

Запорізький національний технічний університет



Радіоелектроніка Інформатика Управління

1(17)'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., Соболев 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 № 10 від 29.05.2007 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)7-698-2-96 – редакційно-видавничий відділ

Факс: (0612)64-21-41

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

ЗМІСТ

РАДІОФІЗИКА	5
<i>В. В. Ващенко, В. П. Чумаченко</i> РЕШЕНИЕ МЕТОДОМ ПРОИЗВЕДЕНИЯ ОБЛАСТЕЙ N-ПЛОСКОСТНОЙ ЗАДАЧИ ДИФРАКЦИИ ВОЛН НА НАКЛОННОЙ ГРАНИЦЕ РАЗДЕЛА ДИЭЛЕКТРИЧЕСКИХ СРЕД В ПРЯМОУГОЛЬНОМ ВОЛНОВОДЕ	5
<i>А. В. Крапивной, В. П. Чумаченко, Я. В. Чумаченко</i> О РАССЕЯНИИ ПЛОСКОЙ СКАЛЯРНОЙ ВОЛНЫ ИМПЕДАНСНОЙ СФЕРОЙ	9
РАДИОЕЛЕКТРОНИКА ТА ТЕЛЕКОМУНІКАЦІЇ	15
<i>Б. А. Артюшенко</i> ПАРАМЕТРИЧНИЙ СИНТЕЗ ЛІНІЙНОЇ АНТЕННОЇ РЕШІТКИ ПРИ ЗАДАНИХ ДОПУСКАХ ЗА ДОПОМОГОЮ ГЕНЕТИЧНОГО АЛГОРИТМУ	15
<i>А. А. Никонова, Д. И. Левинзон, О. Ю. Небеснюк</i> ТЕХНОЛОГИЧЕСКИЕ ОСОБЕННОСТИ ИЗГОТОВЛЕНИЯ КОНТАКТНЫХ СИСТЕМ К ПОЛУПРОВОДНИКОВЫМ СТРУКТУРАМ....	34
<i>А. Ю. Воронай, В. Н. Кришук, Л. М. Карпуков, А. Ю. Фарафонов</i> ДОПУСКОВЫЙ АНАЛИЗ КОНСТРУКЦИЙ ФИЛЬТРОВ С НЕМОНОТОННОСТЬЮ АЧХ	18
<i>Ю. К. Тараненко</i> ИСПОЛЬЗОВАНИЕ ПАРАМЕТРИЧЕСКОГО КОНТУРА РЕГУЛИРОВАНИЯ АМПЛИТУДЫ В АВТОКОЛЕБАТЕЛЬНЫХ СИСТЕМАХ ВИБРОЧАСТОТНЫХ ПЛОТНОМЕРОВ С МЕХАНИЧЕСКИМ РЕЗОНАТОРОМ.....	37
<i>Н. П. Гапоненко, Е. В. Огренич</i> ПРОЕКТИРОВАНИЕ ПЛАСТИНЧАТЫХ РАДИАТОРОВ С ОПТИМАЛЬНЫМИ МАССОГАБАРИТНЫМИ ПАРАМЕТРАМИ	22
<i>Г. М. Шило, Д. А. Коваленко</i> ОСОБЛИВОСТІ КОМПЕНСАЦІЇ ЗОВНІШНІХ ВПЛИВІВ ПРИ ПРИЗНАЧЕННІ НОРМАЛЬНИХ ДОПУСКІВ	44
<i>А. О. Намлинський</i> ІНТЕРВАЛЬНІ КОЕФІЦІЄНТИ ЗОВНІШНІХ ВПЛИВІВ В ЗАДАЧАХ ВИБОРУ ЕЛЕМЕНТІВ ...	26
МАТЕМАТИЧНЕ ТА КОМП'ЮТЕРНЕ МОДЕЛЮВАННЯ	48
<i>А. И. Вершина, Т. Н. Семерюк, Б. Т. Солдатов</i> ПРИМЕНЕНИЕ ЦЕПЕЙ МАРКОВА ПРИ МОДЕЛИРОВАНИИ УЧЕБНОГО ПРОЦЕССА	48
<i>И. М. Килимник, Д. С. Ярымбаш</i> ПОВЫШЕНИЕ ЭФФЕКТИВНОСТИ ОБРАБОТКИ ИНФОРМАЦИИ ПРИ РЕГИСТРАЦИИ КОНТРОЛИРУЕМЫХ ПАРАМЕТРОВ В СИСТЕМАХ АВТОМАТИЧЕСКОГО УПРАВЛЕНИЯ	68
<i>С. Н. Герасин, Е. В. Слипченко</i> МЕТОДЫ ОЦЕНКИ РЕНТАБЕЛЬНОСТИ ФИНАНСОВОЙ ДЕЯТЕЛЬНОСТИ ИЕРАРХИЧЕСКИХ БИЗНЕС-СТРУКТУР	52
<i>М. К. Демин</i> МЕТОДЫ ПРЕДСТАВЛЕНИЯ ИЕРАРХИЧЕСКОЙ ИНФОРМАЦИИ В РЕЛЯЦИОННЫХ БАЗАХ ДАНЫХ ИНФОРМАЦИОННО- УПРАВЛЕНЧЕСКИХ АРХИТЕКТУР	56
<i>Г. П. Коломеец</i> ИССЛЕДОВАНИЕ ВОЗМОЖНОСТЕЙ ПРОГРАММЫ «АНАЛИЗАТОР СЕТЕЙ ETHERNET» ДЛЯ МОДЕЛИРОВАНИЯ ПРОЦЕССА ВОЗНИКНОВЕНИЯ КОЛЛИЗИЙ В КОМПЬЮТЕРНЫХ СЕТЯХ ETHERNET	73
<i>М. Б. Ильяшенко</i> АЛГОРИТМ НАХОЖДЕНИЯ ГРАФ-ПОДГРАФ ИЗОМОРФИЗМА ДЛЯ ВЗВЕЩЕННЫХ ГРАФОВ И ЕГО ПРИМЕНЕНИЕ	62
<i>М. В. Михайлова</i> ПОРІВНЯННЯ АЛГОРИТМІВ НЕЧІТКОГО ПОШУКУ В ТЕКСТАХ УКРАЇНСЬКОЮ МОВОЮ	77

<i>М. Ю. Сенашова, А. Г. Рубцов, М. Г. Садовский</i> КИНЕТИЧЕСКАЯ МАШИНА КИРДИНА И ЗАДАЧА ВОССТАНОВЛЕНИЯ УТЕРЯННЫХ ДАННЫХ	<i>А. Л. Становский, Т. В. Лысенко, Н. П. Худенко</i> УПРАВЛЕНИЕ СОСТОЯНИЯМИ В СЛОЖНЫХ ТЕХНИЧЕСКИХ ОБЪЕКТАХ	87	93
ПРОГРЕСИВНІ ІНФОРМАЦІЙНІ ТЕХНОЛОГІЇ.....		97	
<i>Н. В. Алипов, Н. М. Кораблев, М. И. Хиль, М. В. Гусятин</i> ПРИМЕРЫ ПОСТРОЕНИЯ ОРИЕНТИРОВАННЫХ ГРАФОВ ПЕРЕХОДОВ ЦИФРОВОГО АВТОМАТА С ПСЕВДОСЛУЧАЙНЫМИ ПЕРЕХОДАМИ	<i>В. И. Дубровин, Ю. С. Афонин, Р. Н. Зинченко, С. С. Емельяненко, Т. В. Манило</i> ПОДХОД К ПОСТРОЕНИЮ МОДЕЛИ ПО ЗВУКОВОМУ СИГНАЛУ С ИСПОЛЬЗОВАНИЕМ ВЕЙВЛЕТ- ПРЕОБРАЗОВАНИЯ	97	112
<i>А. О. Говоров, Г. И. Никулищев, Г. Л. Козина</i> АНАЛИЗ СТАНДАРТОВ ЭЛЕКТРОННОЙ ЦИФРОВОЙ ПОДПИСИ	<i>І. О. Романенко, С. В. Алексєєв, І. В. Рубан, С. А. Войтович, В. В. Калачова</i> ДИСТАНЦІЙНЕ НАВЧАННЯ В ОСВІТНЬОМУ ПРОЦЕСІ ЗБРОЙНИХ СИЛ УКРАЇНИ	105	115
<i>А. С. Горецкий</i> МУЛЬТИАГЕНТНЫЕ ТЕХНОЛОГИИ В СИСТЕМАХ УПРАВЛЕНИЯ ПЕРСОНАЛОМ.....	<i>С. С. Ющенко</i> ІНФОРМАЦІЙНО-ПРАВОВІ СИСТЕМИ:РОСІЙСЬКИЙ ДОСВІД	108	119
ТЕОРІЯ І МЕТОДИ АВТОМАТИЧНОГО УПРАВЛІННЯ		124	
<i>Н. М. Ветрова</i> ЭКОЛОГИЧЕСКОЕ УПРАВЛЕНИЕ: ВИДЫ И ОСОБЕННОСТИ СТРАТЕГИЙ ЭКОЛОГИЧЕСКОЙ НАПРАВЛЕННОСТИ	<i>Е. А. Шушляпин, Е. В. Виноградов</i> ОБОБЩЕНИЕ МЕТОДА КОНЕЧНОГО СОСТОЯНИЯ НА НЕЛИНЕЙНЫЕ СИСТЕМЫ С ПЕРЕМЕННЫМИ ЗАПАЗДЫВАНИЯМИ	124	146
<i>И. А. Орловский, А. А. Синяевский</i> РАЗРАБОТКА МОДЕЛЕЙ НЕЛИНЕЙНЫХ ЭЛЕКТРОТЕХНИЧЕСКИХ ОБЪЕКТОВ В ВИДЕ СТЕПЕННЫХ РЕКУРРЕНТНЫХ НЕЙРОННЫХ СЕТЕЙ.....	<i>И. В. Щербань, В. А. Бертенев</i> НЕПАРАМЕТРИЧЕСКАЯ СТОХАСТИЧЕСКИХ ИДЕНТИФИКАЦИЯ НЕЛИНЕЙНЫХ СИСТЕМ НА ОСНОВЕ КВАДРАТИЧНОГО КРИТЕРИЯ ...	128	149
<i>А. А. Худяев</i> ВЛИЯНИЕ НЕСЛУЧАЙНОЙ СОСТАВЛЯЮЩЕЙ ПОЛЕЗНОГО СИГНАЛА НА ТОЧНОСТЬ ИТЕРАЦИОННОЙ ДВУХКАНАЛЬНОЙ ВОСПРОИЗВОДЯЩЕЙ СИСТЕМЫ ПРИ СЛУЧАЙНЫХ ВОЗМУЩЕНИЯХ		137	
АВТОРЫ НОМЕРА		154	

РАДІОФІЗИКА

РАДИОФИЗИКА

RADIOPHYSICS

УДК 621.372.8

В. В. Ващенко, В. П. Чумаченко

РЕШЕНИЕ МЕТОДОМ ПРОИЗВЕДЕНИЯ ОБЛАСТЕЙ H -ПЛОСКОСТНОЙ ЗАДАЧИ ДИФРАКЦИИ ВОЛН НА НАКЛОННОЙ ГРАНИЦЕ РАЗДЕЛА ДИЭЛЕКТРИЧЕСКИХ СРЕД В ПРЯМОУГОЛЬНОМ ВОЛНОВОДЕ

Рассматривается известная задача дифракции электромагнитных волн на прямолинейной границе раздела диэлектрических сред. Решение находится с помощью недавно предложенного варианта метода произведения областей, который использует определенного вида тригонометрические разложения для представления искомой функции внутри выпуклых многогранников. Путем сравнения полученных результатов с данными других авторов, установлены применимость и эффективность этого подхода при исследовании волноводных устройств, содержащих наклонную границу раздела сред.

