьютерные технологии и информационное общество: год прошедший и год грядущий // Науч. и техн. б-ки. – 2003. – № 1. – С. 29–59.
20. *Падерина О.* Новый подход к комплектованию в вузовской библиотеке, в аспекте внедрения новых автоматизированных технологий // Новая библиотека. – 2004. – № 1. – С. 24–26.
21. *Антопольский А. Б.* Информационные ресурсы России и политика их эффективного использования // Проблемы информатизации. – 1997. – № 4. – С. 4–9.

Надійшла 2.11.07

Цель исследования – определение основных исходных данных и направлений развития электронной библиотеки, технологическое определение принципов и источников комплектования для обеспечения создания электронной библиотеки.

The aim of investigation is determination of the main foundations and directions of development of electronic library, technological definition of principles and sources of acquisition for providing of creation of electronic library.

УДК 614.841

М. П. Мусиенко, В. И. Томенко

БЕСПРОВОДНЫЕ ИНФОРМАЦИОННО-КОММУНИКАЦИОННЫЕ И НАВИГАЦИОННЫЕ ТЕХНОЛОГИИ НА ТРАНСПОРТНЫХ СРЕДСТВАХ

В работе описано построение различных типов информационно-коммуникационных систем на базе беспроводных технологий, которые могут быть использованы на транспортных средствах. Рассмотрены навигационные и информационные каналы передачи данных, виды топологий сетей. Приведен пример аппаратной и программной реализации навигационного GPS комплекса с передачей данных по GSM каналу связи.

© Мусиенко М. П., Томенко В. И., 2007

ВВЕДЕНИЕ

В различных передвижных системах (транспортных средствах) важным аспектом является информация о местоположении транспортного средства, а также информационное обеспечение между различными объектами и диспетчерским (информационным, связующим)

центром. В качестве мониторинговой информации могут выступать координаты местоположения объекта и соответствующие им различные параметры (экологические, технические и др.) среды и других исследуемых объектов, видеоизображения и многое другое. В качестве предоставляемой информации на транспортном средстве могут быть различного рода оперативные сообщения (на автомобилях экстренных служб), реклама и объявления остановок (в городском транспорте) и многое другое.

Важным элементом в таких системах является выбор типа передаваемого канала, способа передачи информации, а также построение информационной сети взаимодействия различных субъектов между собой. Используемая на сегодняшний день коротковолновая радиосвязь имеет ряд существенных недостатков: сравнительно малые дальность действия и количество абонентов в сети, отсутствие автоматизации в передаче информации и др. Кроме того, такой тип связи не позволяет передать видеoinформацию и ряд других информационных параметров неголосовой типа.

Последнее время характеризуется бурным развитием различных информационных технологий, в частности – коммуникационных беспроводных технологий передачи информации. Новые типы связи обладают значительно расширенными функциональными возможностями и являются очень перспективными для применения на транспортных средствах.

ПОСТАНОВКА ЗАДАЧИ

Целью работы является описание построения различных типов информационно-коммуникационных систем на базе беспроводных технологий, которые могут быть использованы на транспортных средствах. Для этого рассмотрены различные типы используемых навигационных и информационных каналов передачи данных, виды топологий сетей, приведен пример аппаратной и программной реализации навигационного GPS комплекса совместно с каналами мобильных операторов связи.

ПРИМЕНЕНИЕ НАВИГАЦИОННОЙ СИСТЕМЫ

В настоящее время все чаще находит свое применение технология NAVSTAR GPS (англ. Navigation Satellite Time and Ranging, Global Positioning System – измерение дальности и времени по навигационному спутнику, глобальная система позиционирования) – спутниковая система навигации, часто сокращенно называемая технология GPS. Эта система позволяет в любой точке Земли почти при любой погоде, а также в космическом пространстве на расстоянии до 100 тыс. км от поверхности Земли определить местоположение (с точностью до нескольких метров) и скорость объектов. Основной сегмент применения GPS-технологии – мониторинг и навигация транспортных средств. GPS навигация дает возможность узнать местоположение объекта, направление и скорость движения, пройденное и оставшееся расстояние, время в пути. При привязке к карте появляется возможность проложить кратчайший маршрут к цели следования, запомнить расположение нужных объектов и т. д. [1].

GPS приемник, установленный на транспортном средстве, регистрирует параметры движения и передает их через интерфейс (например RS-232, USB и др.) по текстовому протоколу NMEA 0183 [2].

После включения приемника через определенное время (от 40 сек. до 15 мин. – в зависимости от последнего времени включения, количества определенных спутников и др.) он каждую секунду начинает посылать на порт компьютера полезный сигнал, состоящий из идентификаторов сообщений: GPGLL, GPGLL, GPGSA, GPGSV, GPRMC, GPVTG, GPMSS, GPZDA, которые несут информацию о времени, координатах на местности, скорости, направлении движения, параметрах спутников, с которых получены сигналы. На рис. 1 приведена строка с расшифровкой одного из сообщений (GPRMC идентификатора) приемника (сигнал от объекта, который находился в г. Черкассы).

Соответствующее аппаратное и программное обеспечение дает возможность определить кратчайший маршрут и время приезда к пункту назначения. Кроме того, имея информацию о типе дорог, заторах и ремонтах (эта информация, например, может поступать

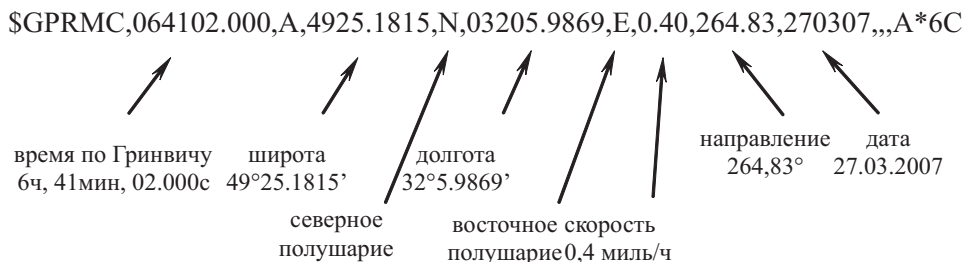


Рисунок 1 – Расшифровка GPRMC идентификатора в сигнале GPS приемника

с WEB-камер, установленных на участках дорог) программа может вводить соответствующие корректировки в движение транспортного средства.

Одной из задач при работе GPS навигации является выполнение какой-либо команды при нахождении в определенной точке местности (например, объявление остановки или рекламы в городском транспорте, отправка видеозображения местности и др.). Алгоритм работы такой программы показан на рис. 2.

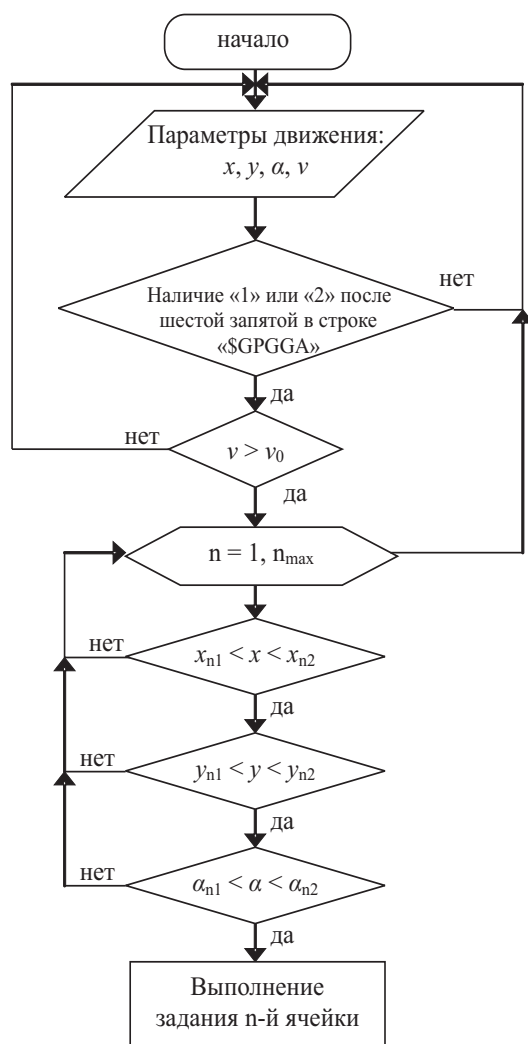


Рисунок 2 – Блок-схема алгоритма работы навигационной программы

Программа из сигналов GPS приемника выделяет координаты движения (x и y), скорость (v) и направление движения (α). Эти параметры сравниваются с данными, занесенными в базу (пример приведен в табл. 1), которые представляют собой значения направления движения (курсы), а также координаты нижнего левого (x_1 и y_1) и верхнего правого (x_2 и y_2) углов квадрата, являющегося территориальной зоной, при заезде в которую только с определенной стороны (для чего сравниваются курсы) необходимо выполнить соответствующее действие (например, для случая объявления остановок – запустить аудио и/или видео-файл).

Некоторые GPS приемники при остановке транспортного средства выдают хаотическое направление движения курса. Для избежания ложного срабатывания, программа устанавливает нижний предел скорости (v_0), при котором происходит опрос значений координат.

ВЫБОР ТИПА БЕСПРОВОДНОЙ СВЯЗИ

Одним из центральных вопросов построения информационно-коммуникационной системы является выбор типа беспроводной технологии связи. Основными критериями выбора являются:

- дальность действия;
- количество абонентов, которые могут находиться в сети;
- скорость и объем передаваемой информации и др.

На сегодняшний день на рынке беспроводных технологий предложено множество решений: Wi-Fi, Bluetooth, ZigBee, Wi-Max, Wi-pro, Mi-Wi и многие другие [3]. Сравнительная характеристика первых трех (все в стандарте 2,4 ГГц), получивших в настоящее время наибольшее распространение, показана в табл. 2. Следует отметить, что приведенные в табл. 2 значения являются усредненными и приведены для наиболее встречающихся стандартов. К примеру, если дальность действия большинства стандартов Bluetooth находится в диапазоне 10–100 метров, то на рынок начинает выходить стандарт с радиусом действия до 300 метров и т. д.

Из таблицы видно, что при наименьшем энергопотреблении стандарт Zig-Bee позволяет создавать сети

Таблица 1 – Параметры для реализации алгоритма GPS приемника работы программы

n	Параметры движения			Примеры заданий для выполнения
	x_1, x_2	y_1, y_2	α_1, α_2	
1	4925.1805 4925.1825	03205.9849 03205.9879	90 270	Объявить остановку (C:\Мои документы\file9.wav)
2	4925.17358 4925.17435	03205.9805 03205.9899	30 210	Передать в информационный центр координаты места (C:\run\place.exe)
3

Таблиця 2 – Сравнительная характеристика беспроводных технологий

Параметр	Тип беспроводной технологии		
	ZigBee (IEEE 802.15.4)	Bluetooth (IEEE 802.15.1)	Wi-Fi (IEEE 802.11.b)
Пропускная способность, кбит/с	250	723,1	11000
Время непрерывной автономной работы от батареи, дни	100–1000	1–10	0,5–5
Максимальное количество узлов в сети	65536	7	10
Дальность действия, м	10–100	10–100	20–300

размером до 65536 узлов, которые могут быть сопряжены как с датчиками, так и с управляющими устройствами и исполнительными механизмами, при этом стандарт предоставляет возможность создания сети сотовой архитектуры, что позволит с легкостью покрыть всю территорию объекта. Wi-Fi имеет слишком высокое энергопотребление, но может передавать наибольшие объемы информации. И так далее.

Альтернативой технологиям диапазона 2,4 ГГц является использование диапазонов 433 и 868 МГц. При потере скорости в передаче данных (что порой не является существенным требованием) эти технологии позволяют увеличить дальность действия до нескольких километров, сигналы этих диапазонов меньше поглощаются средой, чем сигналы с меньшей длиной волны. Все это создает благоприятные предпосылки разработки и использования беспроводных систем коммуникации на базе устройств, работающих в диапазонах 433 и 868 МГц [4].

Другими часто используемым информационными каналами являются каналы связи GSM-сети мобильных операторов. Для передачи данных используют CSP, GPRS, SMS и MMS сервисы. Выбор типа связи определяется в соответствии со стоимостью трафика, временем и объемом передаваемой информации.

ВЫБОР ТИПА ТОПОЛОГИИ СЕТИ

Одним из важнейших вопросов является выбор вида сети связи нескольких транспортных сетей между собой и с информационным центром. Наибольшее рас-

пространение получили сети типа «цепь», «звезда» (рис. 3, а), «сота» (рис. 3, б), «дерево», (рис. 3, в).

В сети типа «звезда» набор датчиков (D) передает сообщения в центральный приемник (S). Такие схемы используются для реализации централизованных систем сбора данных на единый диспетчерский пункт. Часто такие системы не имеют обратной связи (поскольку ради их удешевления датчики оснащаются не трансиверами, а передатчиками) [3].

Преимуществом сотовой архитектуры (рис. 3, б) является то, что приемник в каком-либо узле может обходить центральную станцию и передавать данные только в те узлы, которым эти данные нужны. Таким образом, за счет снижения передачи ненужной информации уменьшается трафик. Кроме того, такая сеть имеет возможность расширения и способность к самостоятельному определению и устранению неполадок, автоматическую перестройку маршрута прохождения трафика (при наличии программной поддержки этой опции в протокольной части используемого стандарта) [3].

В сети, показанной на рис. 3, в предусматривается наличие трех типов узлов. Узлы-приемники передают информацию на пункт сбора данных – либо непосредственно, либо через цепочку узлов-ретрансляторов, в зависимости от удаления от пункта сбора и условий прохождения сигнала. Возможны смешанные типы узлов, например, узел-приемник может исполнять и функцию ретранслятора [5].

Существуют и другие, непопулярные широкого распространения, но тем не менее удобные для примене-

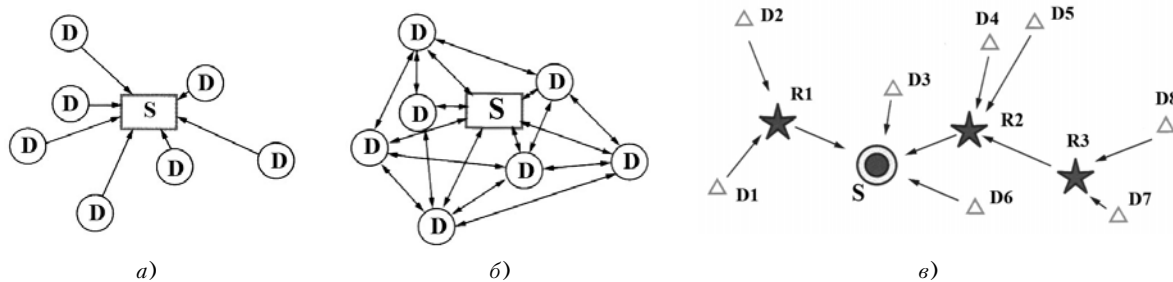


Рисунок 3 – Вид сетевой архитектуры:

а – «звезда», б – «сота»; в – «дерево» (смешанная); S – пункты сбора данных, R – маршрутизаторы или ретрансляторы, D – датчики (приемники)

ния в ряде случаев, сети. К примеру, при использовании сети для обмена данными между объектами, вытянутыми вдоль одной линии, удобно применять топологию «BackBone». В данном случае появляется возможность сообщений между устройств на расстояния до 4 километров [4].

ПРИМЕР АППАРАТНОЙ РЕАЛИЗАЦИИ

В качестве примера реализации описанных технологий был разработан навигационный программно-аппаратный GPS/GSM/GPRS комплекс с передачей данных по каналу GSM связи, позволяющий (в качестве одного из применений) объявлять остановки и предоставлять рекламу в городском транспорте. Устройство позволяло при въезде в определенную зону объявлять соответствующую остановку или прокручивать аудио/видео ролик. При этом обновление информации осуществлялось через GPRS канал с помощью GSM модема.

В разработанном изделии был использован SIM508 – трехдиапазонный 900/1800/1900 MHz GSM/GPRS/GPS модуль. Программное обеспечение написано с использованием программ Java (навигационная программа) и Delphi (программа обновления информации через GSM канал). Операционная система – Windows XP.

ВЫВОДЫ

Являясь одним из самых быстрорастущих сегментов рынка, беспроводные информационно-коммуникационные и навигационные технологии являются весьма перспективными для применения на транспортных средствах различного назначения. Однозначно отдать пред-

почтение той или иной технологии беспроводной связи, а также выбора типа используемой топологии сети невозможно. Выбор должен основываться на комплексном анализе нескольких параметров и учитывать объем, скорость, стоимость и надежность передачи информационных данных.

ПЕРЕЧЕНЬ ССЫЛОК

1. Мусиенко М. П., Томенко В. И., Савчук О. Л., Рудь М. П. Разработка навигационных программно-аппаратных GPS/GPRS комплексов на движущихся объектах // Вісник Черкаського державного технологічного університету. – 2007. – № 1.
2. GPS Engine Board. User's Guide. Holux Technology Inc.: www.holux.com.tw.
3. Байчаров С. Выбор технологии беспроводного обмена данными // Беспроводные технологии. – 2007. – № 2. – С. 59–61.
4. Кривченко Т. Радиомодули и радиомодемы компании OneRF // Беспроводные технологии. – 2007. – № 2. – С. 38–40.
5. Жиганов Е, Краснов С., Мошевикин А. Исследование условий применимости приемопередатчиков // Беспроводные технологии. – 2007. – № 1. – С. 65–68.