ВВЕДЕНИЕ

Применение метода произведения областей к решению двумерных внешних и внутренних задач дифракции волн обычно связывается с возможностью представления произвольно-многоугольной области определения поля в виде общей части (произведения) простых базовых областей, каждая из которых является собой всю плоскость вне определенного звена граничного контура. В такой области уравнение Гельмгольца допускает разделение переменных в эллиптической системе координат, что приводит к построению искомой функции в виде разложений по функциям Матье (см., например, [1, 2]).

Применение этого метода приводит к надежным и устойчивым вычислительным алгоритмам. Следует, однако, заметить, что расчет значений функций Матье

является более сложной задачей, чем расчет значений других специальных функций, а соответствующее программное обеспечение не является общедоступным. Это заметно суживает круг возможных пользователей названного метода.

В работе [3], основываясь на основной идее метода произведения областей, была предложена более простая форма представления решения уравнения Гельмгольца для частного (но все же весьма общего) вида многоугольной области определения – произвольного выпуклого многоугольника. Было показано, что, рассматривая этот объект в виде пересечения полуплоскостей, решение можно представить в виде определенного типа тригонометрических разложений. Предполагалось, что граничные задачи для более сложных многоугольных областей могут быть решены путем их расчленения на выпуклые многоугольники и последующего применения метода сшивания.

Эта идея была эффективно использована для решения ряда волноводных задач, в том числе включающих границы раздела диэлектриков [4–6]. В работах [4–6] прямоугольные границы раздела расположены перпендикулярно стенкам волноводов. В настоящей работе метод применяется для решения задачи дифракции H_{m0} -волн на наклонной границе раздела сред в H -плоскости прямоугольного волновода. Отметим, что рассматриваемая конфигурация является хорошо

изученной, и исследовалась в прошлом многими авторами с помощью различных методов [7–11]. Исчерпывающее решение этой задачи было дано в работе [12] методом полуобращения.

Целью нашей работы является проверка на известной задаче применимости и эффективности развиваемого подхода при наличии наклонной границы раздела с тем, чтобы в последующем расширить область его приложений на более сложные конфигурации.

ПОСТАНОВКА ЗАДАЧИ И ПРЕДСТАВЛЕНИЕ ИСХОДНОЙ ФУНКЦИИ

Рассматриваемая структура изображена на рис. 1. Плоскость раздела двух сред с относительными диэлектрическими проницаемостями ϵ_1 и ϵ_2 перпендикулярна широким стенкам и наклонена под углом α к его оси. Слева на неоднородность набегают H_{m0} -волны единичной амплитуды. Задача заключается в нахождении z -компоненты электрического поля $E_z = u \cdot e^{i\omega t}$. Искомая функция u удовлетворяет уравнению Гельмгольца

$$\Delta u + k_j^2 \cdot u = 0, \tag{1}$$

где $k_j^2 = k_1^2 = k_0^2 \cdot \epsilon_1$ слева от границы раздела, $k_j^2 = k_2^2 = k_0^2 \cdot \epsilon_2$ справа от нее, $k_0 = 2\pi/\lambda$, а λ является длиной волны в свободном пространстве.

Обозначим u как $u^{(I)}$ в области I, $u^{(1)}$ в области 1, $u^{(2)}$ в области 2 и $u^{(II)}$ в области II.

После этого граничные условия для функции u могут быть записаны в виде:

$$u^{(I)} = u^{(1)}, \quad \frac{\partial u^{(I)}}{\partial x} = \frac{\partial u^{(1)}}{\partial x} \quad \text{при } x = -l, \tag{2}$$

$$u^{(2)} = u^{(II)}, \quad \frac{\partial u^{(2)}}{\partial x} = \frac{\partial u^{(II)}}{\partial x} \quad \text{при } x = l, \tag{3}$$

$$u^{(1)} = 0 \quad \text{при } y = a, \tag{4}$$

$$u^{(2)} = 0 \quad \text{при } y = -a, \tag{5}$$

$$u^{(1)} = u^{(2)}, \quad \frac{\partial u^{(1)}}{\partial y} = \frac{\partial u^{(2)}}{\partial y} \quad \text{при } \tilde{y} = 0. \tag{6}$$

Запишем теперь представления функции u во всех областях, введенных на рис. 1. Для регулярных участков волноводов мы, как обычно, получим:

$$u^{(I)} = \varphi_m(y, a) \cdot e^{-\gamma_m^{(I)} \cdot (x+l)} + \sum_{n=1}^{\infty} R_n \cdot \varphi_n(y, a) \cdot e^{\gamma_n^{(I)} \cdot (x+l)}, \tag{7}$$

$x < -l,$

$$u^{(II)} = \sum_{n=1}^{\infty} T_n \cdot \varphi_n(y, a) \cdot e^{-\gamma_n^{(II)} \cdot (x-l)}, \quad x > l, \tag{8}$$

где $\varphi_n(y, a) = \sin\left(\frac{n\pi(y+a)}{2a}\right)$, $\gamma_n^{(I)} = \sqrt{\left(\frac{n\pi}{2a}\right)^2 - k_1^2}$, $\gamma_n^{(II)} = \sqrt{\left(\frac{n\pi}{2a}\right)^2 - k_2^2}$, а коэффициенты отражения R_n и T_n подлежат определению.

Далее, пользуясь техникой, описанной в [3], поле в области 1 запишем в виде суммы:

$$u^{(1)} = \sum_{i=1}^3 u_i^{(1)}, \tag{9}$$

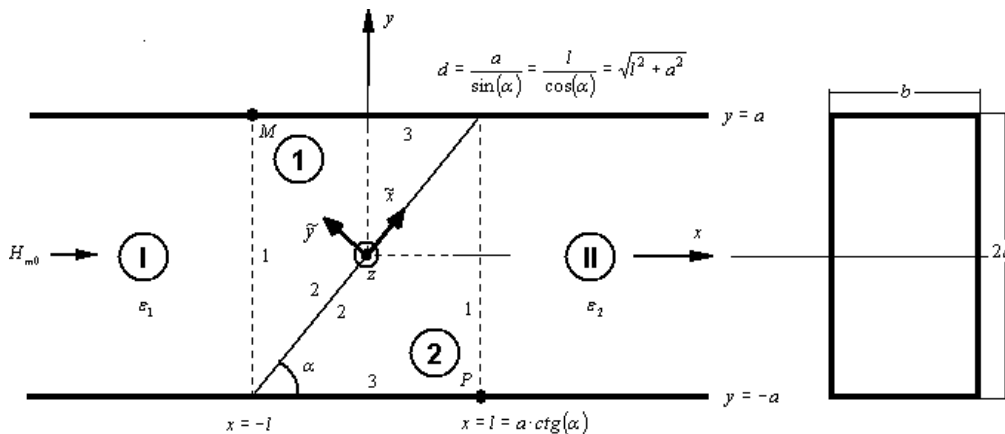


Рисунок 1 – Геометрия задачи

где нумерация функций-слагаемых соответствует нумерации сторон треугольника, а сами они определяются следующими формулами:

$$u_1^{(1)} = \sum_{n=1}^{\infty} {}^{(1)}A_n \cdot \varphi_n(y, a) \cdot e^{-\gamma_n^{(1)} \cdot (x+l)}, \quad (10)$$

$$u_3^{(1)} = \sum_{n=1}^{\infty} {}^{(3)}A_n \cdot \varphi_n(x, l) \cdot e^{\gamma_n^{(1)} \cdot (y-a)}, \quad (11)$$

$$u_2^{(1)} = \sum_{n=1}^{\infty} {}^{(2)}A_n \cdot \varphi_n(\tilde{x}, d) \cdot e^{-\gamma_n^{(1)} \cdot \tilde{y}} + C^{(1)} \cdot \varphi_1(\tilde{x}, d) \cdot e^{\gamma_1^{(1)} \cdot \tilde{y}}, \quad (12)$$

где $\gamma_n^{(1)} = \sqrt{\left(\frac{n\pi}{2l}\right)^2 - k_1^2}$, $\gamma_n^{(1)} = \sqrt{\left(\frac{n\pi}{2d}\right)^2 - k_1^2}$.

Коэффициент $C^{(1)}$ определяется из дополнительного точечного условия:

$$u_2^{(1)}(M) = 0. \quad (13)$$

Аналогично строим представление и для функции $u^{(2)}$:

$$u^{(2)} = \sum_{i=1}^3 u_i^{(2)}, \quad (14)$$

$$u_1^{(2)} = \sum_{n=1}^{\infty} {}^{(1)}B_n \cdot \varphi_n(y) \cdot e^{\gamma_n^{(2)} \cdot (x-l)}, \quad (15)$$

$$u_3^{(2)} = \sum_{n=1}^{\infty} {}^{(3)}B_n \cdot \varphi_n(x, l) \cdot e^{-\gamma_n^{(2)} \cdot (y+a)}, \quad (16)$$

$$u_2^{(2)} = \sum_{n=1}^{\infty} {}^{(2)}B_n \cdot \varphi_n(\tilde{x}, d) \cdot e^{\gamma_n^{(2)} \cdot \tilde{y}} + C^{(2)} \cdot \varphi_1(\tilde{x}, d) \cdot e^{-\gamma_1^{(2)} \cdot \tilde{y}}, \quad (17)$$

где $\gamma_n^{(2)} = \sqrt{\left(\frac{n\pi}{2d}\right)^2 - k_2^2}$, $\gamma_n^{(2)} = \sqrt{\left(\frac{n\pi}{2l}\right)^2 - k_2^2}$, а коэффициент $C^{(2)}$ находится с помощью условия:

$$u_2^{(2)}(M) = 0. \quad (18)$$

АЛГЕБРАИЗАЦИЯ ЗАДАЧИ

Изложим теперь кратко процесс получения системы алгебраических уравнений, которой удовлетворяют неизвестные коэффициенты разложений.

Из граничного условия (13) имеем:

$$C^{(1)} = \sum_{n=1}^{\infty} {}^{(C)}t_n^{(1)} \cdot {}^{(2)}A_n, \quad (19)$$

где ${}^{(C)}t_n^{(1)} = -[\varphi_n(\tilde{x}_M, d)/\varphi_1(\tilde{x}_M, d)] \cdot e^{-[\gamma_1^{(1)} + \gamma_n^{(1)}] \cdot \tilde{y}_M}$.

После подстановки (9)–(12) в граничное условие (4) умножим полученное равенство на $\frac{1}{l} \cdot \varphi_k(x, l)$ и проинтегрируем от $-l$ до l . Мы получим:

$$\sum_{n=1}^{\infty} {}^{(3)}\tilde{a}_{kn}^{(1)} \cdot {}^{(2)}A_n + {}^{(C)}a_k^{(1)} \cdot C^{(1)} - {}^{(3)}A_k = 0, \quad k = \overline{1; \infty}, \quad (20)$$

где ${}^{(3)}\tilde{a}_{kn}^{(1)}$ и ${}^{(C)}a_k^{(1)}$ – известные интегралы.

Последнее уравнение, с учетом (19), перепишем в виде:

$${}^{(3)}A_k = \sum_{n=1}^{\infty} {}^{(3)}\tilde{a}_{kn}^{(1)} \cdot {}^{(2)}A_n, \quad (21)$$

где ${}^{(3)}\tilde{a}_{kn}^{(1)} = {}^{(3)}\tilde{a}_{kn}^{(1)} + {}^{(C)}a_k^{(1)} \cdot {}^{(C)}t_n^{(1)}$.

Воспользуемся далее первым из граничных условий (2). Подставим в него выражения (7) и (9)–(12), а затем спроектируем полученное равенство на ортогональный функциональный базис $\{\varphi_n(y, a)\}$. Мы получим, с учетом (19):

$$R_k = -\delta_{km} + {}^{(1)}A_k + \sum_{n=1}^{\infty} {}^{(2)}\rho_{kn}^{(1)} \cdot {}^{(2)}A_n, \quad (22)$$

где ${}^{(2)}\rho_{kn}^{(1)}$ – некоторые известные выражения.

Из равенства нормальных производных в плоскости $x = -l$, с учетом (19), (21) и (22), следует:

$${}^{(1)}A_k = \delta_{km} + \sum_{n=1}^{\infty} {}^{(1)}\tilde{a}_{kn}^{(1)} \cdot {}^{(2)}A_n, \quad (23)$$

где ${}^{(1)}\tilde{a}_{kn}^{(1)}$ – известно.

Таким образом, все коэффициенты разложений, относящиеся к областям 1 и I, выражаются через неизвестные $\{{}^{(2)}A_n\}$. Аналогично в области 2 и II, все коэффициенты разложений выражаются через $\{{}^{(2)}B_n\}$. Используя теперь граничные условия (7) и производя очевидные исключения неизвестных, мы получим бесконечную систему линейных алгебраических уравнений, содержащих только $\{{}^{(2)}A_n\}$ и $\{{}^{(2)}B_n\}$. Последняя система решается методом усечения.

ЧИСЛЕННЫЕ РЕЗУЛЬТАТЫ

Описанный алгоритм был реализован в виде программы для ЭВМ. Ниже приводятся результаты тестовых расчетов, подтверждающие адекватность и эффективность используемой методики.

В качестве первого теста для анализа корректности алгоритма и скорости сходимости вычислительной процедуры был использован закон сохранения энергии.

Для его выполнения $|A|^2$ в выражении

$$\sum_n \left(\text{Im}(\gamma_n^{(I)}) \cdot |R_n|^2 + \text{Im}(\gamma_n^{(II)}) \cdot |T_n|^2 \right) = \text{Im}(\gamma_n^{(I)}) \cdot |A|^2 \quad (24)$$

должно быть равно 1. На рис. 2 представлены результаты расчета $|A|^2$ по формуле (24) в зависимости от порядка усечения N , которое, для удобства представления данных, выбрано одинаковым для всех рядов. Видно, что значение $|A|^2$ достаточно быстро сходится к его точному значению.

На рис. 3 представлены коэффициенты прохождения, рассчитанные для структуры с параметрами $\epsilon_1 = 1, \alpha = 45^\circ, 2a/\lambda = 0,7$ при изменении относительной диэлектрической проницаемости ϵ_2 в пределах от 1 до 11. Из рисунка следует, что результаты расчета хорошо согласуются с данными работы [12].

Рис. 4 представляет графики зависимости коэффициентов отражения для различных значений частотного параметра от угла наклона границы раздела сред.

Как можно видеть, и в этом случае результаты близки к известным.

Эффективность алгоритма и его достоверность подтверждаются также и для других значений параметров структуры.

ЗАКЛЮЧЕНИЕ

В работе подтверждена возможность расширения области приложений недавно предложенного варианта метода произведения областей на волноводные устройства, содержащие наклонные границы раздела сред. Развиваемый подход основывается на представлении искомого решения уравнения Гельмгольца в выпуклой многоугольной области в виде определенного вида тригонометрических разложений. В качестве тестовой рассмотрена известная H -плоскостная задача дифракции волн на границе раздела двух диэлектриков в прямоугольном волноводе. Ее решение показало, что, хотя изначально метод ведет к довольно громоздкой бесконечной системе алгебраических уравнений, структура этой системы такова, что возможно аналитическое исключение неизвестных коэффициентов разложений,

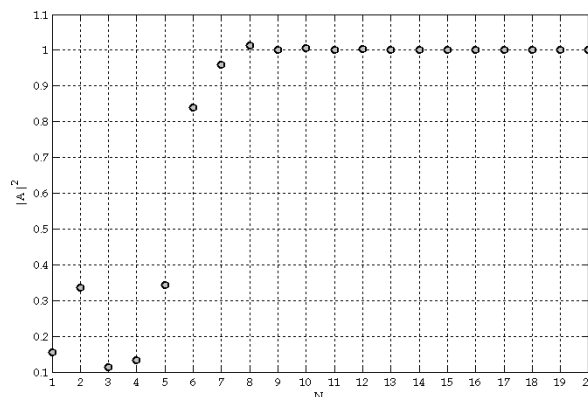


Рисунок 2 – Зависимость $|A|^2$ от порядка усечения N при $\epsilon_1 = 1, \epsilon_2 = 6, 8, 2a/\lambda = 0,7, \alpha = 30^\circ$

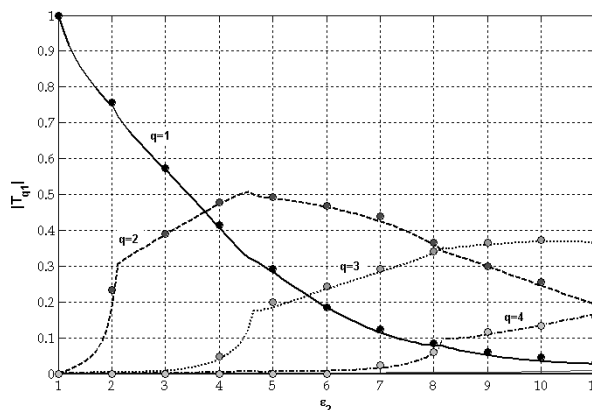


Рисунок 3 – Зависимость коэффициентов прохождения от ϵ_2 при $\epsilon_1 = 1, \alpha = 45^\circ$ и $2a/\lambda = 0,7$ (линии – предложенный метод, кружки – данные из [12])

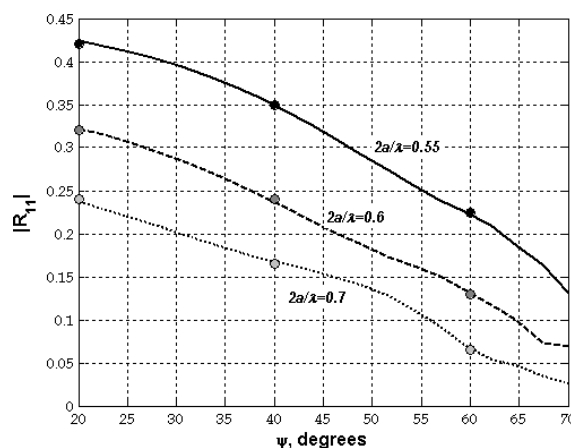


Рисунок 4 – Зависимости коэффициента отражения от значения $2a/\lambda$ и угла наклона границы раздела диэлектрических сред при $\epsilon_1 = 1, \epsilon_2 = 2, \psi = 90^\circ - \alpha$ (линии – предложенный метод, кружки – данные из [12])

относящихся ко всем звеньям граничного контура, кроме самой границы раздела. Последняя система разрешима методом усечения. Численные эксперименты и сравнение с результатами других авторов подтвердили, что построенный таким образом алгоритм обеспечивает эффективное и надежное определение характеристик рассеяния структуры во всем диапазоне значений ее геометрических и электрических параметров, представляющих интерес для приложений.

ПЕРЕЧЕНЬ ССЫЛОК

1. Чумаченко В. П. О расчете H -плоскосных волноводных узлов с многоугольной границей // Радиотехника и электроника. – 1986. – Т. 31, № 12. – С. 2335–2342.
2. Chumachenko V. P. Domain-product technique solution for the problem of electromagnetic scattering from multi-angular composite cylinders // IEEE Transactions on Antennas and Propagation. – 2003. – Vol. 51, No. 10. – P. 2845–2851.
3. Chumachenko V. P. Efficient field representation for polygonal region // Electronics letters. – 2001. – Vol. 37, No. 19. – P. 1164–1165.
4. Chumachenko V. P., Karaçhuha E., Petrusenko I. V. Accurate analysis of waveguide junctions with rectangular coupling cavity // Microwave and Optical Technology Letters. – 2001. – Vol. 31, No. 4. – P. 305–308.
5. Chumachenko V. P., Tarapov S. I., Eker S. Scattering by a lossy dielectric cylinder in a waveguide cross-junction // IEE Proceedings. – Microwaves, Antennas and Propagation. – 2002. – Vol. 149, No. 4. – P. 229–236.
6. Chumachenko V. P. Accurate model of E-plane waveguide junction with loaded rectangular coupling cavity // Microwave and Optical Technology Letters. – 2002. – V. 34, No. 5. – P. 351–354.
7. Chow Y. L., Wu S.-C. A moment method with mixed basis functions for scattering by waveguide junctions // IEEE Transactions on Microwave Theory and Techniques. – 1973. – Vol. MTT-21, No. 5. – P. 333–339.