Надійшла 7.12.07

У роботі описана побудова різних типів інформаційно-комунікаційних систем на базі бездротових технологій, які можуть бути використані на транспортних засобах. Розглянуто навігаційні й інформаційні канали передачі даних, види топологій мереж. Наведено приклад апаратної й програмної реалізації навігаційного GPS комплексу з передачею даних по GSM каналі зв'язку.

The construction of different types of the informative-communication systems on the base of off-wire technologies which can be used at the transport vehicles, is described in the article. The navigative and informative ductings of data transferring, types of topology networks are considered. The example of hardware and programm realization of the navigative GPS complex with data transferring by GSM.

ТЕОРІЯ І МЕТОДИ АВТОМАТИЧНОГО УПРАВЛІННЯ

ТЕОРИЯ И МЕТОДЫ АВТОМАТИЧЕСКОГО УПРАВЛЕНИЯ

THEORY AND METHODS OF AUTOMATIC CONTROL

УДК 62-55:681.515

В. И. Гостев, С. Н. Скуртов, И. В. Панченко

ОПРЕДЕЛЕНИЕ УПРАВЛЯЮЩИХ ВОЗДЕЙСТВИЙ НА ВЫХОДЕ НЕЧЕТКОГО РЕГУЛЯТОРА ПРИ ИДЕНТИЧНЫХ ТРЕУГОЛЬНЫХ ФУНКЦИЯХ ПРИНАДЛЕЖНОСТИ С ОГРАНИЧЕНИЕМ

Получены аналитические выражения для управляющих воздействий на выходе нечеткого регулятора при идентичных треугольных функциях принадлежности с ограничением.

ВВЕДЕНИЕ

При синтезе нечетких регуляторов в системах автоматического управления наиболее часто используются треугольные функции принадлежности (ФП) для лингвистических величин. При расчете управляющих воздействий на выходе нечеткого регулятора абсциссу «центра тяжести результирующей фигуры» определяют обычно приближенным методом численного интегрирования. В данной работе получены аналитические выражения для управляющих воздействий на выходе нечеткого регулятора при идентичных треугольных функциях принадлежности с ограничением.

РЕШЕНИЕ ЗАДАЧИ

В данной работе рассмотрим нечеткий регулятор, структурная схема которого подробно описана в [1].

© Гостев В. И., Скуртов С. Н., Панченко И. В., 2007

На вход регулятора поступают ошибка системы θ , скорость изменения (первая производная) ошибки $\dot{\theta}$, ускорение (вторая производная) ошибки $\ddot{\theta}$.

На универсальном множестве $U = [0, 1]$ заданы два нечетких подмножества, функции принадлежности (ФП) которых для каждой лингвистической величины определяются по формулам:

$$\mu_1(u) = \begin{cases} 1, & 0 \leq u \leq a, \\ \frac{1-u}{1-a}, & a \leq u \leq 1; \end{cases}$$

$$\mu_2(u) = \begin{cases} \frac{u}{1-a}, & 0 \leq u \leq 1-a, \\ 1, & 1-a \leq u \leq 1. \end{cases} \quad (1)$$

При поступлении на нечеткий регулятор в какой-то момент времени значений входных переменных θ^* , $\dot{\theta}^*$ и $\ddot{\theta}^*$ с шагом квантования h осуществляется пересчет входных переменных в переменные u_1^* , u_2^* , u_3^* на универсальном множестве $U = [0, 1]$ и расчет значений

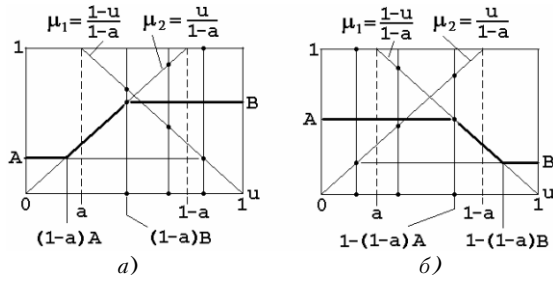


Рисунок 1

ФП для этих переменных (см. рис. 1). Точками на универсальном множестве отмечены возможные для какого-то момента времени значения переменных u_1^* , u_2^* , u_3^* .

Для упрощения нормировки (пересчета значений сигналов в значения элементов единого универсального множества) диапазоны изменения входных и выходного сигналов (параметров нечеткого регулятора) принимаем симметричными:

$$\begin{aligned} \theta_{\max} &= -\theta_{\min}; & \dot{\theta}_{\max} &= -\dot{\theta}_{\min}; & \ddot{\theta}_{\max} &= -\ddot{\theta}_{\min}; \\ m_{\max} &= -m_{\min}. \end{aligned}$$

Тогда формулы для нормировки (пересчета) принимают вид

$$\left. \begin{aligned} u_1^* &= -(\theta^* - \theta_{\min}) / (2\theta_{\min}); \\ u_2^* &= -(\dot{\theta}^* - \dot{\theta}_{\min}) / (2\dot{\theta}_{\min}); \\ u_3^* &= -(\ddot{\theta}^* - \ddot{\theta}_{\min}) / (2\ddot{\theta}_{\min}). \end{aligned} \right\} \quad (2)$$

Лингвистическое правило управления нечеткого регулятора формулируется в виде:

$$\begin{aligned} \text{Если } (\theta^* = a_1^j) \text{ и } (\dot{\theta}^* = a_2^j) \text{ и } (\ddot{\theta}^* = a_3^j), \\ \text{то } (m^* = a_c^j), \quad j = \overline{1, 2}, \end{aligned} \quad (3)$$

где a_1^j , a_2^j и a_3^j – лингвистические оценки ошибки, первой производной ошибки и второй производной ошибки, рассматриваемые как нечеткие терм-множества, определенные на универсальном множестве, $j = \overline{1, 2}$; a_c^j – лингвистические оценки управляющего воздействия на объект, выбираемые из терм-множества переменной m . Лингвистические оценки выбираются из терм-множеств лингвистических переменных θ^* , $\dot{\theta}^*$, $\ddot{\theta}^*$ и m^* :

$$a_i^j \in \{\text{отрицательная (1), положительная (2)}\}.$$

Другими словами, все сигналы (определенные выше лингвистические переменные) характеризуются как отрицательные ($j = 1$) или положительные ($j = 2$).

Функция принадлежности управляющего воздействия нечеткому множеству «отрицательная» определяется из системы нечетких логических уравнений (алгоритм Мамдани):

$$\mu^{m_1}(u_1, u_2, u_3) = \mu_1(u_1) \wedge \mu_1(u_2) \wedge \mu_1(u_3). \quad (4)$$

Функция принадлежности управляющего воздействия нечеткому множеству «положительная» определяется из системы нечетких логических уравнений:

$$\mu^{m_2}(u_1, u_2, u_3) = \mu_2(u_1) \wedge \mu_2(u_2) \wedge \mu_2(u_3). \quad (5)$$

Результирующая функция принадлежности для управляющего воздействия в соответствии с рабочим правилом нечеткого регулятора записывается в виде

$$\mu^m(u_1, u_2, u_3) = \mu^{m_1}(u_1, u_2, u_3) \vee \mu^{m_2}(u_1, u_2, u_3). \quad (6)$$

В выражениях (4)–(6) \wedge – логическое *и*, \vee – логическое *или*.

В соответствии с лингвистическими правилами управления, функция принадлежности управляющего воздействия $\mu_{1c}(u)$ нечеткому множеству «отрицательный» ограничена сверху значением:

$$A = \min[\mu_1(u_1^*), \mu_1(u_2^*), \mu_1(u_3^*)], \quad (7)$$

функция принадлежности управляющего воздействия $\mu_{2c}(u)$ нечеткому множеству «положительный» ограничена сверху значением:

$$B = \min[\mu_2(u_1^*), \mu_2(u_2^*), \mu_2(u_3^*)]. \quad (8)$$

Результирующая функция принадлежности для управляющего воздействия на основании выражения (6) определяется как

$$\mu_c(u) = \mu_{1c}(u) \vee \mu_{2c}(u), \quad (9)$$

т. е. получается формированием максимума

$$\mu_c(u) = \max[\mu_{1c}(u), \mu_{2c}(u)]. \quad (10)$$

Для определения конкретного значения управляющего воздействия m^* формируется «результирующая фигура», ограниченная результирующей ФП и производится поиск абсциссы «центра тяжести результирующей фигуры» u_c .

Отметим весьма существенный факт. Какие бы значения не принимали переменные u_1^* , u_2^* , u_3^* на универсальном множестве $U = [0, 1]$ в зависимости от соотношений величин A и B «результирующая фигура» может принимать только две конфигурации: при $A \leq B$ первая конфигурация показана на рис. 1, а; при $A \geq B$ вторая конфигурация показана на рис. 1, б.

Общая формула для определения абсциссы «центра тяжести результирующей фигуры» записывается в виде

$$u_c = \frac{\int_0^1 u \mu(u) du}{\int_0^1 \mu(u) du}. \quad (11)$$

Абсцисса «центра тяжести результирующей фигуры» при $A \leq B$ определяется по формуле

$$u_c = \frac{A \int_0^{(1-a)A} u du + \int_{(1-a)A}^{(1-a)B} \frac{u}{1-a} u du + B \int_{(1-a)B}^1 u du}{A \int_0^{(1-a)A} du + \int_{(1-a)A}^{(1-a)B} \frac{u}{1-a} du + B \int_{(1-a)B}^1 du} \quad \text{при } A \leq B. \quad (12)$$

После несложных вычислений находим:

$$u_c = \frac{B/2 + (1-a)^2(A^3 - B^3)/6}{B + (1-a)(A^2 - B^2)/2} \quad \text{при } A \leq B. \quad (13)$$

Абсцисса «центра тяжести результирующей фигуры» при $A \geq B$ определяется по формуле

$$u_c = \frac{\int_0^{1-(1-a)A} u du + \int_{1-(1-a)A}^{1-(1-a)B} \frac{1-u}{1-a} u du + B \int_{1-(1-a)B}^1 u du}{A \int_0^{1-(1-a)A} du + \int_{1-(1-a)A}^{1-(1-a)B} \frac{1-u}{1-a} du + B \int_{1-(1-a)B}^1 du} \quad \text{при } A \geq B. \quad (14)$$

После несложных вычислений находим:

$$u_c = \frac{A/2 - (1-a)(A^2 - B^2)/2 + (1-a)^2(A^3 - B^3)/6}{A - (1-a)(A^2 - B^2)/2} \quad \text{при } A \geq B. \quad (15)$$

Полученное значение u_c затем преобразуется в значение управляющего воздействия на объект управления

$$m^* = m_{\min}(1 - 2u_c).$$

В качестве примера приведем следующие результаты расчетов.

При $A = 0,1, B = 0,3, a = 0,25$ получаем $u_c = 0,6082$.

При $A = 0,3, B = 0,1, a = 0,25$ получаем $u_c = 0,3918$.

ВЫВОД

Полученные формулы позволяют использовать точный метод вычисления абсциссы «центра тяжести результирующей фигуры» при идентичных треугольных функциях принадлежности с ограничением и дают возможность упростить алгоритм расчета управляющих воздействий на выходе нечеткого регулятора. Изложенную методику можно также использовать для расчета управляющих воздействий на выходе нечеткого регулятора при функциях принадлежности другого вида.

ПЕРЕЧЕНЬ ССЫЛОК

1. Гостев В. И. Синтез нечетких регуляторов систем автоматического управления. – К.: Издательство «Радиоаматор», 2005. – 708 с.

Надійшла 5.06.07

Отримано аналітичні вираження для керуючих впливів на виході нечіткого регулятора при ідентичних трикутних функціях приналежності з обмеженням.

Analytical expressions for controlling actions on an output of an fuzzy controller are received at identical triangular membership functions with restriction.

УДК 681.5.01.23

Т. Б. Никитина

СИНТЕЗ ЦИФРОВОГО НЕЛИНЕЙНОГО РОБАСТНОГО УПРАВЛЕНИЯ МНОГОКАНАЛЬНЫМИ СИСТЕМАМИ

Разработан метод синтеза нелинейного робастного управления многоканальными системами, работающими по принципу грубого и точного управления в соответствии с итерационным алгоритмом. Приведен пример динамических характеристик синтезированной двухканальной электромеханической системы.

ВВЕДЕНИЕ

Многоканальные системы, работающие по принципу грубого и точного управления, позволяют существенно повысить точность управления, а в ряде случаев обес-

печить точность, недостижимую в одноканальных системах [1–4]. Такие системы обычно применяются в тех случаях, когда с помощью одноканальных систем принципиально невозможно обеспечить требуемую точность управления. Это, в частности, касается систем управления большими антеннами и радиотелескопами, тяжелыми станками с ЧПУ, роботами и манипуляторами, прокатными станами, системами частотной и фазовой автоподстройки генераторов, оптических дисковых систем записи информации, источников питания с высоким качеством выходной энергии и многих других. Так, например, в тяжелых металлорежущих станках находит применение схема двухдвигательного электропривода, суммирование силовых воздействий от которых осуществляется через механический дифференциал. Особенно эффективно совместное использование электромеханических и гидравлических приводов, когда реализация уточняющего движения может осуществляться за счет управляющего зазора в паре винт – гайка и т. д. Третий точный канал может быть реализован в виде магнестрикционного либо пьезострикционного двигателя, жестко закрепленного на подвижной каретке, управляемой основными силовыми приводами.

ПОСТАНОВКА ПРОБЛЕМЫ, СВЯЗЬ С НАУЧНЫМИ И ПРАКТИЧЕСКИМИ ЗАДАЧАМИ

Повышение точности многоканальных систем сдерживается наличием нелинейных и упругих элементов, что проявляется в неплавном движении рабочего органа, сопровождающегося остановками и рывками при движении рабочего органа на нижнем пределе рабочих скоростей. Рассмотрим синтез многоканальных систем цифрового нелинейного робастного управления с учетом нелинейных и упругих элементов. При этом рассматриваются аналитические нелинейности, которые могут быть аппроксимированы аналитическими выражениями в виде степенных рядов по векторам состояния и управления.

АНАЛИЗ ПОСЛЕДНИХ ДОСТИЖЕНИЙ И ПУБЛИКАЦИЙ ПО ДАННОЙ ПРОБЛЕМЕ

В работах [1–2] рассмотрены вопросы синтеза многоканальных систем робастного управления при непрерывном управлении. Современные системы управления реализуются на цифровой элементной базе. В работах [3–4] рассмотрены вопросы синтеза многоканальных систем цифрового робастного управления при последовательном синтезе отдельных каналов.

ЦЕЛЬ РАБОТЫ

Целью данной работы является синтез цифрового нелинейного робастного управления многоканальными системами, позволяющего обеспечивать работоспособность системы при изменении моделей объектов управления и внешних воздействий и высокой точности многоканальной системы.

ИЗЛОЖЕНИЕ МАТЕРИАЛА ИССЛЕДОВАНИЯ, ПОЛУЧЕННЫХ НАУЧНЫХ РЕЗУЛЬТАТОВ

При работе многоканальных систем в реальных условиях отработки задающих и компенсации возмущающих воздействий, требующих движения исполнительных механизмов многоканальных систем на малых скоростях, нередко возникают скачкообразные движения, остановки, автоколебания и т. д., хотя при синтезе линеаризованной системы условия устойчивости заведомо выполнялись [5–6]. Это связано с тем, что при синтезе системы использовались линейные модели каналов и идеализированная характеристика модели трения в виде знаковой характеристики от скорости движения. Однако при движении на низких и сверхнизких – ползучих скоростях наблюдается падающий участок в характеристике внешнего трения так, что с увеличением скорости движения момент сопротивления уменьшается. Наличие такого падающего участка в характеристике внешнего трения характерно при движении на низких скоростях, как при вращательном, так и при поступательном движении. При увеличении скорости движения падающий участок переходит в горизонтальный.

Для многих механизмов работа на падающем участке может приводить к возникновению автоколебаний и такой режим является аварийным [7–9]. Для других механизмов такой режим является нормальным [10], хотя и приводит к повышенному износу. В системах стабилизации и промышленных следящих системах неплавное движение вызывает ухудшение качества выпускаемой продукции. Так, например, в процессе прокатки многоканальная система регулирования геометрических параметров проката определяет точность поддержания толщины, профиля и формы готового проката. Однако в режиме захвата полосы при заправочной скорости проскальзывание валков относительно прокатываемой полосы может привести к серьезным авариям.

Таким образом, при синтезе многоканальных систем необходимо при работе на малых скоростях учитывать наличие падающего участка в характеристике внешнего трения, что обуславливает положительную обратную связь в системе [11–13]. Однако при работе на больших скоростях целесообразно учитывать горизонтальный и восходящий участки в характеристике

внешнего трения, так как системы, синтезированные с учетом падающего участка в линейном приближении при работе с большой скоростью на горизонтальном и восходящем участках обладают излишней инерционностью. С этой целью аппроксимируем нелинейную зависимость момента сопротивления от скорости проскальзывания аналитическим выражением в виде степенного ряда от скорости проскальзывания и при этом исходная система становится нелинейной [2–3].