8. Kashyap S. C. Slant Dielectric Interface Discontinuity in a Waveguide // IEEE Transactions on Microwave Theory and Techniques. – 1975. – Vol. MTT-23, No. 2. – P. 257–260.
9. Каплевич Б. Ю., Силин Н. С. Отражение от диэлектрического клина в прямоугольном волноводе // Изв. вузов. Радиофизика. – 1976. – Т. 19, № 1. – С. 135–140.
10. Zuckerman D. N., Diamant P. Rank reduction of ill-conditioned matrices in waveguide junction problems // IEEE Transactions on Microwave Theory and Techniques. – 1977. – Vol. MTT-25, No.5. – P. 613–619.
11. Ильинский А. С., Воронов А. А. Метод интегральных уравнений в задаче о дифракции волн на наклонной границе раздела двух сред в волноводе // Вычислительные методы и программирование. – М.: Изд-во Моск. ун-та, 1978. – Вып. 28. – С. 177–194.
12. Кириленко А. А., Рудь Л. А. Дифракция волн на наклонной границе раздела диэлектрических сред в прямоугольном волноводе // Радиотехника и электроника. – 1977. – № 10. – С. 2057–2067.

Надійшла 15.01.07

Розглядається відома задача дифракції електромагнітних хвиль на прямолінійній межі між двома діелектричними середовищами. Розв'язок знаходиться за допомогою нещодавно запропонованого варіанта методу добуток областей, який використовує певного виду тригонометричні розвинення для зображення шуканої функції всередині опуклих багатокутників. Шляхом порівняння отриманих результатів з даними інших авторів встановлено можливість застосування та ефективність цього підходу при дослідженні хвильоводних пристроїв, які містять похилу межу розділу середовищ.

The known problem of electromagnetic wave scattering by a rectilinear dielectric interface is considered. The solution is found with the help of a recently proposed version of the domain-product technique, which exploits certain trigonometric expansions to represent the sought-for function inside a convex polygon. The comparison of the data obtained with the results of other authors has shown the validity and efficiency of the technique in studying waveguide devices with a slant interface.

УДК 537.874.6

А. В. Крапивной, В. П. Чумаченко, Я. В. Чумаченко

О РАССЕЙАНИИ ПЛОСКОЙ СКАЛЯРНОЙ ВОЛНЫ ИМПЕДАНСНОЙ СФЕРОЙ

Рассматривается задача рассеяния плоской скалярной волны реактивно нагруженной сферой. Теория основывается на свойстве локальности коротковолнового рассеяния и дает в явном виде значение поля на поверхности цели. Она учитывает кривизну поверхности, а также значение поля на ее затененной части. Многочисленные сравнения с точными решениями для сфер различных размеров с постоянными значениями поверхностного импеданса показали, что метод является более точным, чем приближение физической оптики, и требует приблизительно тех же вычислительных затрат. Приведен при-

мер использования развитого подхода в случае сферы с изменяющейся поверхностной нагрузкой.

ВВЕДЕНИЕ

Задача рассеяния гармонических волн некоторым телом сводится, как известно [1], к граничным задачам для уравнения Гельмгольца. Их точное аналитическое решение существует только для нескольких относительно простых частных случаев, для которых возможно

использование метода разделения переменных [2]. Для более сложных поверхностей разработаны как приближенные аналитические методы, включая такие асимптотические подходы как метод физической оптики и геометрическая теория дифракции, так и численные или численно-аналитические методы, включая методы конечных и граничных элементов, метод сшивания и другие методики. Области приложений всех этих методов ограничены в том или ином аспекте либо частотным диапазоном, либо классом поверхностей, либо возможностью их эффективной численной реализации.

Метод физической оптики (приближение Кирхгофа) [3, 4] применяется обычно в задачах рассеяния коротких волн. Рассеянная волна находится в виде интеграла по элементам поверхности, для каждого из которых предполагается, что его коэффициент отражения совпадает с коэффициентом отражения бесконечной касательной плоскости. На затененной части поверхности рассеивающегося поля принимается равным нулю. Метод имеет ряд недостатков, в частности, точность расчета поля, отраженного в некотором направлении, ухудшается, когда это направление отклоняется от зеркального. Однако этот метод широко используется, так как, исключая матричное решение, является простым и эффективным инструментом при решении задач высокочастотного рассеяния для гладких тел общего вида.

В недавних работах [5, 6] был предложен подход к оцениванию поля, рассеянного выпуклыми двумерными объектами, который, сохраняя простоту метода физической оптики, позволяет заметно улучшить оценки для не зеркальных направлений. Подобно методу физической оптики он дает аналитическое выражение для поля у рассеивающей поверхности, однако при этом учитывает ее кривизну и не предполагает обращение возбуждающего поля в нуль в тени. В работе [7], на примере рассеяния плоской волны мягкой сферой, была показана возможность применения развитой методики к решению трехмерных задач. В настоящей работе метод используется для решения задачи рассеяния плоской скалярной волны на сфере с переменным поверхностным импедансом. По своему физическому содержанию задача является акустической, однако может также рассматриваться как необходимый этап в развитии методов решения задач электромагнитного рассеяния.

ПОСТАНОВКА И РЕШЕНИЕ ЗАДАЧИ

Рассмотрим стационарное волновое поле, которое устанавливается в среде, характеризуемой волновым сопротивлением Z_0 и скоростью распространения волн c , при наличии в ней сферы, показанной на рис. 1. Поле будем характеризовать переменной p и гармонической зависимостью от времени $e^{i\omega t}$, где ω круговая час-

тота колебаний. Если ось z направлена навстречу падающей плоской волне, то возбуждающее поле можно записать в виде

$$p_i = e^{ikz}, \quad k = \frac{\omega}{c} = \frac{2\pi}{\lambda}, \tag{1}$$

где λ – длина волны в среде. Полное поле

$$p = p_i + p_s \tag{2}$$

удовлетворяет уравнению Гельмгольца вне сферы и граничному условию

$$\frac{\partial p}{\partial n} - ik\chi p = 0 \tag{3}$$

на ее поверхности. Здесь

$$\chi = \frac{Z_0}{Z}, \tag{4}$$

где Z – поверхностный импеданс. Направление нормали \vec{n} указано на рис. 1. Отметим также, что значение $\chi = 0$ отвечает случаю жесткой сферы.

Рассеянное поле p_s на бесконечности удовлетворяет условию излучения. Вне сферы оно может быть выражено через значения p и $\partial p/\partial n$ на ее поверхности S с помощью интеграла Кирхгофа

$$p_s(M) = \frac{1}{4\pi} \iint_S \left[p \frac{\partial}{\partial n} \frac{e^{-ikr_{PM}}}{r_{PM}} - \frac{\partial p}{\partial n} \frac{e^{-ikr_{PM}}}{r_{PM}} \right] dS_p. \tag{5}$$

С учетом соотношения (3) ясно, что p_s полностью определяется значениями полного поля $p(P)$ на поверхности рассеивателя.

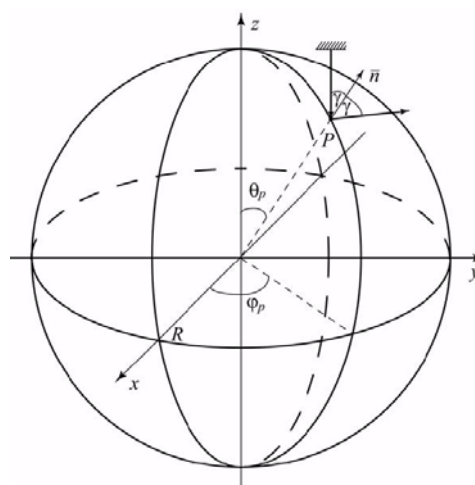


Рисунок 1 – Геометрия задачи

С помощью метода физической оптики значение поля на поверхности легко может быть оценено как

$$p(P) = \begin{cases} e^{ikR\cos\gamma} \frac{2\cos\gamma}{\cos\gamma + \chi} & \gamma \leq \frac{\pi}{2} \\ 0 & \gamma > \frac{\pi}{2} \end{cases}, \quad (6)$$

где γ – угол падения в точке P . Для того, чтобы уточнить это приближение, мы будем рассматривать окрестность точки P не как часть тангенциальной плоскости, а как часть кругового цилиндра, который перпендикулярен плоскости падения, касается линии пересечения плоскости падения со сферой и имеет радиус R , совпадающий с радиусом кривизны этой линии.

Пусть (r, θ, φ) – сферическая система координат с началом в центре шара и полярной осью вдоль Oz , а (ρ, ϕ, ζ) – цилиндрическая система координат такая, что ζ направлено вдоль оси цилиндра и $\phi = \theta - \theta_P$ вблизи точки P . В малой окрестности точки P , воспользовавшись разложением $\cos\theta$ по степеням ϕ , падающую волну можно приближенно описать следующим выражением

$$p_i = e^{ikR\cos\theta} = e^{ikR\cos(\phi + \theta_P)} = e^{ikR(\cos\theta_P - \phi\sin\theta_P + \dots)} \approx e^{ikR(\cos\theta_P - \phi\sin\theta_P)}. \quad (7)$$

Далее мы предположим, что точно также как и в случае отражающей плоскости, функциональные зависимости p_s и p_i в пределах названной окрестности на поверхности совпадают с точностью до постоянного множителя. После этого, разделяя переменные в уравнении Гельмгольца в цилиндрических координатах и сравнивая полученное выражение с формулой (7), мы получим, что вблизи P

$$p_s = C \frac{H_{\mu}^{(2)}(k\rho)}{H_{\mu}^{(2)}(kR)} e^{-i\mu\phi}, \quad \mu = kR\sin\gamma. \quad (8)$$

Подставив (8) в граничное условие (3), мы найдем константу C и, далее, следующее приближение для поля на поверхности

$$p(P) = e^{ikR\cos\gamma} \left[1 + i \frac{\chi - \cos\gamma}{\sin\gamma - H_{\mu+1}^{(2)}(kR)/H_{\mu}^{(2)}(kR) - i\chi} \right]. \quad (9)$$

Знание поля на поверхности сферы позволяет с помощью формулы (5) определить полностью все характеристики рассеянного поля.