В настоящее время наиболее широкое распространение получило решение задачи нелинейного цифрового робастного управления во временной области. Нахождение нелинейного робастного управления сводится к решению уравнения Гамильтона – Якоби – Беллмана для нелинейных моделей задающего и возмущающего воздействий и модели объекта управления многоканальной системы. В данной работе используется итеративное решение уравнения Гамильтона – Якоби – Беллмана в соответствии с разложением аналитических нелинейностей исходной системы в степенной ряд по переменным состояния системы. В частности, для нелинейной характеристики внешнего трения с учетом падающего, горизонтального и восходящего участков с достаточной для практики точностью необходим третий порядок аппроксимирующего полинома.

Рассмотрим многоканальную систему, состоящую из m автономных каналов, каждый j -й канал которой описывается разностным уравнением состояния

$$\begin{aligned} \vec{x}_j(k+1) &= \Phi_j(\vec{x}_j(k), \vec{u}_j(k)), \\ \vec{y}_j(k) &= \varphi_j(\vec{x}_j(k), \vec{u}_j(k)), \end{aligned}$$

в которой векторные функции $\Phi_j(\vec{x}_j(k), \vec{u}_j(k))$, $\varphi_j(\vec{x}_j(k), \vec{u}_j(k))$ могут быть представлены в следующем виде:

$$\begin{aligned} \Phi_j(\vec{x}_j(k), \vec{u}_j(k)) &= A_j \vec{x}_j(k) + \\ &+ B_j \vec{U}_j(k) + \sum_{i=2}^n f_i(\vec{x}_j(k), \vec{u}_j(k)), \\ \varphi_j(\vec{x}_j(k), \vec{u}_j(k)) &= C_j \vec{x}_j(k) + \\ &+ D_j \vec{u}_j(k) + \sum_{i=2}^n h_i(\vec{x}_j(k), \vec{u}_j(k)), \end{aligned}$$

где символ i указывает порядок формы от векторов состояния $\vec{x}_j(k)$ и управления $\vec{u}_j(k)$.

Рассмотрим модель задающего воздействия j -го канала, которая описывается уравнением состояния

$$\begin{aligned} \vec{x}_{jr}(k+1) &= \Phi_{jr}(\vec{x}_{jr}(k)), \\ \vec{y}_{jr}(k) &= \varphi_{jr}(\vec{x}_{jr}(k)), \end{aligned}$$

в которой векторные функции $\Phi_{jr}(\vec{x}_{jr}(k))$, $\varphi_{jr}(\vec{x}_{jr}(k))$ могут быть представлены в следующем виде:

$$\begin{aligned} \Phi_{jr}(\vec{x}_{jr}(k)) &= A_{jr} \vec{x}_{jr}(k) + \sum_{i=2}^n f_i(\vec{x}_{jr}(k)), \\ \varphi_{jr}(\vec{x}_{jr}(k)) &= C_{jr} \vec{x}_{jr}(k) + \sum_{i=2}^n h_{ijr}(\vec{x}_{jr}(k)), \end{aligned}$$

где символ i указывает порядок формы от векторов состояния $\vec{x}_j(k)$ и управления $\vec{u}_j(k)$.

Рассмотрим модель задающего воздействия j -го канала, которая описывается уравнением состояния

$$\begin{aligned} \vec{x}_{jr}(k+1) &= \Phi_{jr}(\vec{x}_{jr}(k)), \\ \vec{y}_{jr}(k) &= \varphi_{jr}(\vec{x}_{jr}(k)), \end{aligned}$$

в которой векторные функции $\Phi_{jr}(\vec{x}_{jr}(k))$, $\varphi_{jr}(\vec{x}_{jr}(k))$ могут быть представлены в следующем виде

$$\begin{aligned} \Phi_{jr}(\vec{x}_{jr}(k)) &= A_{jr} \vec{x}_{jr}(k) + \sum_{i=2}^n f_i(\vec{x}_{jr}(k)), \\ \varphi_{jr}(\vec{x}_{jr}(k)) &= C_{jr} \vec{x}_{jr}(k) + \sum_{i=2}^n h_{ijr}(\vec{x}_{jr}(k)), \end{aligned}$$

где символ i указывает порядок формы от векторов состояния $\vec{x}_j(k)$ и управления $\vec{u}_j(k)$.

Рассмотрим также модель возмущающего воздействия j -го канала, которая описывается уравнением состояния

$$\begin{aligned} \vec{x}_{jd}(k+1) &= \Phi_{jd}(\vec{x}_{jd}(k)), \\ \vec{y}_{jd}(k) &= \varphi_{jd}(\vec{x}_{jd}(k)), \end{aligned}$$

в которой векторные функции $\Phi_{jd}(\vec{x}_{jd}(k))$, $\varphi_{jd}(\vec{x}_{jd}(k))$ могут быть представлены в следующем виде:

$$\begin{aligned} \Phi_{jd}(\vec{x}_j(k), \vec{u}_j(k)) &= A_j \vec{x}_j(k) + \\ &+ B_j \vec{U}_j(k) + \sum_{i=2}^n f_i(\vec{x}_j(k), \vec{u}_j(k)), \\ \varphi_{jd}(\vec{x}_{jd}(k)) &= C_{jd} \vec{x}_{jd}(k) + \sum_{i=2}^n h_{ijd}(\vec{x}_{jd}(k)), \end{aligned}$$

где символ i указывает порядок формы от вектора состояния $\vec{x}_{jd}(k)$.

Введем согласно [2] вектор $\vec{x}(k)$ состояния многоканальной системы, компонентами которого являются вектора $\vec{x}_j(k)$ состояния всех каналов, входящих в систему, а также вектора состояний $\vec{x}_{jr}(k)$, $\vec{x}_{jd}(k)$ задающих и возмущающих воздействий этих каналов и за-

пишем уравнения состояния многоканальной системы в следующем виде:

$$\begin{aligned}\vec{x}(k+1) &= \Phi(\vec{x}(k), \vec{u}(k)), \\ \vec{y}(k) &= \varphi(\vec{x}(k), \vec{u}(k)),\end{aligned}$$

в которой векторные функции $\Phi(\vec{x}(k), \vec{u}(k))$, $\varphi(\vec{x}(k), \vec{u}(k))$ могут быть представлены в следующем виде:

$$\begin{aligned}\Phi(\vec{x}(k), \vec{u}(k)) &= A\vec{x}(k) + B\vec{U}(k) + \sum_{i=2}^n f_i(\vec{x}(k), \vec{u}(k)), \\ \varphi(\vec{x}(k), \vec{u}(k)) &= C\vec{x}(k) + D\vec{u}(k) + \sum_{i=2}^n h_i(\vec{x}(k), \vec{u}(k)),\end{aligned}$$

где символ i указывает порядок формы от векторов состояния $\vec{x}(k)$ и управления $\vec{u}(k)$.

Рассмотрим задачу синтеза оптимального управления $\vec{u}(k)$, минимизирующего функционал

$$J = \sum_{k=1}^{\infty} \psi(\vec{x}(k), \vec{u}(k)).$$

Предполагается, что функция $\psi(\vec{x}(k), \vec{u}(k))$ является аналитической и разлагается в степенной ряд

$$\psi(\vec{x}(k), \vec{u}(k)) = \sum_{i=2}^{n+1} \psi_i(\vec{x}(k), \vec{u}(k)).$$

Будем искать оптимальное управление $\vec{u}(k)$ в форме обратных связей по полному вектору состояния

$$\vec{u}(k) = \sum_{i=1}^n \vec{u}_i(\vec{x}(k)).$$

Введем функцию Ляпунова

$$V(\vec{x}(k)) = \sum_{i=2}^{n+1} V_i(\vec{x}(k)).$$

Тогда на основании достаточного условия оптимальности минимум критерию обеспечивает оптимальное управление, образующее систему уравнений, эквивалентную уравнению Гамильтона – Якоби – Беллмана [5],

$$\Phi^T(\vec{x}(k), \vec{u}(k)) \frac{\partial V(\vec{x}(k))}{\partial \vec{x}(k)} + \psi(\vec{x}(k), \vec{u}(k)) = 0,$$

$$\frac{\partial \Phi^T(\vec{x}(k), \vec{u}(k))}{d\vec{u}(k)} \frac{\partial V(\vec{x}(k))}{d\vec{x}(k)} + \frac{\partial \psi(\vec{x}(k), \vec{u}(k))}{\partial \vec{u}(k)} = 0.$$

Решение многих задач оптимизации нелинейных систем приводит к необходимости решения уравнения

Гамильтона – Якоби [2]. Решение уравнения Гамильтона – Якоби в теории оптимизации нелинейных систем является таким же важным, как и решение уравнения Риккати в теории линейных оптимальных систем. В настоящее время к решению уравнения Гамильтона – Якоби проявляется повышенный интерес в связи с разработкой нелинейной теории H^∞ – нелинейного робастного управления. Это происходит аналогично тому, как решение линейной теории H^∞ – линейного робастного управления сводится к итеративной процедуре решения уравнения Риккати при заданном уровне толерантности. С практической точки зрения целесообразно рассматривать приближенное решение нелинейной теории H^∞ как приближенное решение уравнения Гамильтона – Якоби в виде итеративной процедуры. При этом робастное управление первым линеаризованным приближением решения уравнения Гамильтона – Якоби нелинейного робастного управления является решение уравнения Риккати в линейном робастном управлении. Такое представление имеет место для исходной нелинейной системы с аналитическими нелинейностями.

Ограничимся сначала в рядах членами низшего порядка относительно векторов состояния $\vec{x}_1(t)$ и управления $\vec{u}_1(t)$ первого канала. Тогда исходная система будет линейной, для которой может быть синтезировано линейное робастное управление. Запишем для исходной линейной дискретной системы разностное уравнение состояния, вектор контролируемых параметров $\vec{z}(k)$ и вектор измеряемых переменных $\vec{y}(k)$ с помощью Δ оператора в стандартной форме, принятой в H^∞ теории [14–15].

В этом случае уравнение состояния дискретной системы примет следующий вид:

$$\begin{aligned}\delta \vec{x}(k) &= A_\delta \vec{x}(k) + B_{1\delta} \vec{w}_1(k) + B_{2\delta} \vec{u}(k), \\ \vec{z}(k) &= C_1 \vec{x}(k) + D_{11} \vec{w}_1(k) + D_{12} \vec{u}(k), \\ \vec{y}(k) &= C_2 \vec{x}(k) + D_{21} \vec{w}_1(k) + D_{22} \vec{u}(k).\end{aligned}$$

Для нахождения цифрового робастного регулятора необходимо решить уравнение Риккати по управлению

$$\begin{aligned}0 &= \widehat{Q} + A^T X + XA + \Delta A^T X A - \left[\widehat{L} + B^T X(\Delta A + I) \right]^T \times \\ &\times \left[\widehat{R} + \Delta B^T X B \right]^{-1} \left[\widehat{L} + B^T X(\Delta A + I) \right],\end{aligned}$$

где

$$\begin{aligned}\widehat{R} + \Delta B^T X B &\stackrel{\Delta}{=} \widehat{R}_q / \Delta, \\ \widehat{L} + B^T X(\Delta A + I) &\stackrel{\Delta}{=} \widehat{L}_q / \Delta, \\ \widehat{Q} &\stackrel{\Delta}{=} \widehat{C}^T \widehat{J} \widehat{C} / \Delta.\end{aligned}$$

При этом замкнутой таким регулятором система

$$A - B(\widehat{R} + \Delta B^T X B)^{-1} (\widehat{L} + B X (\Delta A + I))$$

является устойчивой.

Для нахождения цифрового робастного наблюдателя необходимо решить уравнение Риккати по наблюдению

$$0 = \widehat{Q} + AZ + ZA^T + \Delta \widehat{A} Z \widehat{A}^T - [\widehat{L} + (\Delta \widehat{A} + I) Z \widehat{C}^T] \times \\ \times [\widehat{R} + \Delta \widehat{C} Z \widehat{C}^T]^{-1} [\widehat{L} + (\Delta \widehat{A} + I) Z \widehat{C}^T]^T,$$

где

$$\widehat{R} + \Delta C Z_{\infty} \widehat{C}^T \stackrel{\Delta}{=} \Delta \widehat{R}_q, \\ \widehat{L} + (\Delta \widehat{A} + I) Z_{\infty} \widehat{C}^T \stackrel{\Delta}{=} \widehat{L}_q, \\ Q \stackrel{\Delta}{=} \widehat{B} \widehat{J} \widehat{B}^T / \Delta.$$

При этом цифровой робастный наблюдатель может быть представлен в виде $A B C D$ реализации

$$\left[\begin{array}{c|c} \widehat{A} & \widehat{B} \\ \hline \widehat{C} & \widehat{D} \end{array} \right] = \left[\begin{array}{c|c|c} A - B_1 R_d^{-1} L_d & B_1 V_{21}^{-1} & 0 \\ \hline V_{12} R_3^{-1} (\widehat{L}_2 - \widehat{R}_2 R_d^{-1} L_d) & V_{12} \widehat{R}_3^{-1} V_{21}^{-1} & I \\ \hline C_2 - D_{21} R_d^{-1} L_d & D_{21} V_{21}^{-1} & 0 \end{array} \right],$$

где

$$R_d = R_1 - R_2^T R_3^{-1} R_2, \\ L_d = L_1 - R_2^T R_3^{-1} L_2, \\ V_{12}^T V_{12} = R_3, \\ V_{21}^T V_{21} = -\gamma^{-2} (R_1 - R_2^T R_3^{-1} R_2).$$

Естественно, что наблюдатель в замкнутом виде

$$\widehat{A} - [\widehat{L} + (\Delta \widehat{A} + I) Z \widehat{C}^T] (\widehat{R} + \Delta \widehat{C} Z \widehat{C}^T)^{-1} \widehat{C}$$

также является асимптотически устойчивой системой.

Тогда цифровой робастный регулятор и цифровой робастный наблюдатель представляют собой цифровой робастный компенсатор, входом которого является измеряемый вектор исходной системы $\vec{y}(k)$, а выходом является вектор управления исходной системы $\vec{u}(k)$. Робастный компенсатор с матрицами A_p, B_p, C_p, D_p описывается следующим уравнением состояния:

$$\vec{x}_p(k+1) = A_p \vec{x}_p(k) + B_p \vec{y}(k),$$

$$\vec{u}(k) = C_p \vec{x}_p(k) + D_p \vec{y}(k),$$

где

$$A_p = \widehat{A} - B_2 V_{12}^{-1} \widehat{C}_1 + B_2 V_{12}^{-1} \tilde{R}_2 \tilde{R}_3^{-1} \widehat{C}_2 - \tilde{L}_2 \tilde{R}_3^{-1} \widehat{C}_2, \\ B_p = -B_2 V_{12}^{-1} \tilde{R}_2 \tilde{R}_3^{-1} + \tilde{L}_2 \tilde{R}_3^{-1}, \\ C_p = -V_{12}^{-1} \widehat{C}_1 + V_{12}^{-1} \tilde{R}_2 \tilde{R}_3^{-1} \widehat{C}_2, \\ D_p = -V_{12}^{-1} \tilde{R}_2 \tilde{R}_3^{-1}.$$

В этих выражениях

$$\begin{bmatrix} \tilde{R}_1 & \tilde{R}_2 \\ \tilde{R}_2^T & \tilde{R}_3 \end{bmatrix} = \widehat{R} + \Delta \widehat{C} Z \widehat{C}^T, \\ \begin{bmatrix} \tilde{L}_1 & \tilde{L}_2 \end{bmatrix} = \widehat{L} + (\Delta \widehat{A} + I) Z \widehat{C}^T.$$

Решение уравнения Риккати по управлению может быть сведено к нахождению спектра матрицы Гамильтона по управлению

$$\overline{M} = \begin{bmatrix} A_m + \Delta B \overline{R}^{-1} B^T (I + \Delta A_m^T) Q_m & -B \overline{R}^{-1} B (I + \Delta A_m^T)^{-1} \\ -(I + \Delta A_m^T) Q_m & -(I + \Delta A_m^T) A_m^T \end{bmatrix},$$

где

$$A_m = A - B \overline{R}^{-1} \overline{L}, \quad Q_m = \overline{Q} - \overline{L}^T \overline{R}^{-1} \overline{L}.$$

Решение уравнения Риккати по наблюдению также может быть сведено к нахождению спектра матрицы Гамильтона по наблюдению

$$\widehat{M} = \begin{bmatrix} \widehat{A}_m + \Delta \widehat{C}^T \widehat{R}^{-1} \widehat{C} (I + \Delta \widehat{A}_m^T) \widehat{Q}_m & -\widehat{C}^T \widehat{R}^{-1} \widehat{C} (I + \Delta \widehat{A}_m^T)^{-1} \\ -(I + \Delta \widehat{A}_m^T)^{-1} \widehat{Q}_m & -(I + \Delta \widehat{A}_m^T) \widehat{A}_m^T \end{bmatrix},$$

где

$$\widehat{A}_m = \widehat{A} - \widehat{L} \widehat{R}^{-1} \widehat{C}, \quad \widehat{Q}_m = \widehat{Q} - \widehat{L} \widehat{R}^{-1} \widehat{L}^T.$$

Подставим синтезированное таким образом линейное управление и квадратичную функцию Ляпунова в выражение для исходного нелинейного оптимального управления и добавим пока с неизвестными коэффициентами члены второго порядка в оптимальное управление и члены третьего порядка $V_3(\vec{x}(k))$ в функцию Ляпунова. Добавим также в исходную нелинейную систему члены второго порядка $f_2(\vec{x}(k), \vec{u}(k))$ и $h_2(\vec{x}(k), \vec{u}(k))$ и добавим к интегральному квадратичному функционалу

члены третьего порядка $\psi_3(\vec{x}(k), \vec{u}(k))$. Откуда найдем исходные уравнения для неизвестных коэффициентов второго порядка $\vec{u}_2(\vec{x}(k))$ оптимального управления и третьего порядка $V_3(\vec{x}(k))$ функции Ляпунова. Аналогично находятся последующие приближения.

РЕЗУЛЬТАТЫ МОДЕЛИРОВАНИЯ

В качестве примера рассмотрим синтез робастных регуляторов двухканальной системы с раздельной нагрузкой [1], у которой каждый канал представляет собой трехмассовую электромеханическую систему, схема которой показана на рис. 1. Система предназначена

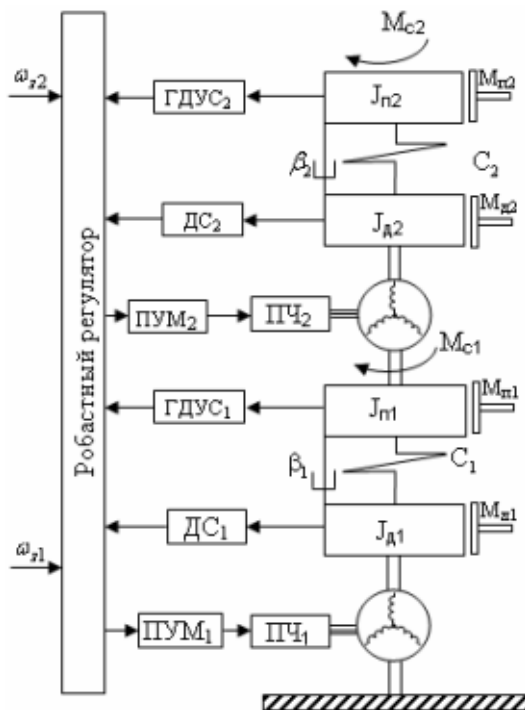


Рисунок 1 – Схема двухканальной системы с раздельной нагрузкой

для обработки заданных значений скоростей вращения ω_{z1}, ω_{z2} первой и второй платформ. На первую и вторую платформы действуют внешние моменты сопротивления $M_{с1}, M_{с2}$, которые являются возмущающими моментами.

Платформы приводятся во вращение с помощью асинхронных двигателей с короткозамкнутым ротором. В системе используется векторное управление асинхронными двигателями с помощью преобразователей частоты (ПЧ). Каждый канал содержит программно-

аппаратный контур прямого управления моментом (ПУМ) [4–7].

Скорости вращения платформ $\omega_{п1}, \omega_{п2}$, измеряются с помощью гироскопических датчиков угловых скоростей (ГДУС), а скорости вращения роторов двигателей $\omega_{д1}, \omega_{д2}$ измеряются с помощью датчиков скорости (ДС).

В системе учтены нелинейные характеристики моментов внутреннего трения на валах двигателей $M_{д1}, M_{д2}$ и платформ $M_{п1}, M_{п2}$.

Составим уравнения динамики основных элементов двухканальной системы с раздельной нагрузкой следя работе [3].

Уравнение равновесия моментов для платформы второго канала с моментом инерции $J_{п2}$

$$J_{п2} \frac{d\omega_{п2}}{dt} = M_{\Sigma 2} - M_{с2} - M_{п2}(\omega_{п2}),$$

где $M_{с2}$ – момент внешнего сопротивления, действующий на вторую платформу; $M_{\Sigma 2}$ – суммарный момент, приведенный ко второй платформе

$$M_{\Sigma 2} = M_{y2} + \beta_2(\omega_{д2} - \omega_{п2}),$$

где $\omega_{д2}, \omega_{п2}$ – скорости вращения соответственно ротора второго двигателя и второй платформы; β_2 – коэффициент вязкого трения трансмиссии; M_{y2} – упругий момент трансмиссии второй платформы, уравнение состояния которого примет следующий вид:

$$\frac{dM_{y2}}{dt} = C_2(\omega_{д2} - \omega_{п2}),$$

где C_2 – жесткость вала, соединяющего вторую платформу с двигателем.

Уравнение равновесия моментов для ротора второго двигателя с моментом инерции $J_{д2}$

$$J_{д2} \frac{d\omega_{д2}}{dt} = M_{л2} - M_{\Sigma 2} - M_{л2}(\omega_{д2}).$$

Уравнение равновесия моментов первой платформы с моментом инерции $J_{п1}$

$$J_{п1} \frac{d\omega_{п1}}{dt} = M_{\Sigma 1} - M_{с1} - M_{д2} - M_{п1}(\omega_{п1}),$$

где $M_{с1}$ – момент внешнего сопротивления, действующий на первую платформу; $M_{\Sigma 1}$ – суммарный момент, приведенный к первой платформе,

$$M_{\Sigma 1} = M_{y1} + \beta_1(\omega_{д1} - \omega_{п1}),$$

где $\omega_{д1}$, $\omega_{п1}$ – скорости вращения соответственно ротора первого двигателя и первой платформы; β_1 – коэффициент вязкого трения трансмиссии; M_{y1} – упругий момент трансмиссии первой платформы, уравнение состояния которого примет следующий вид:

$$\frac{dM_{y1}}{dt} = C_1(\omega_{д1} - \omega_{п1}),$$

где C_1 – жесткость вала, соединяющего первую платформу с двигателем.

Уравнение равновесия моментов для ротора первого двигателя с моментом инерции $J_{д1}$

$$J_{д1} \frac{d\omega_{д1}}{dt} = M_{д1} - M_{\Sigma 1} - M_{д1}(\omega_{д1}).$$

Нелинейные характеристики моментов внутреннего трения на валах двигателей $M_{д1}$, $M_{д2}$ и платформ $M_{п1}$, $M_{п2}$ являются нелинейными функциями скоростей движения соответствующих подвижных частей объекта управления и могут быть аппроксимированы степенными рядами:

$$M_c(\omega_{ck}) = \sum_{i=0}^n \beta_{ci} \omega_{ck}^i.$$

Для практического использования обычно достаточно третьего порядка этого степенного ряда. При этом обычно линейный член имеет отрицательное значение, квадратичный член имеет практически нулевое значение, а кубический член имеет положительное значение.

Для построения математической модели двухканальной системы с отдельной нагрузкой при цифровом управлении рассмотрим вспомогательную непрерывную динамическую систему.

Введем вектор состояния этой системы, компонентами которой являются скорость вращения первого двигателя $\omega_{д1}$, скорость вращения первой платформы $\omega_{п1}$, момент упругости первой платформы M_{y1} , скорость вращения второго двигателя $\omega_{д2}$, скорость вращения второй платформы $\omega_{п2}$, момент упругости второй платформы M_{y2} , так что вектор состояния примет следующий вид:

$$\vec{X}(t) = \{\omega_{д1}, \omega_{п1}, M_{y1}, \omega_{д2}, \omega_{п2}, M_{y2}\}^T.$$

Введем вектор внешних воздействий $\vec{U}(t)$, компонентами которого являются моменты первого и второго двигателей $M_{д1}$, $M_{д2}$, моменты внешних сопротивлений, действующих на первую и вторую платформы M_{c1} и M_{c2} , так что вектор внешних воздействий примет следующий вид:

$$\vec{U}(t) = \{M_{д1}, M_{д2}, M_{c1}, M_{c2}\}^T.$$

Тогда матрицы состояния и управления линеаризованной системы примут следующий вид:

$$A = \begin{bmatrix} \frac{-\beta_1 - \beta_{д1}}{J_{д1}} & \frac{\beta_1}{J_{д1}} & -\frac{1}{J_{д1}} & & & \\ \frac{\beta_1}{J_{п1}} & \frac{-\beta_1 - \beta_{п1}}{J_{п1}} & \frac{1}{J_{п1}} & & & \\ C_1 & -C_1 & & & & \\ & & & \frac{-\beta_2 - \beta_{д2}}{J_{д2}} & \frac{\beta_2}{J_{д2}} & -\frac{1}{J_{д2}} \\ & & & \frac{\beta_2}{J_{п2}} & \frac{-\beta_2 - \beta_{п2}}{J_{п2}} & \frac{1}{J_{п2}} \\ & & & C_2 & -C_2 & \end{bmatrix};$$

$$B = \begin{bmatrix} \frac{1}{J_{д1}} & & & & & \\ & \frac{1}{J_{п1}} & -\frac{1}{J_{п1}} & & & \\ & & & & & \\ & \frac{1}{J_{д2}} & & & & \\ & & & & -\frac{1}{J_{п2}} & \\ & & & & & \end{bmatrix}$$

В матрице состояния линеаризованной системы учтены лишь линейные составляющие нелинейных характеристик моментов внутреннего трения на валах двигателей $M_{д1}$, $M_{д2}$ и платформ $M_{п1}$, $M_{п2}$ в виде соответствующих линейных коэффициентов $\beta_{д1}$, $\beta_{д2}$, $\beta_{п1}$, $\beta_{п2}$.

Для этого непрерывного объекта управления построим его дискретный аналог

$$\vec{x}_д(k+1) = A_д \vec{x}_д(k) + B_д \vec{u}(k),$$

$$\text{где } A_д = I + \frac{A_0 \Delta}{1!} + \frac{A_0^2 \Delta^2}{2!} + \dots + \frac{A_0^n \Delta^n}{n!}.$$

Для полученной исходной дискретной системы представим матрицы состояния $A_д$ и управления $B_д$ этой вспомогательной системы в виде следующих блочных матриц:

$$A_д = \begin{bmatrix} A_{11} & & & \\ & & & \\ & & & \\ & & & A_{22} \end{bmatrix};$$

$$B_д = \begin{bmatrix} B_{11} & B_{12} & B_{13} & \\ & B_{22} & & B_{24} \end{bmatrix}.$$

При векторном управлении асинхронным двигателем с короткозамкнутым ротором обычно реализуется

го напряжения первого интегратора z_1 , помеха f_{ω_2} измерения скорости вращения $\omega_{д2}$ второго двигателя, помеха f_{z2} измерения выходного напряжения второго интегратора z_2 , так что

$$\vec{w}_1 = \{M_{c1}, \omega_{31}, M_{c2}, \omega_{32}, f_{\omega_1}, f_{z1}, f_{\omega_2}, f_{z2}\}^T.$$

Компонентами вектора управления \vec{u} являются заданные значения M_{31}, M_{32} моментов двигателей $M_{д1}, M_{д2}$, которые подаются на регуляторы моментов при прямом управлении моментом асинхронного двигателя при векторном управлении так, что $\vec{u} = \{M_{31}, M_{32}\}^T$.

Компонентами вектора контролируемых переменных $\vec{z}(k)$ примем ошибку ε_1 отработки заданной скорости вращения ω_{31} с помощью двигателя первой платформы $\omega_{д1}$ так, что $\varepsilon_1 = \omega_{31} - \omega_{д1}$; переменную состояния первого интегратора z_1 ; управляющее воздействие первого канала u_1 , ошибку ε_2 отработки заданной скорости вращения ω_{32} с помощью двигателя второй платформы $\omega_{д2}$ так, что $\varepsilon_2 = \omega_{32} - \omega_{д2}$; переменную состо-

яния второго интегратора z_2 и управляющее воздействие второго канала u_2 . При этом вектор контролируемых переменных имеет следующий вид:

$$\vec{z}(k) = \{\varepsilon_1(k), z_1(k), u_1(k), \varepsilon_2(k), z_2(k), u_2(k)\}^T.$$

Компонентами вектора измеряемых переменных $\vec{y}(k)$ примем измеренные $\varepsilon_{1u}(k), \varepsilon_{2u}(k)$ с помехами $f_{\omega_1}(k), f_{\omega_2}(k)$ ошибки $\varepsilon_1(k), \varepsilon_2(k)$ регулирования скоростей вращения двигателей первой и второй платформ так, что

$$\begin{aligned} \varepsilon_{1u}(k) &= \varepsilon_1(k) + f_{\omega_1}(k), \\ \varepsilon_{2u}(k) &= \varepsilon_2(k) + f_{\omega_2}(k), \end{aligned}$$

а также измеренные $z_{1u}(k), z_{2u}(k)$ с помехами $f_{z1}(k), f_{z2}(k)$ вспомогательные переменные состояния $z_1(k), z_2(k)$ интеграторов так, что

$$\begin{aligned} z_{1u}(k) &= z_1(k) + f_{z1}(k), \\ z_{2u}(k) &= z_2(k) + f_{z2}(k), \end{aligned}$$

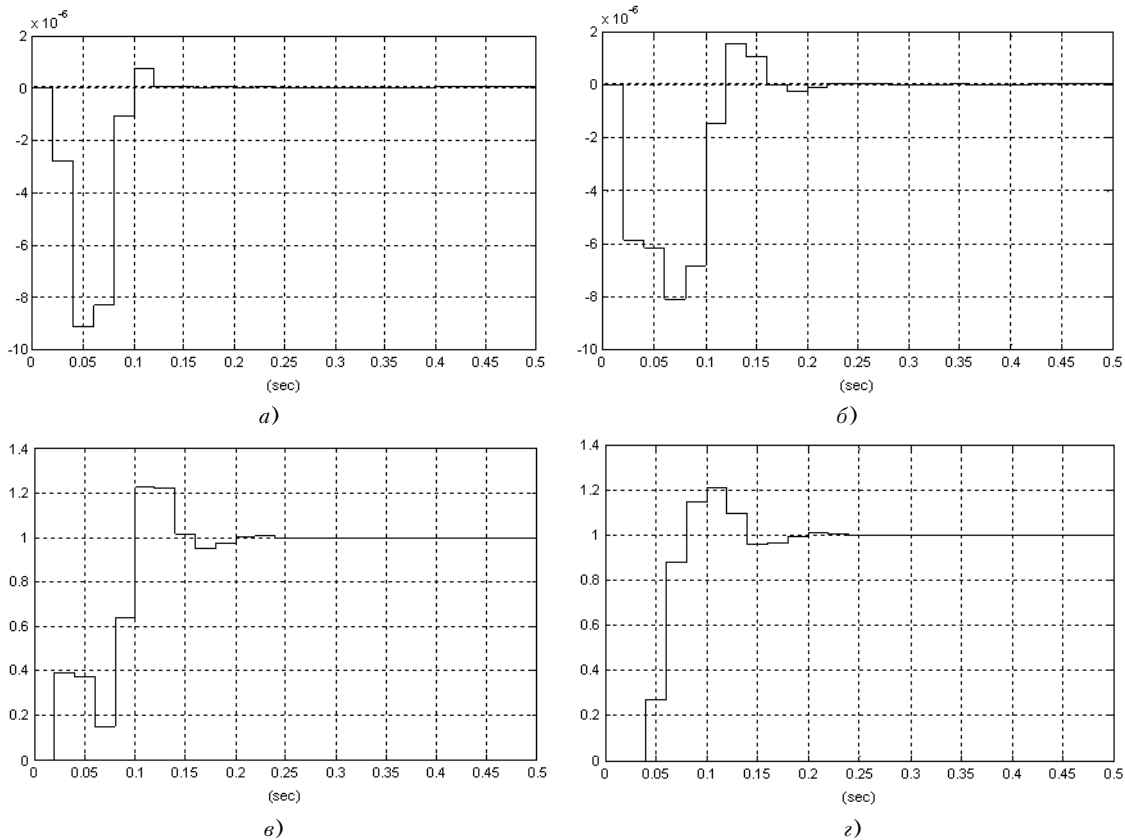


Рисунок 2 – Переходные процессы:

a – скорости вращения второго двигателя $\omega_{д2}$; *б* – скорости вращения второй платформы $\omega_{п2}$; *в* – момента упругости второй платформы M_{y2} ; *г* – момента второго двигателя $M_{д2}$ синтезированной системы по возмущению на второй канал

и, следовательно, вектор $\vec{y}(k)$ измеряемых переменных примет следующий вид:

$$\vec{y}(k) = \{\varepsilon_{1u}(k), z_{1u}(k), \varepsilon_{2u}(k), z_{2u}(k)\}^T.$$

В качестве примера на рис. 2 показаны переходные процессы: a – скорости вращения второго двигателя $\omega_{д2}$; b – скорости вращения второй платформы $\omega_{п2}$; v – момента упругости второй платформы M_{y2} ; z – момента второго двигателя $M_{д2}$ синтезированной системы по возмущению на второй канал. Как видно из этих графиков, в системе имеется астатизм по моменту сопротивления, а переходные процессы заканчиваются примерно за 0,25 с.