ЧИСЛЕННЫЕ РЕЗУЛЬТАТЫ

При $M \rightarrow \infty$ формула (5) преобразуется к виду

$$p_s(M) \approx f(\varphi_M, \theta_M) \frac{e^{-ikr}}{r}. \quad (10)$$

Здесь $\bar{r} = \overline{OM}$ и функция рассеяния $f(\varphi_M, \theta_M)$ определяется интегралом

$$f(\varphi_M, \theta_M) = -\frac{1}{4\pi} \iint_S e^{ik|\overline{OP}| \cos\psi} \left[\frac{\partial p}{\partial n} - ikp \cos\alpha \right] dS_p, \quad (11)$$

где ψ – угол между \bar{r} и \overline{OP} , а α – угол между \bar{r} и \bar{n} . Учитывая граничные условия (3) и то, что для сферы $\psi = \alpha$ и $|\overline{OP}| = R$, формулу (11) можно переписать в виде

$$f(\varphi_M, \theta_M) = -\frac{ik}{4\pi} \iint_S (\chi - \cos\alpha) p e^{ikR\cos\alpha} dS_p. \quad (12)$$

Величина

$$\sigma = 4\pi |f|^2 \quad (13)$$

представляет поперечник рассеяния. Если $\chi(P)$ не зависит от φ , то интегрируя в (12) по φ и учитывая известные [8] табличные интегралы, можно получить

$$\sigma(\theta_M) = \pi R^4 k^2 \left[\int_0^\pi p [(\chi - \cos\theta_M \cos\theta) J_0(kR \sin\theta_M \sin\theta) - i \sin\theta_M \sin\theta J_1(kR \sin\theta_M \sin\theta)] e^{ikR\cos\theta_M \cos\theta} \sin\theta d\theta \right]^2. \quad (14)$$

Ниже для представления данных применяются нормированный поперечник рассеяния $\sigma_s = \sigma/(\pi R^2)$ и нормированная амплитуда рассеянного поля $\sqrt{\sigma_s} = 2|f|/R$. В случае постоянного поверхностного импеданса для сравнения используется точное решение, под которым мы подразумеваем решение в виде ряда, полученное методом разделения переменных.

На рис. 2 показаны рассчитанные распределения полного поля на поверхности сферы радиусом 2λ при различных значениях поверхностного импеданса. Во всех случаях наше приближение лучше, чем приближение физической оптики. Точность аппроксимации повышается при изменении типа поверхности от жесткой к мягкой.

Рис. 3 представляет угловые зависимости рассеянного поля в дальней зоне. Для всех значений R и χ при использовании нашего приближения точность оценок для не зеркальных направлений повышается. Особенно это заметно для относительно небольших R .

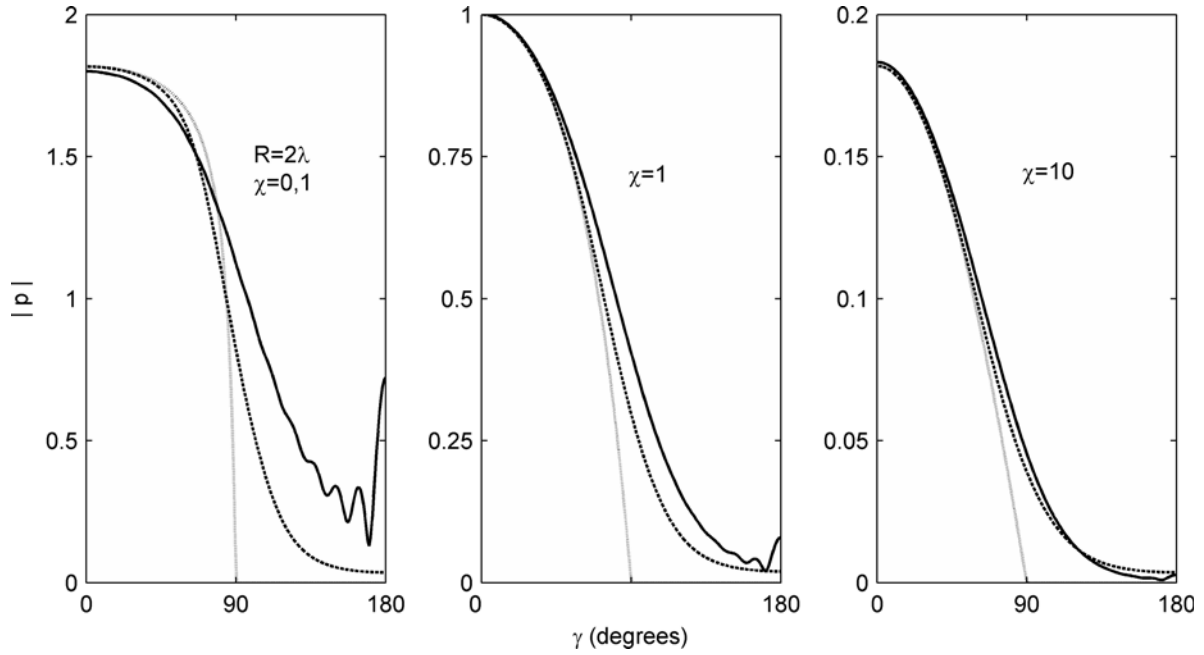


Рисунок 2 – Полное поле на поверхности сферы радиусом 2λ при $\chi = 0, 1$, $\chi = 1$ и $\chi = 10$
 (сплошная линия – точно, пунктир – метод физической оптики,
 штриховая линия – предложенный метод)

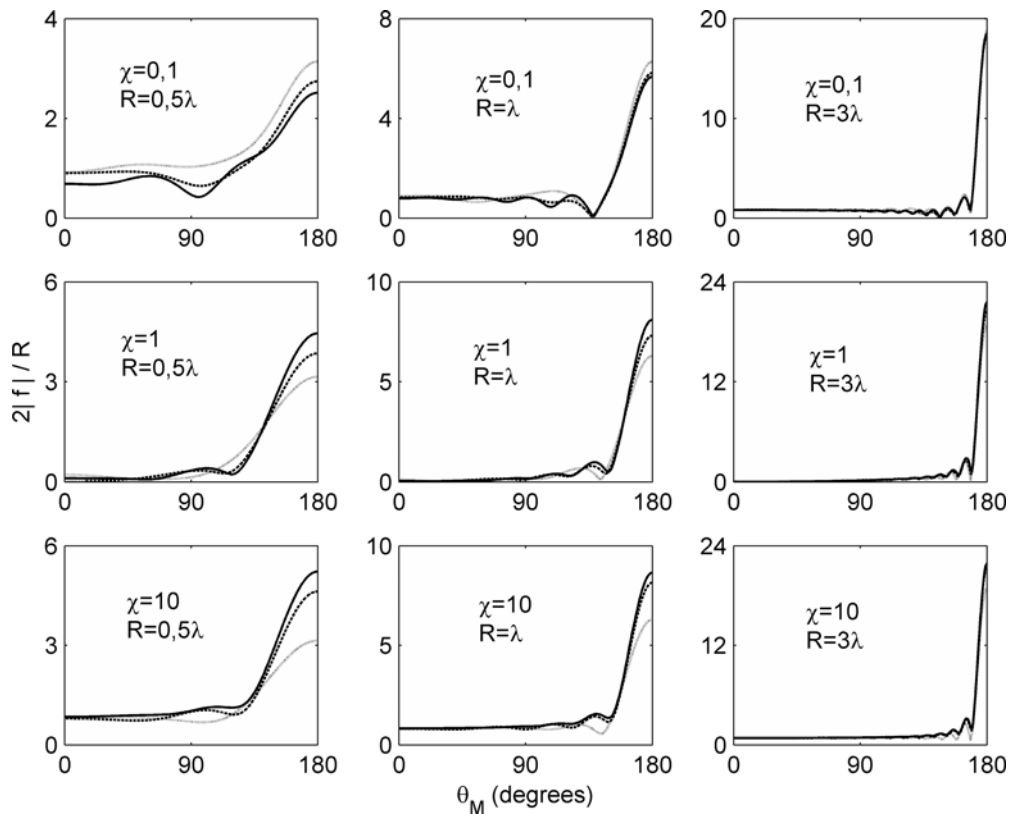


Рисунок 3 – Угловая зависимость рассеянного поля в дальней зоне
 (сплошная линия – точно, пунктир – метод физической оптики,
 штриховая линия – предложенный метод)

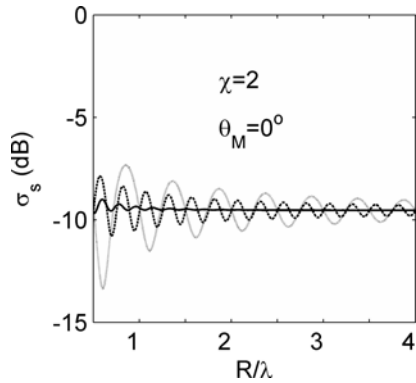


Рисунок 4 – Зависимость поперечника обратного рассеяния от радиуса сферы при $\chi = 2$ (сплошная линия – точно, пунктир – метод физической оптики, штриховая линия – предложенный метод)

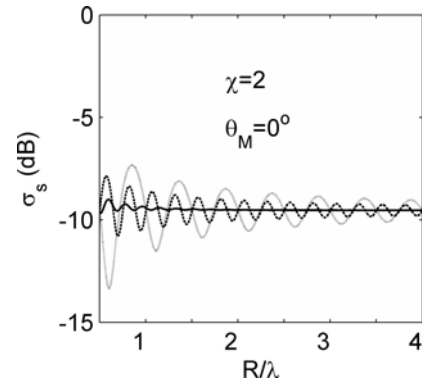


Рисунок 5 – Зависимость поперечника рассеяния в прямом направлении от радиуса сферы при $\chi = 2$ (сплошная линия – точно, пунктир – метод физической оптики, штриховая линия – предложенный метод)

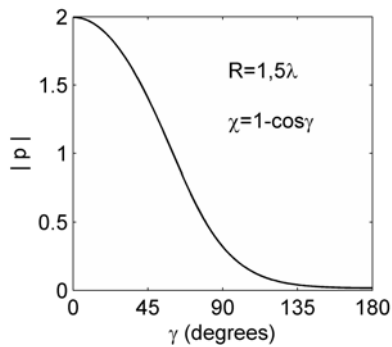


Рисунок 6 – Полное поле на поверхности при $R = 1,5\lambda$ и $\chi = 1 - \cos \gamma$

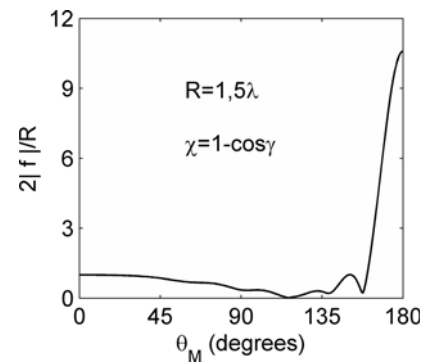


Рисунок 7 – Угловая зависимость рассеянного поля в дальней зоне при $R = 1,5\lambda$ и $\chi = 1 - \cos \gamma$

На рис. 4 и 5 показаны полученные зависимости от R поперечников обратного и прямого рассеяния. Видно, что использование формулы (9) вместо (6) приводит к некоторому повышению точности и для обратно отраженного поля. Для прямого же направления улучшение точности весьма существенно и эта точность быстро увеличивается с ростом R .

На рис. 6 и 7 приведен пример эксплуатации алгоритма для случая импеданса, обладающего угловой зависимостью, которая реализует плавный переход от жесткой поверхности на освещенной части к $\chi = 2$ в тени.

ЗАКЛЮЧЕНИЕ

Рассмотрена задача высокочастотного рассеяния плоской скалярной волны на сфере с переменным поверхностным импедансом. Для ее решения применен предложенный ранее приближенный метод, использующий свойство локальности коротковолнового рассеяния и основанную на этом возможность аналитичес-

кой оценки поля на рассеивающей поверхности. Путем сравнения получаемых приближенных решений с точными решениями для ряда объектов с постоянными значениями поверхностного импеданса установлено, что предложенный алгоритм является более точным, чем метод физической оптики, и обладает приблизительно той же эффективностью. Приведен пример расчета характеристик рассеяния для сферы с импедансом, зависящим от угла падения возбуждающей волны.

ПЕРЕЧЕНЬ ССЫЛОК

1. Виноградова М. Б., Руденко О. В., Сухоруков А. П. Теория волн. – М.: Наука, 1979. – 384 с.
2. Морс Ф. М. и Фешбах Г. Методы теоретической физики. Том 1. – М.: Изд-во иностр. лит., 1958. – 930 с.
3. Knott E. F., Shaeffner J. F., Tuley M. T. Radar cross section. – Norwood, MA: Artech House, 1993. – 611 p.
4. Gaunard G. C. Sonar cross sections of bodies partially insensitized by finite sound beams // IEEE J. Ocean Eng. – Vol. 10. – 1985. – P. 213–230.
5. Chumachenko V. P. On the estimation of scattering from convex conducting cylinders // Microwave and Optical Technology Letters. – Vol. 45. – May 2005. – P. 191–194.

6. Крапувну А. В., Чумаченко В. Р. On the estimation of scattering from convex impedance cylinders // Proceedings of the 11th Int. Conf. on Mathematical Methods in Electromagnetic Theory, Kharkiv, Ukraine. – June 26–29, 2006. – P. 248–250.
7. Крапивной А. В., Чумаченко В. П. Об оценивании волнового поля рассеянного мягкой сферой // Радиоэлектроника. Информатика. Управління. – № 2, 2006. – С. 12–15.
8. Градштейн И. С., Рыжик И. М. Таблицы интегралов, сумм, рядов и произведений. – М.: Наука, 1971. – 1108 с.

Надійшла 31.01.07

Розглядається задача розсіювання плоскої скалярної хвилі реактивно навантаженою сферою. Теорія ґрунтується на властивості локальності короткохвильового розсіювання і дає значення поля на поверхні в явному вигляді. Вона враховує кривизну поверхні, а також значення поля на її затіненій частині. Багаторазові порівняння з точними розв'язками для сфер різних розмірів зі сталими значеннями поверхневого імпедансу показали, що метод є більш точним, ніж наближення фізичної оптики,

і потребує приблизно тих самих обчислювальних затрат. Наведено приклад застосування розвинутого підходу в випадку сфери з поверхневим імпедансом, що змінюється.

The problem of scalar plane-wave scattering off a reactively loaded sphere is considered. The theory is based on the locality property of short-wave scattering and yields an explicit formula for the surface field. It takes into account the curvature of the surface and supposes that the incident field is nonzero in both the lit and shadow regions. Extensive comparisons with exact solutions for spheres of different sizes with constant surface impedances have shown that the approach is more accurate than the physical optics approximation and requires about the same computational costs. Although the method is strictly valid for high frequencies, it gives good qualitative results down to resonant frequency range. An example of application of the technique to the sphere with a varying reactive load is presented.

РАДІОЕЛЕКТРОНІКА ТА ТЕЛЕКОМУНІКАЦІЇ

РАДИОЕЛЕКТРОНИКА И ТЕЛЕКОММУНИКАЦИИ

RADIO ELECTRONICS AND TELECOMMUNICATIONS

УДК 621.396.6

Б. А. Артюшенко

ПАРАМЕТРИЧНИЙ СИНТЕЗ ЛІНІЙНОЇ АНТЕННОЇ РЕШІТКИ ПРИ ЗАДАНИХ ДОПУСКАХ ЗА ДОПОМОГОЮ ГЕНЕТИЧНОГО АЛГОРИТМУ

Запропоновано метод, що базується на використанні генетичного алгоритму синтезу антенних решіток при заданих допусках на параметри. Розглянуто можливість використання інтервалів для представлення допускової області. Показано, що природне інтервальне розширення може бути застосовано для оцінки граничних лінійних антенних решіток.

ВСТУП

При виробництві та експлуатації антен виникають каналні, фазові та інші помилки і відхилення. Ці зміни параметрів чинять суттєвий вплив на характеристики антени. Існуючі методи проектування антен не завжди дозволяють враховувати допуски на параметри. Актуальною задачею є вибір номінальних параметрів антенної решітки за заданими діапазонами відхилень. Для розв'язання цієї задачі може бути застосовано генетичний алгоритм (ГА), який вже використовується при синтезі номінальних параметрів антен [1–3] та допусковому синтезі [4].

Метою статті є розробка генетичного алгоритму параметричного синтезу лінійної антенної решітки з заданими допусками на параметри.

Для вирішення поставленої задачі необхідно:

- розробити засоби представлення допускових значень для задачі розробки антен;
- розробити засоби оцінки характеристик антенної решітки при неточних значеннях її параметрів;
- адаптувати ГА для задачі параметричного синтезу;
- провести аналіз впливу параметрів ГА на ефективність його застосування.

ЗАДАЧА ПАРАМЕТРИЧНОГО СИНТЕЗУ ЛІНІЙНОЇ АНТЕННОЇ РЕШІТКИ

Діаграма напрямленості (ДН) представленої на рис. 1 лінійної антенної решітки (ЛАР):

$$E_{\varphi} = \tilde{f}_{\varphi} \sum_{n=-M}^M I_n \exp(j(k\Delta r_n + \phi_n)), \quad (1)$$

де \tilde{f}_{φ} – діаграма направленості елементу ЛАР; $k = 2\pi/\lambda$; λ – довжина хвилі; $j = \sqrt{-1}$, I_n – величина амплітуди струму випромінювача n ; Δr_n – відстань між елементом n та центром ЛАР; ϕ_n – фазова відстань; $2M + 1$ – кількість елементів.

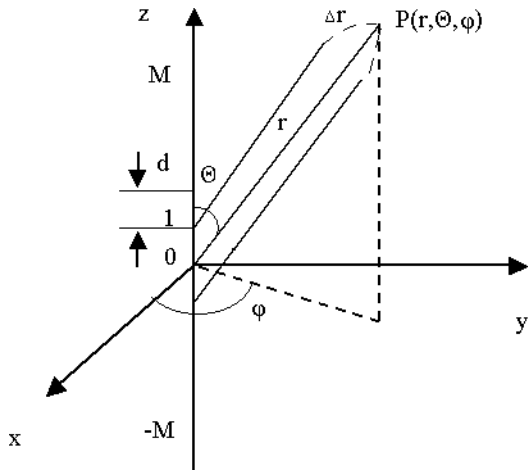


Рисунок 1 – Лінійна антенна решітка

Найчастіше як показник якості ЛАР використовують інтегральну характеристику амплітудної діаграми напрямленості:

$$\int_{\varphi=0}^{2\pi} \left| |E_{\varphi}| - |E_{\varphi}^{\text{target}}| \right|, \quad (2)$$

де $E_{\varphi}^{\text{target}}$ – оптимальна діаграма напрямленості.

Через виникаючі відхилення ДН набуває вигляд:

$$E_{\varphi} = \tilde{f}_{\varphi} \sum_{n=-M}^M (I_n \pm \Delta I_n) e^{j(kn(d \pm \Delta d_n) \cos(\varphi) + \phi_n \pm \Delta \phi_n)}, \quad (3)$$

де ΔI_n , $\Delta \phi_n$, Δd_n – величини відхилень; d – відстань між сусідніми елементами.

Для оцінки максимального і мінімального можливих значень діаграми напрямленості може бути використа-

но інтервальне розширення. Вираз (3) є раціональним виразом відносно параметрів ΔI_n . Параметри $\Delta \phi_n$, Δd_n є аргументами функцій \cos , \sin , отже природне інтервальне розширення є досить вузьким зовнішнім розширенням та може бути застосовано для аналізу.

Припустимо, що ΔI_n , $\Delta \phi_n$, Δd_n відомі заздалегідь. Отже задача полягає у підборі номінальних параметрів (I_n, ϕ_n, d_n) при заданих $(I_n = \pm \Delta I_n, \phi_n = \pm \Delta \phi_n, d_n = \pm \Delta d_n)$ з тим, щоб мінімізувати

$$\int_{\varphi=0}^{2\pi} \text{dist} \left(\frac{|E_{\varphi}^{\text{target}}|}{\max(|E_{\varphi}^{\text{target}}|)}, \left. \tilde{f}_{\varphi} \sum_{n=-M}^M (I_n + I_n) \exp(j(k \cos(\varphi)(nd + \mathbf{d}) + n\phi + \mathbf{d}\phi)) \right) \right), \quad (4)$$

де dist – відстань Хаусдорфа.

ГЕНЕТИЧНИЙ АЛГОРИТМ

Особливістю ГА є те, що він проводить пошук від багатьох точок і використовує випадковий перехід з одних рішень до наступних замість детермінованого переходу. ГА виконує пошук, проходячи задану кількість кроків. Протягом кроку ГА (рис. 2) визначається група (популяція) різних рішень (геномів). Геноми представлено бітовими рядками. До схрещення допускаються попарно геноми (батьки), що обираються процедурою селекції відповідно до їх пристосованості. Для задачі синтезу ЛАР пристосованість визначається за виразом (4). Процес схрещення полягає в обміні ге-

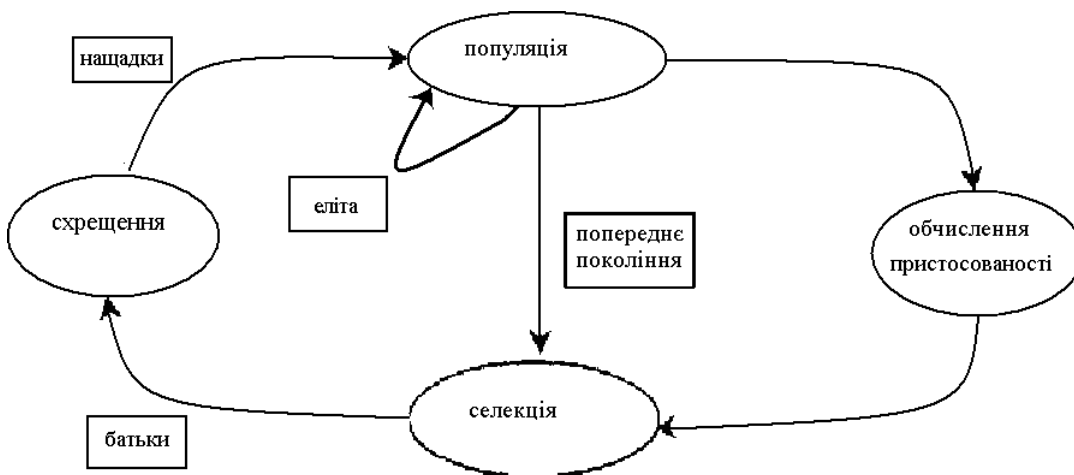


Рисунок 2 – Крок генетичного алгоритму

нетичного матеріалу між батьками (кросовер) та внесення випадкових змін (мутації) з утворенням генонащадка. Найкращі батьки (еліта) та нащадки утворюють нове покоління і переходять до наступного кроку.