ВЫВОДЫ ИЗ ПРИВЕДЕННОГО ИССЛЕДОВАНИЯ, ПЕРСПЕКТИВЫ ЭТОГО НАПРАВЛЕНИЯ

Рассматривается достаточно широкий класс многоканальных систем, работающих по принципу грубого и точного управления, у которых нелинейные эффекты могут быть аппроксимированы аналитическими нелинейностями в виде степенных рядов по векторам состояния и управления. Наличие нелинейных элементов и упругих звеньев в кинематических звеньях многоканальных систем сдерживают получение высокой точности, потенциально присущей таким системам. Рассмотрены вопросы приближенного синтеза многоканальных систем цифрового нелинейного робастного управления. Нахождение робастного нелинейного управления каждого канала сводится к решению уравнения Гамильтона – Якоби – Беллмана для нелинейных моделей задающего и возмущающего воздействий и модели объекта управления канала.

Приводится пример синтеза робастного регулятора двухканальной электромеханической системы с раздельной нагрузкой, у которой каждый канал является двухмассовой электромеханической системой с векторным управлением асинхронными приводами.

Показано, что с помощью синтезированного нелинейного робастного регулятора удалось достаточно эффективно демпфировать собственные механические колебания, обусловленные наличием упругих элементов, и уменьшить негативное влияние нелинейных элементов на динамические характеристики системы.

ПЕРЕЧЕНЬ ССЫЛОК

1. Никитина Т. Б. Робастное управление многоканальными итерационными электроприводами по H^2 и H критериям // Електромашинобудування та електрообладнання. Міжвідомчий науково-технічний збірник. – Випуск № 67. – Одеса, 2006. – С. 13–17.
2. Никитина Т. Б. Синтез робастных регуляторов многоканальных итерационных систем // Радиозлектроника

- и информатика. Научно-технический журнал. – Харьков: ХНУР, 2007. – № 4(35). – С. 24–30.
3. Никитина Т. Б. Синтез цифровых робастных регуляторов многоканальных электромеханических систем // Збірник наукових праць Дніпродзержинського державного технічного університету (технічні науки). Тематичний випуск «Проблеми автоматизованого електропривода. Теорія і практика». – Дніпродзержинськ: ДГТУ, 2007. – С. 391–392.
4. Никитина Т. Б. Синтез цифровых нелинейных робастных регуляторов каналов многоканальных систем при последовательном синтезе. // Автоматизация виробничих процесів. – Київ, 2006. – № 2 (23). – С. 109–114.
5. Никитина Т. Б. Синтез приближенно-оптимальных нелинейных систем цифрового управления технологическими процессами с аналитическими нелинейностями // Автоматизация виробничих процесів. – Київ, 2003. – № 2(17). – С. 62–65.
6. Кузнецов Б. И., Никитина Т. Б., Коломиец В. В. Синтез электромеханических систем со сложными кинематическими целями. – Харьков: УИПА, 2005. – 511с.
7. Никитина Т. Б. Стабилизация динамических характеристик двухканальной системы управления обмоточной машиной с помощью робастного управления // Вестник НТУ «ХПИ»: Сборник научных трудов. Тематический выпуск «Системный анализ, управление и информационные технологии». – 2006. – № 39. – С. 38–45.
8. Никитина Т. Б. Робастное управление двухмассовой электромеханической системой в режиме буксования // Вестник НТУ «ХПИ»: Сборник научных трудов. Тематический выпуск «Информатика и моделирование». – 2006. – № 40. – С. 138–144.
9. Никитина Т. Б. Исследование динамических характеристик цифрового робастного управления блюмингом в режиме пробуксовки валков // Вестник НТУ «ХПИ». Сборник научных трудов. Тематический выпуск «Автоматика и приборостроение». – 2006. – № 31. – С. 93–101.
10. Никитина Т. Б. Ограничение нагрузок в нелинейных многоканальных электромеханических системах при управлении по состоянию // Механіка та машинобудування. – Харків: 2006. – №1. – С. 259–264.
11. Никитина Т. Б. Робастная система управления двухмассовой обмоточной машиной // Вісник Хмельницького національного університету (технічні науки). – Хмельницький: Хмельницький національний університет, 2007. – № 5(96). – С. 182–187.
12. Никитина Т. Б. Цифрове робастне управління двомасовою обмотувальною машиною // Вісник Національного університету «Львівська політехніка». Електроенергетичні та електромеханічні системи. – Львів, 2007. – № 587. – С. 70–75.
13. Никитина Т. Б. Робастное управление трехпорной управляемой платформой // Вестник НТУ «ХПИ»: Сб. науч. раб. Тематический выпуск «Системный анализ – управление и информационные технологии». – Харьков: НТУ «ХПИ». – 2007. – № 18. – С. 29–36.
14. Никитина Т. Б. Робастная стабилизация дискретно-континуального объекта // Технічна електродинаміка. Тематичний випуск. Силова електроніка та енергоефективність. – Київ.: 2007. – Ч. 1. – С. 56–61.
15. Никитина Т. Б. Цифровая робастная стабилизация танкового вооружения // Електромашинобудування та електрообладнання. Міжвідомчий науково-технічний збірник. – Одеса.: 2007. – Випуск № 68. – С. 16–21.

Надійшла 27.11.07

Розроблено метод синтезу нелінійного робастного управління багатоканальними системами, що працюють за принципом грубого і точного управління відповідно до ітераційного алгоритму. Наведено приклад динамічних характеристик синтезованої двокаанальної електромеханічної системи.

A method of synthesis for robust control multichannel systems, workings on rough and exact control principle in accordance with iteration algorithm is developed. The example of dynamic descriptions for synthesized twochannel electromechanical system is resulted.

Е. М. Потапенко, А. Е. Казурова, Е. В. Душинова, В. И. Левыкина

АСИМПТОТИЧЕСКОЕ ВЫПОЛНЕНИЕ МАТЕМАТИЧЕСКИХ ОПЕРАЦИЙ НА МНОГОФАЗНЫХ СИГНАЛАХ

Разработан аппарат для проведения математических операций на многофазных зашумленных сигналах с неизвестной частотой и с постоянными неизвестными смещениями ноля. В частности, аппарат позволяет получить производные и интегралы главных гармоник фазовых сигналов. В качестве примера получены два алгоритма определения потокосцепления асинхронного двигателя по его ЭДС.

ВВЕДЕНИЕ

В науке и технике широко используются однофазные, двухфазные, трехфазные и т. д. сигналы. Однофазные сигналы встречаются практически везде. Двухфазные сигналы появляются, в частности, при рассмотрении плоского вращения вектора относительно какой-либо двумерной системы координат (базиса) с помощью его проекций на оси этого базиса. Трехфазные сигналы имеют место, например, при рассмотрении плоского вращения вектора в проекциях на три оси, лежащие в плоскости вращения. Типичной областью применения двухфазных и трехфазных сигналов является электротехника переменного тока и, в частности, электродвигатели переменного тока. Трехфазную систему можно свести к эквивалентной двухфазной системе следующим образом [1].

Пусть x_a , x_b , x_c – переменные (токи, напряжения, потокосцепления) каждой из фаз статора. Эти переменные можно представить в виде трех векторов с общим результирующим вектором. Результирующий вектор можно разложить на две составляющие x_α , x_β вдоль осей ортогональной системы координат (α, β) , жестко связанной со статором двигателя. В трехфазной электротехнической системе оси фаз расположены под углом 120° друг к другу. Если переменные подчиняются зависимости

$$x_a + x_b + x_c = 0,$$

то имеют место следующие соотношения [1]:

$$\begin{aligned} x_\alpha &= \sqrt{\frac{3}{2}}x_a, \quad x_\beta = \frac{1}{\sqrt{2}}(x_a + 2x_b); \\ x_a &= \sqrt{\frac{2}{3}}x_\alpha, \quad x_b = \sqrt{\frac{2}{3}}\left(-\frac{1}{2}x_\alpha + \frac{\sqrt{3}}{2}x_\beta\right), \end{aligned}$$

$$x_c = -\sqrt{\frac{2}{3}}\left(\frac{1}{2}x_\alpha + \frac{\sqrt{3}}{2}x_\beta\right).$$

В каждой фазе могут присутствовать постоянное смещение ноля и высокочастотные шумы. Для качественного управления электротехнической системой требуется выделение главных гармоник фаз и проведение с ними различных математических операций: дифференцирование, интегрирование и т. п. После дифференцирования реального (исходного) сигнала уровень высокочастотных шумов возрастает, что может сделать невозможным дальнейшее использование полученного сигнала. О важности учета этого эффекта говорит существование специального раздела в теории фильтрации «Дифференцирование сигналов» [2]. После интегрирования сигнала уровень шумов снижается, но появляются постоянная ошибка, равная неизвестной постоянной интегрирования, и линейно возрастающая со временем ошибка за счет интегрирования неизвестного смещения ноля. В отличие от однофазного сигнала, многофазные сигналы обладают дополнительной информацией за счет детерминированной взаимосвязи между фазами (одни и те же частота и амплитуда, постоянный сдвиг фазы между сигналами). В работе [3] этот факт был успешно использован для фильтрации двухфазного сигнала.

Целью данной статьи является использование свойств многофазных сигналов для выполнения ряда математических операций с ними, таких как дифференцирование, интегрирование, фильтрация и др.

Поскольку существуют эквивалентные преобразования между трехфазными и двухфазными сигналами, то ниже будут рассматриваться только двухфазные сигналы.

ОСНОВНОЙ РЕЗУЛЬТАТ

Рассматривается двухфазная система x_α , x_β , представляющая собой проекции вектора x на оси ортогонального базиса (α, β) . Пусть главные гармоники фаз описываются уравнениями

$$x_{0\alpha} = |x_0| \cos \omega t, \quad x_{0\beta} = |x_0| \sin \omega t. \quad (1)$$

Производные и интегралы выражений (1) при условии медленности изменения амплитуды x_0 и частоты ω приближенно можно описать уравнениями

$$\begin{aligned} \dot{x}_{0\alpha} &= -\omega|x_0|\sin\omega t = -\omega x_{0\beta}, \\ \dot{x}_{0\beta} &= \omega|x_0|\cos\omega t = \omega x_{0\alpha}, \end{aligned} \quad (2)$$

$$\begin{aligned} \int x_{0\alpha} dt &= \omega^{-1}|x_0|\sin\omega t = \omega^{-1}x_{0\beta}, \\ \int x_{0\beta} dt &= -\omega^{-1}|x_0|\cos\omega t = -\omega^{-1}x_{0\alpha}. \end{aligned} \quad (3)$$

Умножим в (2) первое уравнение на $x_{0\beta}$, второе уравнение на $-x_{0\alpha}$ и затем полученные уравнения сложив, найдем

$$\omega = \frac{x_{0\alpha}\dot{x}_{0\beta} - x_{0\beta}\dot{x}_{0\alpha}}{x_{0\alpha}^2 + x_{0\beta}^2}. \quad (4)$$

На практике имеют место соотношения

$$x_\alpha = x_{0\alpha} + \vartheta_\alpha, \quad x_\beta = x_{0\beta} + \vartheta_\beta, \quad (5)$$

где ϑ_α и ϑ_β представляют собой высокочастотные по сравнению с основной частотой ω погрешности входных сигналов x_α, x_β . Для устранения высокочастотных помех сигналы (5) предлагается пропустить через идентичные фильтры низких частот с передаточной функцией $W_f(p)$, имеющей амплитудную $A(\omega)$ и фазовую $\varphi(\omega)$ частотные характеристики. Не принимая во внимание сдвиг по фазе помех $\vartheta_\alpha, \vartheta_\beta$ на выходах фильтров, будем иметь

$$\begin{aligned} x_\alpha^* &= A(\omega)|x_0|\cos(\omega t + \varphi) + A(\omega_\vartheta)\vartheta_\alpha, \\ x_\beta^* &= A(\omega)|x_0|\sin(\omega t + \varphi) + A(\omega_\vartheta)\vartheta_\beta \end{aligned}$$

или, что то же самое,

$$\begin{aligned} x_\alpha^* &= A(\omega)|x_0|(\cos\omega t \cos\varphi - \sin\omega t \sin\varphi) + A(\omega_\vartheta)\vartheta_\alpha, \\ x_\beta^* &= A(\omega)|x_0|(\sin\omega t \cos\varphi + \cos\omega t \sin\varphi) + A(\omega_\vartheta)\vartheta_\beta, \end{aligned} \quad (6)$$

где ω_ϑ – частота погрешностей. Из (6) на основании (2), (5) можно найти точные значения проекций вектора x_0 (искомые сигналы) в виде

$$\begin{aligned} x_{0\alpha} &= |x_0|\cos\omega t = A^{-1}(\omega)(x_\alpha^* \cos\varphi + x_\beta^* \sin\varphi) + \\ &+ A^{-1}(\omega)A(\omega_\vartheta)(\vartheta_\alpha \cos\varphi + \vartheta_\beta \sin\varphi), \\ x_{0\beta} &= |x_0|\sin\omega t = A^{-1}(\omega)(x_\beta^* \cos\varphi - x_\alpha^* \sin\varphi) + \\ &+ A^{-1}(\omega)A(\omega_\vartheta)(\vartheta_\beta \cos\varphi - \vartheta_\alpha \sin\varphi). \end{aligned} \quad (7)$$

Поскольку погрешности ϑ_α и ϑ_β не известны, то вместо точных значений $x_{0\alpha}$ и $x_{0\beta}$ будем использовать их оценки $\hat{x}_{0\alpha}, \hat{x}_{0\beta}$ по зависимостям

$$\begin{aligned} \hat{x}_{0\alpha} &= A^{-1}(\omega)(x_\alpha^* \cos\varphi + x_\beta^* \sin\varphi), \\ \hat{x}_{0\beta} &= A^{-1}(\omega)(x_\beta^* \cos\varphi - x_\alpha^* \sin\varphi). \end{aligned} \quad (8)$$

В выражения (8) входят выходные сигналы фильтров x_α^*, x_β^* , с пониженным содержанием шумов, следовательно, эти выражения можно рассматривать как оценки главных гармоник исходных сигналов x_α, x_β . Определим погрешности этих оценок. Будем полагать, что амплитуды сигналов ϑ_α и ϑ_β равны ϑ_a . Тогда погрешности $\tilde{x}_\alpha = \hat{x}_{0\alpha} - x_\alpha, \tilde{x}_\beta = \hat{x}_{0\beta} - x_\beta$ оценок (8) по сравнению с точными значениями (6) можно оценить выражением

$$\tilde{x}_\alpha = \tilde{x}_\beta = A^{-1}(\omega)A(\omega_\vartheta)\vartheta_a. \quad (9)$$

Фильтры выбираются из условия близости $A(\omega)$ к единице, а $A(\omega_\vartheta)$ к нулю.

Пусть собственно фильтры представляют собой инерционные звенья первого порядка

$$W_f(p) = \frac{k_f}{Tp + 1}, \quad k_f > 0, T > 0,$$

(T – постоянная времени) с частотными характеристиками

$$A(\omega) = \frac{k_f}{\sqrt{(T\omega)^2 + 1}}, \quad \text{tg}\varphi(\omega) = -T\omega. \quad (10)$$

С помощью соотношений

$$\sin\varphi(\omega) = \frac{\text{tg}\varphi(\omega)}{\sqrt{1 + \text{tg}^2\varphi(\omega)}}, \quad \cos\varphi(\omega) = \frac{1}{\sqrt{1 + \text{tg}^2\varphi(\omega)}}$$

и выражений (10) можно записать

$$\sin\varphi = -T\omega k_f^{-1}A(\omega), \quad \cos\varphi = k_f^{-1}A(\omega). \quad (11)$$

Подстановка (10), (11) в (8) дает

$$\begin{aligned} \hat{x}_{0\alpha} &= k_f^{-1}(x_\alpha^* - T\hat{\omega}x_\beta^*), \\ \hat{x}_{0\beta} &= k_f^{-1}(x_\beta^* + T\hat{\omega}x_\alpha^*), \end{aligned} \quad (12)$$

где ω заменена ее оценкой $\hat{\omega}$, которая по аналогии с (4) определяется выражением

$$\hat{\omega} = \frac{x_{\alpha}^* \dot{x}_{\beta}^* - x_{\beta}^* \dot{x}_{\alpha}^*}{x_{\alpha}^{*2} + x_{\beta}^{*2}}. \quad (13)$$

С использованием (2), (3), (12) запишем

$$\begin{aligned} \dot{\hat{x}}_{0\alpha} &= -\hat{\omega} k_f^{-1} (x_{\beta}^* + T \hat{\omega} x_{\alpha}^*), \\ \dot{\hat{x}}_{0\beta} &= \hat{\omega} k_f^{-1} (x_{\alpha}^* - T \hat{\omega} x_{\beta}^*), \end{aligned} \quad (14)$$

$$\begin{aligned} \int_0^t \hat{x}_{0\alpha} dt &= (\hat{\omega} k_f)^{-1} (x_{\beta}^* + T \hat{\omega} x_{\alpha}^*), \\ \int_0^t \hat{x}_{0\beta} dt &= -(\hat{\omega} k_f)^{-1} (x_{\alpha}^* - T \hat{\omega} x_{\beta}^*) \end{aligned} \quad (15)$$

или при $k_f = T$

$$\begin{aligned} \int_0^t \hat{x}_{0\alpha} dt &= (x_{\alpha}^* + (T \hat{\omega})^{-1} x_{\beta}^*), \\ \int_0^t \hat{x}_{0\beta} dt &= (x_{\beta}^* - (T \hat{\omega})^{-1} x_{\alpha}^*). \end{aligned} \quad (16)$$

Выражение скорости (13) содержит производные от сигналов, пропущенных через инерционные звенья первого порядка, то есть $\hat{\omega}$, а следовательно и переменные (12), (14)–(16) фактически будут содержать не отфильтрованные высокочастотные шумы. Для повышения помехозащищенности можно сигнал $\hat{\omega}$ пропустить через фильтр низкой частоты. Рассмотрим другой путь.