Використання ГА для задачі параметричного синтезу ЛАР має певні особливості:

- для антенних решіток бінарне кодування [1, 3], рівномірний кросовер та мультимутація [2] дають кращі результати. На відміну від класичного кросоверу з однією точкою при рівномірному кросовері всі гени нащадка формуються випадковим чином незалежно один від одного. Якщо $u = \{u_1, u_2, \dots, u_N\}$, $v = \{v_1, v_2, \dots, v_N\}$ - батьки, то їх нащадком за рівномірного схрещення може бути $z = \{v_1, u_2, u_3, v_4 \dots u_N\}$. Рівномірне схрещення дозволяє отримати покоління, що буде в більшій ступені відтворювати попереднє і буде менш детермінованим;

- на параметри накладено виключно обмеження зверху та знизу, отже має бути застосовано спеціальне відображення [5, 6];

- через обчислювальну складність визначення ДН має бути застосовано елітизм [6]. Під елітизмом розуміється збереження кращої частини попереднього покоління;

- початкове покоління формується із рішення отриманого аналітичними методами з внесенням 50 % мутації в геноми (без внесення мутації процес пошуку не зможе початись). Якщо допускається вибір кількості елементів, то кожному значенню M відводиться рівна частина і один геном генерується без внесення мутації;

- при синтезі обирається кількість елементів ЛАР, отже слід доробити процедуру кросоверу. Приклад адаптованого кросоверу наведено в таблиці 1.

Для допускової задачі з 8–20 параметрами має бути застосовано [4, 5] популяцію розміром в 50–100 та стандартні величини щільності мутації ($1/M$) та кросоверу (0,5–0,6). Через малу популяцію імовірність мутації має бути високою.

Таблиця 1 – Приклад кросоверу

Хромосома	Ген 1 (M)	Ген 2	Ген 3	Ген 4
Батько 1	000010 ($M = 2$)	010110	010101	*****
Батько 2	000011 ($M = 3$)	010100	101100	111011
Нащадок 1	000011 ($M = 3$)	0101*0	***10*	111011
Нащадок 2	000010 ($M = 3$)	0101*0	***10*	*****

Примітка – У таблиці 1 * приймає випадково значення 1 або 0

Таблиця 2 – Розподіл I_n між елементами

n	1, 21	2, 20	3, 19	4, 18	5, 17	6, 16	7, 15	8, 14	9, 13	10, 12	11
I_n	0,293	0,361	0,606	0,708	1,066	1,088	1,267	1,367	1,480	1,502	1,524

ПРИКЛАДИ

Розглянемо задачу синтезу номінальних значень параметрів M, I_n, ϕ, d ЛАР з ненапрямлених елементів при заданій діаграмі напрямленості (рис. 3, а). Область пошуку параметрів: $6 < M < 10, 0,1 < I_n < 2,$

$\phi < 1, 0,02 < d < 0,5, \sum_{n=-M}^M I_n = 21$. Допустимі відхилення параметрів: $\Delta I = 0,02, \Delta \phi = 0,01, \Delta d = 0,01$.

Було встановлено наступні значення параметрів генетичного алгоритму: величина популяції – 60, щільність мутації – 0,01, вірогідність кросоверу – 1, вірогідність мутації – 1, еліта – 0,3, кількість поколінь (кроків) – 100. Хід роботи ГА наведено на рис. 4.

Результат роботи ГА на 30 кроці показано на рис. 3, б. Було отримано ЛАР з 10 елементів, номінальні значення параметрів: $\phi = 0,0039, d = 0,4098$, розподіл I_n наведено в таблиці 2

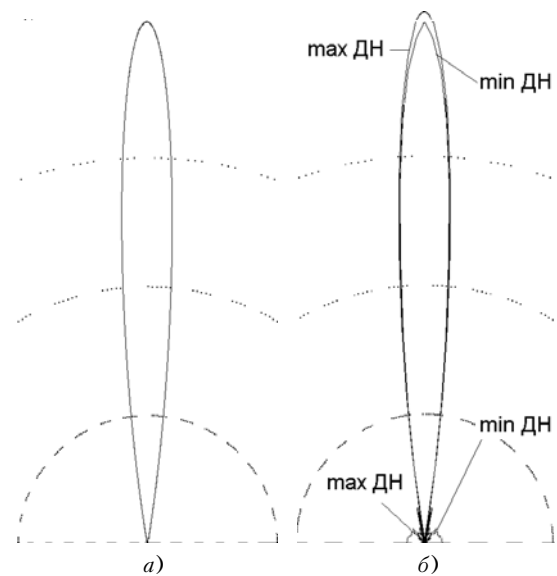


Рисунок 3 – Діаграма напрямленості ЛАР: а – оптимальна; б – отримана з використанням ГА

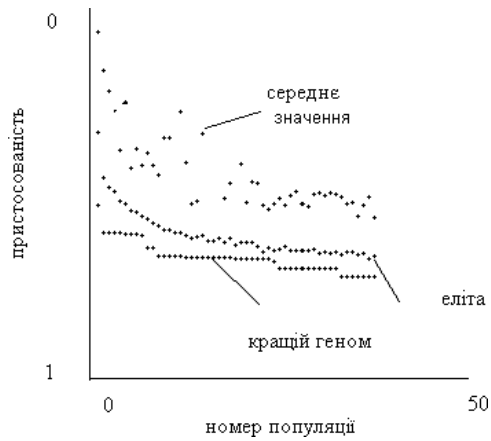


Рисунок 4 – Хід роботи генетичного алгоритму

Як і слід було очікувати, оптимальним виявився симетричний розподіл амплітуди між елементами з максимумом посередині ($n = 11$). Як видно з рис. 3, б, відхилення параметрів найбільш суттєво впливають на бокові пелюстки. Проведено розрахунок номінальних значень параметрів в двох випадках з урахуванням впливу заданих відхилень параметрів та без урахування відхилень параметрів. Урахування відхилень призвело до погіршення ДН на 1–1,5 % для номінальних значень параметрів, але покращило ДН на 2–3 % для максимальних значень відхилень.

ВИСНОВКИ

В роботі запропоновано генетичний алгоритм параметричного синтезу лінійної антенної решітки. Для отримання граничних діаграм напрямленості використовувалось природне інтервальне розширення. Проведені чисельні дослідження показали, що застосування

генетичного алгоритму для задачі синтезу антенної решітки при заданих відхиленнях параметрів дозволяє отримати стійкі рішення.

ПЕРЕЛІК ПОСИЛАНЬ

1. Lee Y., Marvin A. and Porter S. Genetic algorithm using real parameters for array antenna design optimization // Proceedings of the High Frequency Postgraduate Student Colloquium. – Leeds, 1999. – P. 8–13.
2. Elkamchouchi H. M., Wagih M. M. Genetic algorithm operation effect in optimizing the antenna array pattern synthesis // Proceedings of the Twentieth National Radio Science Conference. – Cairo (Egypt), 2003. – P. 1–7.
3. Haupt R. Thinned arrays using genetic algorithms // IEEE Trans. on Antennas Propag. – 1994. – Vol. 42, No. 7. – Pp. 993–999.
4. Artyushenko B., Krischuk V., Shilo G. Tolerable area creation with genetic algorithm // Proceedings of the International Conference TSCET'2006 «Modern problems of radio engineering, telecommunications and computer science». – Lviv-Slavsko (Ukraine), 2006. – Pp. 121–124.
5. Кришук В. М., Шило Г. М., Артюшенко Б. А. Генетичний алгоритм призначення допусків на радіоелементи з інтервальним оцінюванням // Радіоелектроніка. Інформатика. Управління. – 2006. – № 2. – С. 29–32.
6. Артюшенко Б. А., Шило Г. М., Кришук В. М. Генетический алгоритм назначения интервальных допусков // Доклады Всероссийского совещания (с международным участием) по интервальному анализу и его приложениям. – Петергоф (Россия), 2006. – С. 5–8.

Надійшла 13.11.06
Після доробки 31.01.07

Предложен метод, базирующийся на генетическом алгоритме синтеза антенных решеток при заданных допусках на параметры. Рассмотрена возможность применения интервалов при представлении допусковой области. Показано, что естественное интервальное расширение может быть использовано для оценивания предельных значений диаграммы направленности линейных антенных решеток.

A new genetic algorithm based method for antenna array, with given parameter tolerances, design is developed. Tolerance domain representation with intervals is considered. It is shown that natural interval extension is usable for maximal and minimal directorial diagram evaluation.

УДК 621.396.6.004 : 004.942

А. Ю. Воропай, В. Н. Кришук, Л. М. Карпуков, А. Ю. Фарафонов

ДОПУСКОВЫЙ АНАЛИЗ КОНСТРУКЦИЙ ФИЛЬТРОВ С НЕМОНОТОННОСТЬЮ АЧХ

Предложено процедуры назначения допусков параметров фильтров с использованием методов математического программирования для определения глобального экстремума АЧХ. Проведено исследование влияния допусков на параметры пассивного микрополоскового полосового фильтра на связанных линиях с боковой связью и активного полосового RC-фильтра с многопетлевой обратной

связью. Показано преимущества предложенного метода в сравнении с данными внутренней интерполяции для интервальных моделей.

Развитие радиоаппаратуры базируется на принципах разработки и освоения новой элементной базы, фи-

© Воропай А. Ю., Кришук В. Н., Карпуков Л. М., Фарафонов А. Ю., 2007

зико-химических, схемотехнических и конструктивных решений. При этом цель развития лежит в оптимизации совокупности показателей функциональных возможностей, сроков разработки и цены. Освоение элементной базы, сроки разработки и цена в огромной степени зависят от технологии производства. При этом допуски на параметры устройства часто определяют технологию его изготовления. Так, отклонения на элементы конструкции СВЧ устройств в микрополосковом исполнении будут влиять на техпроцесс изготовления микрополосковой топологии, а допуски на параметры элементов активных фильтров требуют подбора параметров элементной базы и могут стать причиной схемотехнических изменений или проведения операций регулировки и настройки.