Вместо фильтров первого порядка применим фильтр второго порядка с передаточной функцией и частотными характеристиками следующего вида:

$$W_f(p) = \frac{k_f}{T^2 p^2 + 2dTp + 1}; \quad k_f, T, d > 0, \quad (17)$$

$$\begin{aligned} A(\omega) &= \frac{k_f}{\sqrt{(1 - T^2 \omega^2)^2 + (2dT\omega)^2}}, \\ \operatorname{tg} \varphi(\omega) &= -\frac{2dT\omega}{1 - T^2 \omega^2}, \end{aligned} \quad (18)$$

откуда

$$\begin{aligned} \sin \varphi(\omega) &= -\frac{2dT\omega}{\sqrt{(1 - T^2 \omega^2)^2 + (2dT\omega)^2}}, \\ \cos \varphi(\omega) &= \frac{1 - T^2 \omega^2}{\sqrt{(1 - T^2 \omega^2)^2 + (2dT\omega)^2}}. \end{aligned} \quad (19)$$

Подстановка (18), (19) в (8) дает уравнения оценок главных гармоник

$$\begin{aligned} \dot{\hat{x}}_{0\alpha} &= k_f^{-1} [x_{\alpha}^* (1 - T^2 \hat{\omega}^2) - x_{\beta}^* 2dT\hat{\omega}], \\ \dot{\hat{x}}_{0\beta} &= k_f^{-1} [x_{\beta}^* (1 - T^2 \hat{\omega}^2) + x_{\alpha}^* 2dT\hat{\omega}] \end{aligned} \quad (20)$$

где ω заменено на $\hat{\omega}$. С использованием (2), (3), (20) запишем

$$\begin{aligned} \dot{\hat{x}}_{0\alpha} &= -\hat{\omega} k_f^{-1} [x_{\beta}^* (1 - T^2 \hat{\omega}^2) + x_{\alpha}^* 2dT\hat{\omega}], \\ \dot{\hat{x}}_{0\beta} &= \hat{\omega} k_f^{-1} [x_{\alpha}^* (1 - T^2 \hat{\omega}^2) - x_{\beta}^* 2dT\hat{\omega}]; \end{aligned} \quad (21)$$

$$\begin{aligned} \int_0^t \hat{x}_{0\alpha} dt &= (\hat{\omega} k_f)^{-1} [x_{\beta}^* (1 - T^2 \hat{\omega}^2) + x_{\alpha}^* 2dT\hat{\omega}], \\ \int_0^t \hat{x}_{0\beta} dt &= -(\hat{\omega} k_f)^{-1} [x_{\alpha}^* (1 - T^2 \hat{\omega}^2) - x_{\beta}^* 2dT\hat{\omega}]. \end{aligned} \quad (22)$$

В рассмотренном случае выражение скорости (13) содержит производные первого порядка от сигналов, пропущенных через колебательные звенья (второго порядка), то есть $\hat{\omega}$, а следовательно, и переменные (20)–(22) не будут содержать не отфильтрованные высокочастотные шумы.

Фильтры (12), (20) подробно изучены в работе [3]. Дифференциаторы (14), (21) и интеграторы (15), (22) рассмотрены впервые. Следует отметить возможность применения интеграторов для оценки не измеряемых потокосцеплений двигателей переменного тока путем интегрирования ЭДС.

Потокосцепление ротора Ψ связано с ЭДС двигателя e зависимостями

$$\dot{\Psi}_{\alpha} = L_r L_m^{-1} e_{\alpha}, \quad \dot{\Psi}_{\beta} = L_r L_m^{-1} e_{\beta}, \quad (23)$$

где L_r, L_m – индуктивность ротора и взаимная индуктивность ротора и статора. По выражениям (23) путем интегрирования можно получить потокосцепление, но с точностью до неизвестных начальных условий и с возрастающими с течением времени ошибками за счет постоянных смещений в фазах ЭДС. Для устранения влияния начальных условий и ограничения ошибок за счет смещения нолей воспользуемся алгоритмом интегрирования (16), синтезированного для инерционного звена первого порядка. Обозначив

$$x_{\alpha} := L_r L_m^{-1} e_{\alpha}, \quad x_{\beta} := L_r L_m^{-1} e_{\beta}, \quad (24)$$

с помощью алгоритма (16) получим

$$\begin{aligned}\hat{\Psi}_\alpha &= L_r L_m^{-1} (e_\alpha^* + (T\hat{\omega})^{-1} e_\beta^*), \\ \hat{\Psi}_\beta &= L_r L_m^{-1} (e_\beta^* - (T\hat{\omega})^{-1} e_\alpha^*),\end{aligned}\quad (25)$$

где

$$e_\alpha^* = \frac{T}{Tp+1} e_\alpha, \quad e_\beta^* = \frac{T}{Tp+1} e_\beta; \quad (26)$$

$$\hat{\omega} = \frac{e_\alpha^* \dot{e}_\beta^* - e_\beta^* \dot{e}_\alpha^*}{e_\alpha^{*2} + e_\beta^{*2}}. \quad (27)$$

Выражения (25)–(27) в точности совпадают с выражениями потокосцепления, полученными в работе [4] совершенно другим методом, а именно, с помощью наблюдателя Луэнбергера. Работа [4] содержит результаты численного моделирования, подтверждающие высокую точность интегрирования.

Пусть теперь используются фильтры второго порядка (17) и обозначения (24). Тогда в соответствии с (22)

$$\begin{aligned}\hat{\Psi}_\alpha &= (\hat{\omega} k_f)^{-1} [e_\beta^* (1 - T^2 \hat{\omega}^2) + e_\alpha^* 2dT\hat{\omega}], \\ \hat{\Psi}_\beta &= -(\hat{\omega} k_f)^{-1} [e_\alpha^* (1 - T^2 \hat{\omega}^2) - e_\beta^* 2dT\hat{\omega}],\end{aligned}\quad (28)$$

где

$$\begin{aligned}e_\alpha^* &= \frac{k_f}{T^2 p^2 + 2dTp + 1} e_\alpha, \\ e_\beta^* &= \frac{k_f}{T^2 p^2 + 2dTp + 1} e_\beta.\end{aligned}\quad (29)$$

Выражения (27)–(29) составляют алгоритм оценки потокосцепления с фильтрами второго порядка.

ЗАКЛЮЧЕНИЕ

Получены соотношения, позволяющие осуществлять такие операции как различной степени фильтрация, дифференцирование и интегрирование многофазных зашумленных сигналов с неизвестной частотой. Разработанный метод можно распространить и на другие математические операции.

ПЕРЕЧЕНЬ ССЫЛОК

1. Чиликин М. Г., Ключев В. И., Сандлер А. С. Теория автоматизированного электропривода. – М.: Энергия, 1979. – 616 с.
2. Емельянов С. В., Коровин С. К. Новые типы обратной связи: Управление при неопределенности. – М.: Наука. Физматлит, 1997. – 352 с.
3. Потапенко Е. Е., Потапенко Е. М. Синтез и анализ компенсационных фильтров многофазных неопределенных сигналов // Вісник Національного технічного університету «ХПІ». Збірник наукових праць «Проблеми автоматизованого електропривода. Теорія і практика». – Харків: НТУ «ХПІ». – 2003. – № 10, т.2. – С. 342–344.
4. Потапенко Е. М., Потапенко Е. Е. Оценка векторов потокосцеплений и их скоростей в двигателях переменного тока // Вісник Національного технічного університету «ХПІ». Збірник наукових праць. Тематичний випуск «Проблеми автоматизованого електропривода. Теорія і практика». – Харків: НТУ «ХПІ». – 2003. – № 10, т. 1. – С. 105–107.

Надійшла 11.09.07

Розроблено апарат для проведення математичних операцій на багатофазних зашумлених сигналах з невідомою частотою та з постійними невідомими зсувами поля. Зокрема, апарат дозволяє отримати похідні та інтеграли головних гармонік фазових сигналів. Як приклад отримано два алгоритми визначення потокосцеплення асинхронного двигуна за його ЕРС.

The apparatus for realizing mathematical operations on multiphase noising signals with unknown frequency and with constant unknown displacements of zeros was developed. In particular the apparatus permits to find derivatives and integrals of main harmonics of phase signals. As an example two algorithms of induction motor flux estimation by electromotive force are obtained.

КЕРУВАННЯ У ТЕХНІЧНИХ СИСТЕМАХ

УПРАВЛЕНИЕ В ТЕХНИЧЕСКИХ СИСТЕМАХ

CONTROL IN TECHNICAL SYSTEMS

УДК 681.527.2

Е. М. Потапенко, А. В. Соломаха

СИНТЕЗ И АНАЛИЗ КОМПЕНСАТОРА ПАССИВНОГО LC-ФИЛЬТРА НА ВЫХОДЕ ПРЕОБРАЗОВАТЕЛЯ ЧАСТОТЫ

Синтезирован алгоритм работы упреждающего компенсатора, устраняющего искажения амплитуды и фазы, вносимые пассивным фильтром. Численное и натурное моделирование подтвердило работоспособность синтезированного алгоритма.

ВВЕДЕНИЕ

Преобразователь частоты (ПЧ) с широтно-импульсной модуляцией (ШИМ) формирует выходное напряжение в виде набора прямоугольных импульсов одинаковой амплитуды. Их крутые фронты вызывают старение изоляции двигателя, питающих кабелей и ухудшение их диэлектрических свойств. Это также приводит к дополнительному нагреву обмоток двигателя и появлению высокочастотного акустического шума. Устранение вышеназванных эффектов особенно актуально для мощных приводов [1, 2]. Поэтому на выходе преобразователя частоты для сглаживания выходного напряжения устанавливают пассивные силовые LC-фильтры [3] (ПСФ). Пассивные фильтры улучшают гармонический состав тока и напряжения, уменьшают акустический шум, уменьшают пусковой ток двигателей, имеют простую конструкцию и относительно дешевые в изготовлении. Ограничение скорости изменения напря-

жения и снижение перенапряжений позволяет отказаться от увеличения номиналов инверторов за счет учета явления обратной волны в кабелях. Без фильтра напряжения срок службы двигателя уменьшается на 10 % по отношению к его декларированному сроку работы [4]. В основном, пассивные фильтры используются в потребителях мощности более 1 мегавольтампера. Пассивные фильтры не требуют технического обслуживания, экономичны, при условии, что сопротивление фильтра должно быть меньше, чем сопротивление системы [1]. К сожалению, пассивные фильтры искажают главные гармоники по амплитуде и фазе.

Целью данной статьи является синтез упреждающего компенсатора пассивного силового фильтра (УК) для устранения искажений амплитуды и фазы главной гармоники выходного напряжения ПЧ.

ПОСТАНОВКА ЗАДАЧИ

На рис. 1 представлены функциональные схемы цепи компенсации неидеальностей ПСФ и автономного инвертора, где СУ – система управления автономным инвертором с ШИМ (АИН с ШИМ); ПСФ – силовой

© Потапенко Е. М., Соломаха А. В., 2007

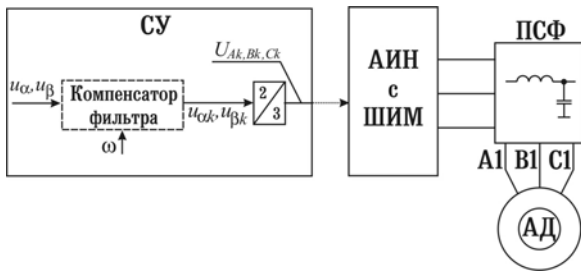


Рисунок 1 – Функциональные схемы цепи компенсации неидеальностей ПСФ и автономного инвертора

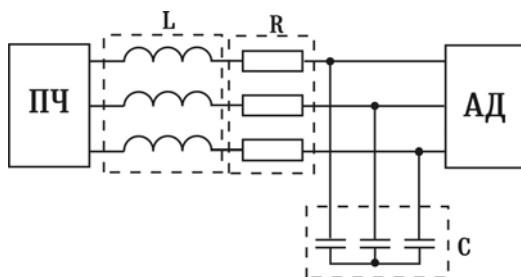


Рисунок 2 – Силовой пассивный фильтр

LC-фильтр; АД – асинхронный двигатель; u_{ap}, u_{bp} – управляющие (программные) сигналы системы управления АИН, u_{ap}^k, u_{bp}^k – программные сигналы с учетом компенсации, u_{Ap}, u_{Bp}, u_{Cp} – фазовые управляющие сигналы.

Схема подключения пассивного силового фильтра в электроприводе показана на рис. 2.

Напряжения, получаемые после LC-фильтров, оказываются искаженными по фазе и амплитуде относительно выходных напряжений ПЧ. Это приводит к ухудшению динамики системы, в которую входит преобразователь частоты. Задача состоит в упреждающем изменении программных сигналов по амплитуде и фазе таким образом, чтобы они компенсировали искажения, вносимые ПСФ. Такая возможность имеется за счет введения перекрестных связей между фазами многофазных сигналов. Это удобнее делать с двухфазными программными сигналами. Предлагаемый компенсационный фильтр состоит из компенсатора и собственно пассивного фильтра.

СИНТЕЗ УПРЕЖДАЮЩЕГО КОМПЕНСАТОРА ФИЛЬТРА

Силовой пассивный фильтр можно описать передаточной функцией

$$W(p) = \frac{1}{LCp^2 + RCp + 1}, \quad (1)$$

где R – активное сопротивление дросселя, или при необходимости, специально введенное в состав фильтра для получения необходимой фильтрации выходного напряжения ПЧ. Передаточной функции (1) будут соответствовать АЧХ и ФЧХ вида

$$A(\omega) = \frac{1}{\sqrt{(1 - LC\omega^2)^2 + (RC\omega)^2}}, \quad (2)$$

$$\psi(\omega) = -\arctg\left(\frac{RC\omega}{1 - LC\omega^2}\right). \quad (3)$$

Частота сигнала обычно используется в алгоритмах управления двигателями переменного тока и считается известной. Из выражения (3) найдем

$$\operatorname{tg}\psi(\omega) = -\frac{RC\omega}{1 - LC\omega^2}. \quad (4)$$

Используя тригонометрические преобразования

$$\sin\psi(\omega) = \frac{\operatorname{tg}\psi(\omega)}{\sqrt{1 + \operatorname{tg}^2\psi(\omega)}}, \quad \cos\psi(\omega) = \frac{1}{\sqrt{1 + \operatorname{tg}^2\psi(\omega)}}, \quad (5)$$

найдем

$$\sin\psi(\omega) = -\frac{RC\omega}{\sqrt{(1 - LC\omega^2)^2 + (RC\omega)^2}}, \quad (6)$$

$$\cos\psi(\omega) = \frac{1 - LC\omega^2}{\sqrt{(1 - LC\omega^2)^2 + (RC\omega)^2}}. \quad (7)$$

Необходимо компенсировать искажения амплитуды (2) и фазы, определяемой выражениями (6), (7). Программный двухфазный сигнал имеет вид

$$u_\alpha = |u|\cos\omega t, \quad u_\beta = |u|\sin\omega t, \quad |u| = \sqrt{u_\alpha^2 + u_\beta^2}. \quad (8)$$

После пропускания двухфазного сигнала (8) или эквивалентного ему трехфазного сигнала через силовой фильтр, эквивалентный двухфазный сигнал примет вид

$$\begin{aligned} u_{\alpha f} &= A(\omega)|u|\cos(\omega t + \psi(\omega)), \\ u_{\beta f} &= A(\omega)|u|\sin(\omega t + \psi(\omega)), \end{aligned} \quad (9)$$

где $A(\omega)$ и $\psi(\omega)$ – АЧХ и ФЧХ LC-фильтра соответственно.

Упреждающий компенсатор фильтра (УК) предназначен для устранения искажений в (9), связанных

с $A(\omega)$ и $\psi(\omega)$. Для этого УК должен иметь следующие частотные характеристики:

$$A_k(\omega) = A^{-1}(\omega), \quad \psi_k(\omega) = -\psi(\omega), \quad (10)$$

а его выходные сигналы с учетом (9) должны иметь вид

$$\begin{aligned} u_{\alpha k} &= A^{-1}(\omega)|u|\cos(\omega t - \psi(\omega)) = \\ &= A^{-1}(\omega)|u|(\cos \omega t \cos \psi(\omega) + \sin \omega t \sin \psi(\omega)) = \\ &= A^{-1}(\omega)(x_\alpha \cos \psi(\omega) + x_\beta \sin \psi(\omega)), \\ u_{\beta k} &= A^{-1}(\omega)|u|\sin(\omega t - \psi(\omega)) = \\ &= A^{-1}(\omega)(u_\beta \cos \psi(\omega) - u_\alpha \sin \psi(\omega)). \end{aligned} \quad (11)$$

Подставляя (2), (6), (7) в (11), получим уравнения работы упреждающего компенсационного фильтра.

$$\begin{aligned} u_{\alpha k} &= u_\alpha(1 - LC\omega^2) - u_\beta(RC\omega), \\ u_{\beta k} &= u_\beta(1 - LC\omega^2) + u_\alpha(RC\omega). \end{aligned} \quad (12)$$

РЕЗУЛЬТАТЫ МОДЕЛИРОВАНИЯ ПРОЦЕССА КОМПЕНСАЦИИ

Моделировалась система (рис. 3), состоящая из следующих последовательно соединенных звеньев: источник сигнала, УК, преобразователь координат $(\alpha, \beta) / (A, B, C)$, блок формирования ШИМ-сигнала, динамические звенья с передаточной функцией (5), стоящие в каждой из фаз. Напряжения U_A, U_B, U_C являются желаемыми сигналами, с которыми будем сравнивать выходные сигналы фильтра U_{A1}, U_{B1}, U_{C1} .

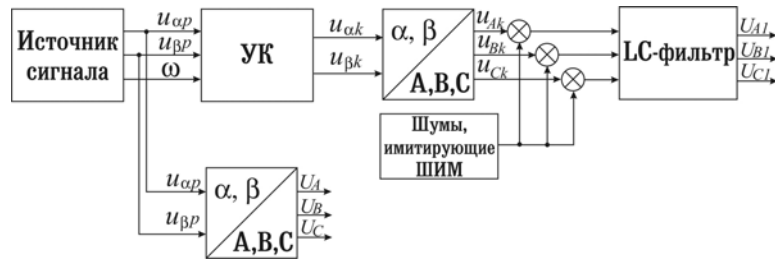


Рисунок 3 – Схема моделирования УК

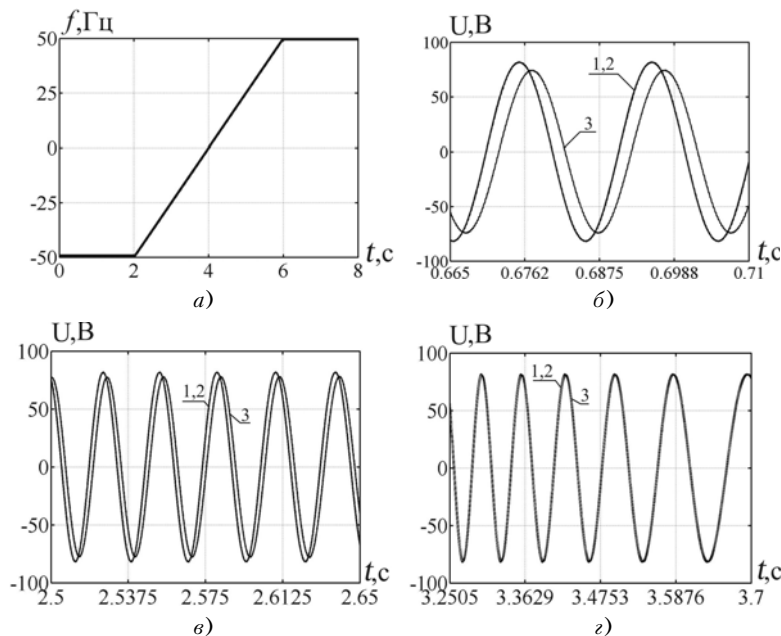


Рисунок 4:

а – частота изменения сигнала источника, б, в, г – желаемый (1), скорректированный сигнал на выходе LC-фильтра (2) и отфильтрованный без коррекции (3) сигналы фазы А на разных частотах

Источник генерировал периодический двухфазный сигнал амплитудой 100 В и переменной частотой, указанной на рис. 4, а. Моделирование проводилось со следующими параметрами фильтра: $R = 0,01$ Ом; $C = 40$ мкФ; $L = 0,01$ Гн.

Исходя из соотношений для преобразования $\alpha, \beta \rightarrow A, B, C$ [5], амплитуда соответствующего ему трехфазного сигнала равна 81,6 В. Перед LC-фильтрами на сигнал накладывались шумы, имитирующие ШИМ, амплитудой в 2 раза большей амплитуды полезного сигнала и частотой 4 кГц. Дискретность алгоритма УК составляла 100 мкс. На рис. 4, б; 4, в; 4, г, соответствующих высоким, средним и низким частотам, показаны желаемый сигнал фазы А U_A – кривая 1, скомпенсированный отфильтрованный U_{A1} – кривая 2 и отфильтрованный без компенсации – кривая 3.

Как видно из рис. 4, компенсационный фильтр работоспособен во всем диапазоне частот ± 50 Гц, при этом имеет 3 % ошибку в амплитуде, обусловленную шумами, по амплитуде превосходящей полезный сигнал в 2 раза. Компенсационный фильтр снижает амплитуду шумов в 54 раза при отсутствии искажений главной гармоники по амплитуде и фазе.

Экспериментальная проверка предлагаемого фильтра представлена в работе [6].

ЗАКЛЮЧЕНИЕ

Синтезирован алгоритм работы упреждающего компенсатора фильтра для коррекции в реальном времени задающего сигнала ШИМ. С помощью этого алгоритма можно получить на выходе пассивного LC-фильтра желаемый отфильтрованный сигнал без искажений главной гармоники по фазе и амплитуде. С помощью компенсационного фильтра предотвращается износ изо-

ляции и уменьшаются потери на нагрев двигателя. Система компенсации входит в состав основной программы управления ПЧ и не требует дополнительной аппаратуры. Численное моделирование подтвердило работоспособность синтезированного алгоритма в диапазоне частот сигнала ± 50 Гц. Предложенный метод компенсации подходит для всех двигателей с АИН с ШИМ.

ПЕРЕЧЕНЬ ССЫЛОК

1. Das J. C. Passive Filters – Potentialities and Limitation // IEEE Trans. Ind. Applicat., Jan./Febr. – 2004. – Vol. 40, No. 1. – Pp. 232–241.
2. Kojima M., Hirabayashi K., Kawabata Y., Ejiogu E. C., Kawabata T. Novel Vector Control System Using Deadbeat-Controlled PWM Inverter With Output LC Filter // IEEE Trans. Ind. Applicat., Jan./Febr. – 2004. – Vol.40, No. 1. – Pp. 162–169.
3. Salomaki J., Hinkkanen M., Luomi J. Sensorless Control of Induction Motor Drives Equipped With Inverter Output Filter // IEEE Trans. on Ind. Electronics. – 2006. – Vol. 53, No. 4. – Pp. 1188–1197.
4. Хойджен П. Фильтры для частотных инверторов // Chip News. – 2002. – № 8. – С. 58–60.
5. Чиликин М. Г., Ключев В. И., Сандлер А. С. Теория автоматизированного электропривода. – М.: Энергия, 1979. – 616 с.
6. Потапенко Е. М., Соломаха А. В., Савранская А. В. Экспериментальное исследование аналого-цифровых адаптивных фильтров трехфазных сигналов // Электромашинобудування та електрообладнання. – 2006. – № 66. – С. 271, 272.

Надійшла 17.10.07

Синтезовано алгоритм роботи попереджувального компенсатора, який усуває скривлення амплітуди та фази, що внесені пасивним фільтром. Чисельне та натурне моделювання підтвердило працездатність синтезованого алгоритма.

The predictive compensator algorithm, eliminating distortions of amplitude and phase, brought in the passive filter, is synthesized. Numerical and natural modeling of serviceability of the synthesized algorithm has confirmed.

УДК 681.52.136

В. Е. Цыганаш, В. М. Зверев, Ю. С. Белоиваненко

АНАЛИЗ ЗАДАЧИ ОПТИМАЛЬНОГО УПРАВЛЕНИЯ ПРОЦЕССОМ ПРЕОБРАЗОВАНИЯ ЭНЕРГИИ В ТУРБОАГРЕГАТЕ

Рассмотрены технологические аспекты и проблемы производства электроэнергии на теплоэлектроцентрали. Изложена концепция оптимального управления процессом преобразования энергии в турбоагрегате.

В настоящее время очень большую роль в повышении использования энергетических затрат играет удельный вес энергетической составляющей в себестоимости промышленной продукции. В отдельных отраслях промышленности доля комплексных энергетических

затрат в себестоимости промышленной продукции составляет более 50 %, в черной металлургии – до 40 %, в машиностроении – 20–30 %. Но и тогда, когда энергетическая составляющая в себестоимости незначительна, необходимо учитывать, что при экономном расходовании энергии появляется возможность выработки дополнительной продукции, а ущерб при недоотпуске энергии во много раз превышает ее стоимость. Следовательно, значение экономии энергии выходит далеко за рамки снижения себестоимости продукции, потери энергии – это бесполезно сгоревший уголь, газ и т. д.

Традиционно в потреблении энергии выделяют три составляющие: потребление электроэнергии, потребление теплоэнергии, сжигание топлива. Следует отметить, что в технологическом процессе большинства теплоэлектроцентралей (ТЭЦ) одновременно задействованы все три составляющие. Современная ТЭЦ представляет собой комплекс на основе различных динамических подсистем, выполняющих разнообразные технологические функции. Это подсистема производства тепловой энергии с помощью паровых котлов; подсистема преобразования тепловой энергии в механическую посредством турбоагрегатов; подсистема генерации электрической энергии посредством синхронных генераторов электрической энергии. Все подсистемы связаны между собой процессами интенсивного взаимодействия и обмена энергией, веществом и информацией.

В настоящее время разработка энергосберегающих технологий для таких систем ведется в двух направлениях. Первое направление связано с процессами регистрации, учета и документирования потребляемой энергии. Оно проработано в достаточной степени и на хорошем уровне, но не затрагивает динамических процессов, происходящих в системах. Второе направление посвящено локальным системам управления энергопотребителями. Оно почти исчерпало свои возможности. Как правило, энергоблоки работают в базовых режимах и практически не используют возможность работы в широком диапазоне нагрузок. В этой связи возникает важная проблема повышения коэффициента полезного действия (КПД) и маневренности агрегатов, решение которой для парогенерирующих установок означает расширение диапазона нагрузок и тепловых режимов, в рамках которой система управления котлом способна обеспечить его эффективную и устойчивую работу. При этом основное препятствие состоит в ограниченных возможностях существующих линейных систем управления энергоблоками. В связи с этим, наряду с заменой элементной базы регуляторов на более современную, основное внимание необходимо уделять совершенствованию принципов и методов управления объектами в целом. Прежде всего, это относится к программному управлению, оптимальному управлению и введению гибкого адаптивного управления. Перспективность этих направлений обусловлена

наличием экстремальной зависимости в процессах энергопреобразования и ее расположением не только на границе, но и в самой области допустимых значений переменных.

Главные проблемы разработок при таком подходе связаны с уникальностью решаемых задач, с изменчивостью задач, объектов управления, структуры системы и ее математического обеспечения, широкими пределами измерений; большим числом датчиков, высоким быстродействием. Для решения задач оптимального управления необходимо знать закономерности, определяющие ход оптимизирующего процесса, его математическую модель. Однако, создание самой модели, синтез уравнений, определяющих множество допустимых решений, в свою очередь, приводят к необходимости решения оптимизационных задач.

Целью настоящей работы является анализ одной из подсистем для поиска новых алгоритмов управления, которые можно было бы распространить на все подсистемы и которые бы обеспечили в дальнейшем надежную генерацию электроэнергии необходимого качества с одновременной минимизацией расхода энергоносителя и, следовательно, себестоимости электроэнергии.

Для реализации подобного подхода удобно выбрать турбоагрегат, преобразующий тепловую энергию в механическую и представить его, в свою очередь, состоящим из двух подсистем – «источника питания» (ИП) и «нагрузки» (Н). Взаимодействие между энергоснабжающей частью (ИП) и энергопотребляющей частью (Н) можно охарактеризовать «сложными понятиями», такими как КПД, напор, импеданс, резонанс и т. п. Основная функция этой системы – энергетические преобразования, которые должны учитывать всю энергию, потерянную или обмененную в турбоагрегате.

Отметим, что «сложные понятия» образуют связь между основными физическими закономерностями и практической задачей. Они могут не учитывать некоторые детали, но должны дать основу для анализа. Основным критерием их выбора является не истинность или строгость, а полезность [1]. Они являются ключевыми элементами математической модели (теории) и, кроме того, позволяют дать сжатую механистическую интерпретацию результатам измерений. Они служат каркасом, поддерживающим более конкретные и подробные знания об объекте [1]. Пользуясь такими качественными понятиями, сформулируем задачу оптимального управления.

При заданном расходе энергоносителя (пара) требуется обеспечить максимальное давление в камере турбины, если же давление в камере необходимо поддерживать постоянным, то расход энергоносителя должен быть минимальным. Для перехода к математической постановке задачи введем обозначения для переменных, характеризующих процесс. С этой целью рассмотрим упрощенную схему турбоагрегата (рис. 1).

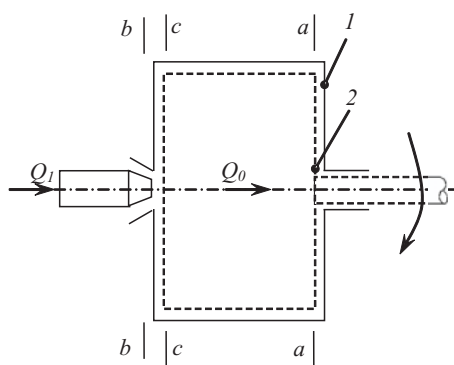


Рисунок 1 – Упрощенная схема турбоагрегата:

1 – корпус турбины; 2 – турбина

Сущность его работы основана на передаче энергии от одной среды (рабочей, в данном случае это пар), движущейся с большой скоростью, к другой среде, обладающей значительно меньшим запасом энергии. Передача энергии происходит в процессе такого взаимодействия, при котором поток пара при входе на лопасти рабочего колеса турбины имеет избыток давления, по сравнению с потоком, выходящим из турбины. При движении пара через криволинейные межлопастные каналы поток оказывает реактивное давление на лопасти, приводя во вращение рабочий вал турбины. При этом кинетическая энергия рабочей среды частично превращается в механическую, а частично теряется вследствие значительного вихреобразования, потерь на трение и потерь в элементах турбины, подводящих и отводящих пар.

Допустив для упрощения анализа, что плотность смешиваемых потоков одинакова, для основных параметров подсистем ИП-Н, можно записать [2]:

1) рабочий напор, затрачиваемый в турбоагрегате

$$H_p = \frac{P_b}{\rho g} + \frac{V_b^2}{2g} - \frac{P_c}{\rho g} - \frac{V_c^2}{2g};$$

2) полезный напор, создаваемый в области (с-с) входа пара в лопасти рабочего колеса турбины и равный разности этого напора и напора области (а-а) выхода пара с лопастей рабочего колеса турбины

$$H_n = \frac{P_c}{\rho g} + \frac{V_c^2}{2g} - \frac{P_a}{\rho g} - \frac{V_a^2}{2g};$$

3) расход пара

$$Q_1 = V_1 S_1;$$

4) полезная подача (эквивалентная величина, определяемая при пересчете из получаемой механической мощности на валу турбины):

$$Q_0 \cong V_0 S_0.$$

В приведенных зависимостях принято:

P_b, P_c, P_a – давление пара на выходе из сопла котлоагрегата, входе на лопасти рабочего колеса и выходе с лопастей рабочего колеса турбины, соответственно;

V_b, V_c, V_a – скорости потока пара на выходе из сопла котлоагрегата, области входа в лопасти рабочего колеса и области выхода с лопастей рабочего колеса турбины;

V_1, V_0 – скорости в рабочем сопле котлоагрегата и турбины;

S_1, S_0 – площадь рабочего сопла и эквивалентная площадь сечения турбины;

ρg – удельный вес единицы объема потока пара, представленный через его плотность ρ и ускорение силы тяжести g .

Учитывая принятые обозначения и традиционный подход, при котором энергетическая эффективность любого процесса передачи или преобразования энергии оценивается коэффициентом полезного действия, выберем в качестве критерия оптимальности КПД. Он для турбоагрегата равен отношению полезной мощности к затраченной:

$$\eta = \frac{H_n Q_0}{H_p Q_1}. \quad (1)$$

Выбор этого критерия приводит к тому, что при выделении множества допустимых значений переменных приходится учитывать три типа ограничений: ограничения, наложенные на вектор режимных и конструктивных параметров, ограничения на вектор, характеризующий параметры рабочего потока пара, а также зависимость, характеризующую полный баланс энергии.

Если первые два типа переменных в реальных условиях учитываются сравнительно просто, то уравнение связи между энергетическими преобразованиями в удобном для практического применения виде представить сложно. Поэтому за основу удобно взять характеристику турбоагрегата, описывающую его работу при переменных режимах. Ее можно получить при условии $H_n + H_p = \text{const}$, близком к типичному случаю эксплуатации турбоагрегата.

Главная особенность характеристики в том, что она имеет выраженный максимум в зоне, где сумма потерь смешения и потерь в тракте преобразования энергии минимальна. В этих условиях можно воспользоваться методом косвенной оптимизации с применением модели, выполненной в форме экстремального принципа и инвариантной к системе координат, в которой рассматривается моделируемый процесс. Как показывает практика, при этом не требуется высокой точности модели. Поэтому представляется целесообразным применение модели, воспроизводящей свойства реальной системы ИП-Н не во всей допустимой области протекания процесса, а только в небольшой ее части, включающей и оптимальную траекторию. В этом случае

схема, имитирующая модель, должна обладать избирательностью и фазовой селективностью. Сочетания этих качеств можно достичь при использовании устройства с управляемыми колебательными контурами высокой добротности [3].

Прежде чем анализировать другие компоненты системы управления, использующей модель [3], выделим доминирующие переменные, которые затем можно использовать в качестве управляющих воздействий, и преобразуем критерий оптимального управления к более рациональному виду.

Если первое условие выполнить не трудно, поскольку ключевыми переменными процесса являются давление и расход энергоносителя, то второе условие не является столь очевидным. Согласно уравнению (1) максимальное значение КПД достигается при максимуме полезной мощности и минимуме затраченной. Практической реализации этого критерия в системе управления препятствует противоречивость условий и трудности оперативного измерения некоторых составляющих мощности. Поэтому, учитывая, что $\frac{P_a}{\rho g} \ll \frac{P_c}{\rho g}$

и пренебрегая разностью $\frac{V_c^2}{2g} - \frac{V_a^2}{2g}$ из-за малости этой разности, запишем

$$H_n \cong \frac{P_c}{\rho g} - \frac{P_a}{\rho g} = \frac{P_c}{\rho g} \left(1 - \frac{P_a}{P_c} \right).$$

Воспользовавшись возможностью представления $\left(1 - \frac{P_a}{P_c} \right)$ в форме биномиального ряда, перестроим выражение для H_n :

$$H_n \cong \frac{P_c}{\rho g} \frac{1}{1 + \frac{P_a}{P_c}}.$$

Если учесть, что в установившемся режиме работы турбоагрегата Q_0 и H_p изменяются мало и эти изменения одного знака, поскольку увеличение (уменьшение) H_p сопровождается увеличением (уменьшением) Q_0 , то и изменение $\frac{Q_0}{H_p}$ будет малым. Поэтому решающее влияние на КПД будет оказывать отношение

$$\frac{H_n}{Q_1} = \frac{P_c}{\rho g} \frac{1}{Q_1 \left(1 + \frac{P_a}{P_c} \right)}.$$

Из этого отношения следует, что максимум КПД достигается при выполнении условия

$$\frac{\rho g}{P_c} \left(1 + \frac{P_a}{P_c} \right) Q_1 \rightarrow \min.$$

Использование критерия оптимальности в такой форме предпочтительнее, так как он сравнительно просто вычисляется, поскольку все переменные, входящие в формулу, допускают непосредственное измерение. Особенно важно то, что вычисления можно вести в реальном времени в течение хода технологического процесса с помощью аналоговой модели [3].

Получив таким путем информацию об оптимальном решении, управлять процессом поддержания механической мощности можно, регулируя расход пара на входе в турбину, изменяя противодействие на выходе или комбинируя оба варианта. Первый вариант управления требует воздействия на мощный энергетический поток, что дорого, и, к тому же, сопровождается значительными дополнительными потерями энергии. Второй вариант не имеет недостатков, отмеченных в первом варианте, но характеризуется малым диапазоном изменения управляющего воздействия. Поэтому для управления более всего подходит третий вариант, когда расход пара на входе устанавливается примерно достаточным для решения технологической задачи, а достижение максимума КПД обеспечивается коррекцией противодействия.

Принципиальная возможность изменения КПД противодействием была экспериментально обоснована в работе [4]. Подтверждено существование оптимальной «зоны смещения» и доказана возможность изменения ее положения. При этом может достигаться максимум КПД, а экономия энергоносителя может составить 8–15 %.

ПЕРЕЧЕНЬ ССЫЛОК

1. Каннинггэм Р. Сжатие газа с помощью жидкоструйного насоса // Труды американского общества инженеров-механиков. Теоретические основы инженерных расчетов. – М.: Мир, 1974. – № 3. – С. 112–128.
2. Цыганаш В. Е., Волошин А. И. Постановка задачи оптимального управления процессом создания и поддержания вакуума с помощью струйного насоса // Вестник национального технического университета «ХПИ». Новые решения в современных технологиях. Сборник научных трудов. – 2001. – № 14. – С. 206–212.
3. Цыганаш В. Е. Метод моделирования электрического режима индукционной электротермической установки // Удосконалення процесів та обладнання обробки тиском в машинобудуванні і металургії: Зб. наук. пр. – Краматорськ-Слов'янськ, 2000. – С. 316–319.
4. Каннинггэм Р., Допкин Р. Длины участка разрушения струи и смешивающей горловины жидкоструйного насоса для перекачки газа // Труды американского общества инженеров-механиков. Теоретические основы инженерных расчетов. – М.: Мир, 1974. – № 3. – С. 128–141.

Надійшла 1.03.07

Розглянуті технологічні аспекти та проблеми виробництва електроенергії на теплоелектроцентралі. Викладена концепція оптимального управління процесом енергоперетворення в турбоагрегаті.

Technological aspects and problems of manufacture of the electric power on thermal power main lines are considered. The concept of optimum control by energy transformation process in a turbine unit is stated.

АВТОРЫ НОМЕРА

Архипова Елена Александровна

студент магистратуры Национального
технического университета Украины
«КПИ»

Афонин Юрий Сергеевич

аспирант Запорожского национального
технического университета

Белованенко Юрий Сергеевич

ассистент Донбасской государственной
машиностроительной академии, г. Кра-
маторск

Берко Андрей Юлианович

кандидат технических наук, доцент
национального университета «Львов-
ская политехника»

Бодянский Евгений Владимирович

доктор технических наук, профессор
Харьковского национального универси-
тета радиоэлектроники

Высоцкая Виктория Анатольевна

ассистент национального университета
«Львовская политехника»

Гапоненко Николай Прокофьевич

кандидат технических наук, доцент
Запорожского национального техниче-
ского университета

Герасин Сергей Николаевич

доктор технических наук, профессор
Харьковского национального универси-
тета радиоэлектроники

Гороховатский Владимир Алексеевич

кандидат технических наук, доцент
Харьковского национального универси-
тета радиоэлектроники

Гостев Владимир Иванович

доктор технических наук, заведующий
кафедрой Государственного универси-
тета информационно-коммуникационных
технологий, г. Киев

Гулин Сергей Петрович

кандидат технических наук, доцент За-
порожского национального техниче-
ского университета

Давыдов Андрей Андреевич

студент Харьковского национального
университета радиоэлектроники

Дейнега Лариса Юрьевна

ассистент Запорожского национального
технического университета

Долгов Виктор Иванович

доктор технических наук, профессор
Харьковского национального универси-
тета радиоэлектроники

Дубровин Валерий Иванович

кандидат технических наук, доцент За-
порожского национального техниче-
ского университета

Душинова Евгения Викторовна

студентка Запорожского национального
технического университета

Дьячук Татьяна Сергеевна

аспирант Запорожского национального
технического университета

Егошина Анна Анатольевна

аспирант Государственного универси-
тета информатики и искусственного
интеллекта, г. Донецк

Журавлев Владимир Николаевич

кандидат технических наук, докторант
Национального технического универси-
тета Украины «КПИ»

Заец Алексей Владимирович

ассистент Черниговского государствен-
ного технологического университета

Захарченко Александр Алексеевич

ведущий инженер-исследователь ННЦ
«Харьковский физико-технический ин-
ститут»

Зверев Михаил Владимирович

начальник цеха Краматорской ТЭЦ

Ильяшенко Матвей Борисович

аспирант Запорожского национального
технического университета

Казурова Алина Евгеньевна

аспирант Запорожского национального
технического университета

Киприч Татьяна Витальевна

аспирант Запорожского национального
технического университета

Киричек Галина Григорьевна

старший преподаватель Запорожского
национального технического универси-
тета

Козлов Максим Александрович

аспирант Харьковского национального
университета радиоэлектроники

Кудерметов Равиль Камирович

кандидат технических наук, заведую-
щий кафедрой Запорожского нацио-
нального технического университета

Кузьмин Сергей Владимирович

аспирант Харьковского национального
университета радиоэлектроники

Кулишова Нонна Евгеньевна

кандидат технических наук, доцент
Харьковского национального универси-
тета радиоэлектроники

Курапов Сергей Всеволодович

кандидат физико-математических наук,
доцент Запорожского национального
университета

Кутний Владимир Евдокимович

кандидат технических наук, начальник
лаборатории ННЦ «Харьковский физи-
ко-технический институт»

Левыкина Виктория Ивановна

ассистент Запорожского национального
технического университета

Липчанский Александр Иванович

кандидат технических наук, докторант
Харьковского национального универ-
ситета радиоэлектроники

Лысенко Елена Валериевна

студентка Одесского национального по-
литехнического университета

Лысенко Татьяна Владимировна

кандидат технических наук, доцент
Одесского национального политехни-
ческого университета

Манило Татьяна Владимировна

студентка Запорожского национального
технического университета

Мусяенко Максим Павлович

доктор технических наук, профессор
Черкасского государственного техноло-
гического университета

Неласая Анна Викторовна

старший преподаватель Запорожского
национального технического универси-
тета

Никитина Татьяна Борисовна

кандидат технических наук, доцент
Национального технического универси-
тета «ХПИ»

Новотарский Михаил Анатольевич

кандидат технических наук, старший
научный сотрудник Института матема-
тики НАН Украины, г. Киев

Носенко Татьяна Ивановна

соискатель Одесского национального
политехнического университета

Огренич Евгений Викторович

аспирант Запорожского национального
технического университета

Панченко Игорь Вячеславович

адъюнкт Военного института телеком-
муникаций и информатизации Нацио-
нального технического университета
Украины «КПИ»

:

Пархоменко Лариса Александровна

кандидат физико-математических наук,
доцент Запорожского национального
технического университета

Пица Дмитрий Макарович

доктор технических наук, проректор
Запорожского национального техничес-
кого университета

Потапенко Евгений Михайлович

доктор технических наук, профессор За-
порожского национального техничес-
кого университета

Прохорец Иван Михайлович

кандидат физико-математических наук,
старший научный сотрудник ННЦ
«Харьковский физико-технический ин-
ститут»

Псарев Валентин Иванович

доктор физико-математических наук,
профессор

Пьянкова Юлия Владимировна

ассистент Запорожского национального
технического университета

Руденко Олег Григорьевич

доктор технических наук, заведующий
кафедрой Харьковского национального
университета радиоэлектроники

Рыбка Александр Викторович

кандидат физико-математических наук,
старший научный сотрудник ННЦ
«Харьковский физико-технический ин-
ститут»

Савельев Александр Викторович

старший научный сотрудник Уфимского
государственного авиационного техни-
ческого университета

Скуртов Сергей Николаевич

кандидат технических наук, старший
научный сотрудник корпорации «Ин-
ком», г. Киев

Соломаха Александр Викторович

аспирант Запорожского национального
технического университета

Становский Александр Леонидович

доктор технических наук, заведующий
кафедрой Одесского национального по-
литехнического университета

Ткач Юрий Эдуардович

ассистент Черниговского государствен-
ного технологического университета

Томенко Виталий Иванович

младший научный сотрудник Черкас-
ского государственного технологичес-
кого университета

Хаджмурадов Манап Ахмадович

доктор технических наук, начальник от-
дела ННЦ «Харьковский физико-техни-
ческий институт»

Цыганаш Виктор Евграфович

кандидат технических наук, доцент
Донбасской государственной машино-
строительной академии, г. Краматорск

Чернобородов Михаил Петрович

инженер-конструктор КП НПК «Искра»,
г. Запорожье

Чернобородова Наталья Петровна

старший преподаватель Запорожского
национального технического универси-
тета

ДЛЯ СВЕДЕНИЯ АВТОРОВ

Журнал «Радиоэлектроника. Информатика. Управління» предназначен для публикации наиболее значимых научных и практических результатов исследований ученых высших учебных заведений и научных организаций.

Журнал включен в перечень научных изданий Украины, в которых могут публиковаться результаты диссертационных работ на соискание ученых степеней доктора и кандидата технических наук.

Статьи, опубликованные в журнале, реферируются в реферативных журналах и базах данных ВИНТИ (Россия) и «Джерело» (Украина). Рефераты статей размещаются на web-сайте Запорожского национального технического университета (<http://www.zntu.edu.ua/base/news2005/ric.htm>).

Журнал издается два раза в год и распространяется по подписке в Украине. Подписной индекс 22914.

Последовательность размещения материала статьи: индекс УДК, название статьи, инициалы и фамилия автора(ов), аннотация, текст, перечень ссылок. Рукопись статьи должна быть подписана всеми авторами.

Текст статьи

Принимаются статьи на русском, украинском и английском языках. Размер статьи до 0,5 авторского листа.

Структура статьи

Редакционная коллегия принимает к рассмотрению только научные статьи, которые содержат такие необходимые элементы: постановка проблемы в общем виде и ее связь с важными научными или практическими заданиями; анализ последних исследований и публикаций, в которых начато решение данной проблемы, и на которые опирается автор; выделение нерешенных ранее частей общей проблемы, которым посвящается статья; формулирование целей статьи (постановка задачи); изложение основного материала исследования с полным обоснованием полученных научных результатов, выводы по данному исследованию и перспективы дальнейших исследований в данном направлении.

Требования к оформлению

Статьи принимаются набранные в редакторе Word for Windows (v.6 и выше).

Параметры страницы:

- размер бумаги – А4 (210x297);
- ориентация – книжная;
- шрифт – Times New Roman Cyr, размер – 12 pt;
- междустрочный интервал – полуторный;
- верхнее поле – 20мм, нижнее – 20 мм, левое – 25 мм, правое – 15 мм.

Аннотация

Аннотация является обязательной и подается в начале статьи на русском, украинском (для граждан Ук-

раины) и английском языках. Объем аннотации не должен превышать 40 слов.

Рисунки

Рисунки помещаются в тексте и подаются в отдельных файлах (формат.TIF с разрешением не менее 200 dpi). Рисунки нумеруют и подписывают внизу.

Таблицы

Заголовки таблиц размещаются по центру.

Формулы

Формулы выполняются с помощью встроенного в Word for Windows редактора Microsoft Equation. Формулы нумеруют в круглых скобках справа:

$$Z(\Theta) = 101 \log \left(\frac{\bar{y}^2}{s^2} \right) \quad (3)$$

Формулы большого размера записываются в несколько строк.

Перечень ссылок

Перечень ссылок в конце рукописи подается языком оригинала и составляется в порядке упоминания ссылок в тексте и в соответствии со стандартом на библиографическое описание. Ссылки на литературу в тексте обозначаются цифрой в квадратных скобках.

В сведениях об авторах необходимо привести фамилию, имя, отчество (полностью), место работы, должность, ученую степень, рабочий и домашний адрес, телефоны, e-mail.

В редакцию журнала необходимо представить:

- 1) распечатанный текст статьи в 2-х экземплярах;
- 2) экспертное обоснование о возможности опубликования;
- 3) сведения об авторах;
- 4) сопроводительное письмо-ходатайство из организации, где была выполнена работа;
- 5) название статьи, фамилии, имена и отчества авторов на английском языке;
- 6) дискета 3,5' с текстом документа, файлами рисунков, данными о статье на английском языке и сведениями об авторах в формате Word for Windows (v.6 и выше). Файлы можно выслать электронной почтой в виде заархивированных (ZIP-архиватором) файлов.

Статьи, которые не отвечают указанным требованиям, не принимаются к рассмотрению.

Рукописи и дискеты не возвращаются.

Адрес редакции: 69063, г. Запорожье, ул. Жуковского, 64, ЗНТУ, редакция журнала "PIU"

Тел.:

(061) 7-698-2-96 – редакционно-издательский отдел

(061) 7-698-3-13 – главный редактор

(061) 7-698-2-96 – зам. главного редактора

Факс: (0612) 64-21-41

E-mail: rvv@zntu.edu.ua, TQM@zntu.edu.ua

Наукове видання

**Радіоелектроніка.
Інформатика.
Управління**

№ 2/2007

Науковий журнал

Головний редактор – д. т. н. Потапенко Є. М.
Заст. головного редактора – к. т. н. Дубровін В. І.

Комп'ютерне моделювання та верстання

Новікова К. О.

Рибалка І. С.

Оригінал-макет підготовлено у редакційно-видавничому відділі ЗНТУ

*Підписано до друку 4.02.2008. Формат 60x84/8.
Папір офс. Різогр. друк. Ум. друк. арк. 18,6.
Тираж 300 прим. Зам. № 249.*

69063 м. Запоріжжя, ЗНТУ, друкарня, вул. Жуковського, 64