Быстроменяющиеся гибкие условия рынка часто приводят к переориентации производств, что обуславливает мелкосерийный или серийный тип производства радиоаппаратуры для отечественных предприятий. В этом случае характерно отсутствие статистических данных в технологическом цикле. Поэтому выпуск продукции возможен при 100 % контроле параметров отдельных элементарных устройств или с использованием принципа максимума-минимума при их изготовлении, а именно применении гарантированных допусков.

Разработке методов анализа и синтеза допусков в радиоэлектронике посвящено значительное количество работ [1, 2]. При расчете гарантированных допусков обычно используются теория чувствительности, метод передаточных коэффициентов и интервальный метод [3]. Отличительной особенностью интервального метода является получение точных значений гарантированных допусков на границах допусковой области. Однако точные значения (внешняя интерполяция [3]) не могут быть получены в случае, если в границах допусковой области выходная функция немонотонна, что сводит данный метод к методу передаточных коэффициентов (внутренняя интерполяция [3]). Для решения задач анализа и синтеза допусков в условиях немонотонности выходной функции могут быть использованы методы детерминированного и случайного поиска. При этом при немонотонности выходной функции по небольшому количеству параметров и наличии 1–2 экстремумов применение детерминированного поиска дает огромный выигрыш в скорости. Это позволяет использовать такие методы не только в задачах анализа допусков, но и в задачах их синтеза и, в перспективе, в задачах оптимизации конструкции.

Целью данной статьи является разработка универсального метода допускового анализа и синтеза устройств фильтрации в условиях немонотонности АЧХ. В качестве объекта исследований были выбраны микрополосковый ППФ на связанных линиях с немоно-

тонным характером АЧХ в области граничных частот полосы пропускания при наличии отклонений параметров конструкции и активный полосовой RC-фильтр.

1 ПРОЕКТИРОВАНИЕ ФИЛЬТРОВ НА СВЯЗАННЫХ ЛИНИЯХ

Для микрополосковых ППФ классической является топология фильтров на одинаковых связанных линиях в виде полуволновых резонаторов с боковой связью (рис. 1).

При проектировании ППФ необходимо определить:

- на этапе синтеза, вид и количество звеньев фильтра n , параметры элементов схемы замещения фильтра g_i и волновые сопротивления четного Z_{oe} и нечетного Z_{oo} типа волн для каждого звена фильтра;
- на этапе расчета размеров топологии фильтра, ширины связанных линий фильтра W_i , зазоры между ними S_i и длины областей связи фильтра l_i (рис. 1).

Подробно методика синтеза данного типа фильтра представлена в [4].

Исследуемой характеристикой фильтра является АЧХ, которая, в случае ППФ на связанных линиях с разомкнутыми на концах резонаторами, может быть описана классическими матрицами передачи. При этом функция вносимого затухания имеет вид [1]:

$$a(f) = 20 \cdot \log \frac{1}{2} \cdot \left(A_{11} + \frac{A_{12}}{Z} + A_{21} \cdot Z_0 + A_{22} \right), \quad (1)$$

где A_{11} , A_{12} , A_{21} , A_{22} – коэффициенты результирующей матрицы передачи фильтра; Z_0 – волновое сопротивление подводящих линий передачи.

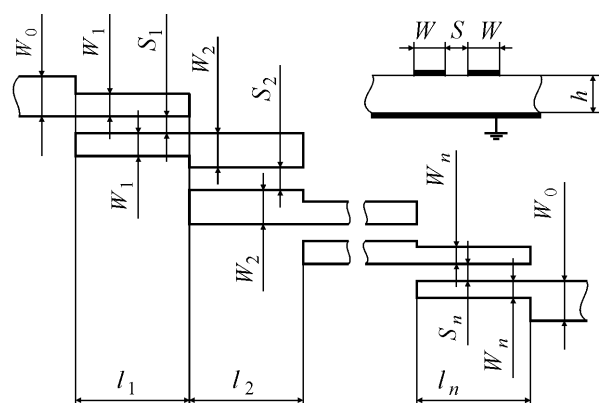


Рисунок 1 – Топология фильтра на связанных линиях:

W_i – ширины связанных линий фильтра; S_i – зазоры между связанными линиями; l_i – длины звеньев фильтра; h – толщина диэлектрической подложки

2 РАСЧЕТ АКТИВНОГО ПОЛОСОВОГО RC-ФИЛЬТРА С МНОГОПЕТЛЕВОЙ ОБРАТНОЙ СВЯЗЬЮ

Классическая схема активного полосового RC-фильтра на операционном усилителе с многопетлевой обратной связью приведена на рис. 2 (обозначения соответствуют авторским) [5].

Передаточная функция по напряжению имеет вид:

$$H = \left| h_0 \frac{j \cdot \omega \cdot b_1}{1 + j \cdot \omega \cdot b_1 + (j \cdot \omega)^2 \cdot b_2} \right|;$$

$$b_1 = (C_1 + C_2) \frac{R_0 \cdot R_2}{R_0 + R_2}; \quad b_2 = R_0 \cdot R_1 \cdot C_1 \cdot C_2;$$

$$h_0 = \left(\frac{R_1 \cdot C_1}{R_2 \cdot C_2} \right) / \left(1 + \frac{C_1}{C_2} \right). \quad (2)$$

3 УЧЕТ НЕМОНОТНОСТИ АЧХ В ПЕРЕХОДНОЙ ОБЛАСТИ

Задача анализа допусков формулируется как задача нахождения максимального и минимального значения в области, ограниченной интервалом допуска входных параметров:

$$y_{\min} = f(X_{\min}), \quad y_{\max} = f(X_{\max}),$$

при ограничениях $X_{\min i} \in [\underline{X}, \bar{X}]$, $X_{\max i} \in [\underline{X}, \bar{X}]$, где X_{\min} и X_{\max} – комбинация параметров x_i , которые обеспечивают соответствующие значения выходного параметра y .

Для немонотонной выходной функции задача требует поиска экстремумов выходной функции на интервале допуска входных параметров.

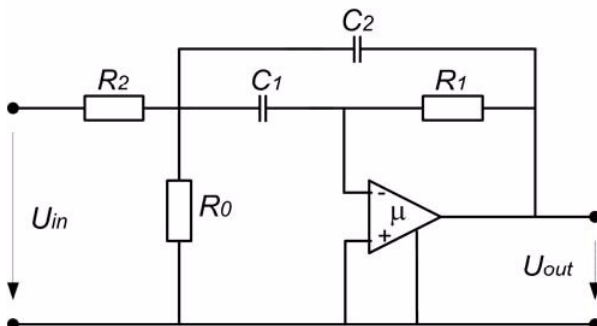


Рисунок 2 – Схема RC-фильтра

В теории математического программирования эта задача получила название оптимизации с ограничениями в виде неравенств [6]. В работе [7] приведен метод анализа допусков для немонотонных выходных функций. В его основе лежит квазиньютоновский метод оптимизации, модифицированный для работы с ограничениями в виде бруса (гиперпараллелепипеда, грани которого параллельны осям координат).

Для обеспечения нахождения входных параметров в границах заданных интервалов $[x_i, \bar{x}_i]$ проводится их принудительное возвращение в эту область. Кроме того, параметры, вышедшие за установленные границы, фиксируются и исключаются из дальнейшего процесса оптимизации.

Метод был применен при исследовании допусков на ширину полосы пропускания микрополоскового ППФ на одинаковых связанных линиях с частотной характеристикой максимально плоского типа и следующими параметрами:

- граничные частоты полосы пропускания фильтра $f_{-п} = 945$ МГц, $f_{п} = 1055$ МГц по уровню вносимого затухания $a_{п} = 3$ дБ;
- граничные частоты полос заграждения $f_{-з} = 800$ МГц, $f_{з} = 1200$ МГц по уровню затухания $a_{з} \geq 30$ дБ;
- волновое сопротивление подводящих линий передачи $Z_0 = 50$ Ом.

Заданным исходным данным соответствует фильтр из четырех звеньев связанных микрополосковых линий ($n = 4$). Он реализован на подложке с диэлектрической проницаемостью $\epsilon_r = 9,6$ и толщиной $h = 1$ мм. Толщина металлизации в расчетах не учитывалась.

Анализ допусков проводился при отклонении параметров элементов конструкции ($W_i, S_i, l_i, h, \epsilon_r$) в пределах 2 %. В этом случае, при внесении отклонений в параметры конструкции фильтра, АЧХ фильтра в переходной области становится немонотонной. Отклонение затухания определялось на верхней границе полосы пропускания фильтра $f_{п} = 1055$ МГц, в связи с особенностями АЧХ (ее меньшей крутизной), вызванными вносимыми потерями и близостью паразитной полосы пропускания.

В этом случае максимальные отклонения функции вносимого затухания были равны:

- для внутренней интерполяции $a_{\max}(f) = 13,635$ дБ;
- для предложенного метода $a_{\max}(f) = 13,846$ дБ.

Полученные результаты свидетельствуют о возможности более точного определения гарантированных допусков при операции их синтеза с использованием предложенного метода анализа, что позволяет снизить требования к технологии изготовления микрополоскового фильтра.

4 УЧЕТ НЕМОНОТОННОСТИ АЧХ, ВЫЗВАННОЙ ЕЕ ФОРМОЙ

При проектировании активных полоснопропускающих высокоизбирательных РС-фильтров актуален вопрос отклонения центральной частоты полосы пропускания. Поэтому возникает задача анализа и синтеза допусков на параметры элементов схемы фильтра в зависимости от требований к смещению полосы пропускания фильтра.

Задача анализа смещения полосы пропускания фильтра по принципу наилучшего случая ставится как:

$$\text{найти } \underline{f}_0 = \text{MIN}(f_0), \quad \bar{f}_0 = \text{MAX}(f_0), \quad (3)$$

где центральная частота f_0 определяется в точке максимального значения коэффициента передачи:

$$H(f_0) \rightarrow \text{MAX}. \quad (4)$$

Минимальное \underline{f}_0 и максимальное \bar{f}_0 значения центральной частоты находятся на множестве точек входных параметров, ограниченном заданным интервалом $X \in [\underline{X}, \bar{X}]$.

Пусть известна функция, позволяющая определить центральную частоту для заданной точки X :

$$f_0 = \theta(X). \quad (5)$$

В этом случае, для решения задачи (3) необходимо найти максимум и минимум функции (5) на области $[\underline{X}, \bar{X}]$:

$$\theta(X) \rightarrow \text{MAX}, \quad \theta(X) \rightarrow \text{MIN} \quad (6)$$

при условии $X \in [\underline{X}, \bar{X}]$.

Зависимость (5) может быть получена из выражений для АЧХ аналитически или методами одномерной оптимизации.

В случае, если функция (5) монотонна, то задача (6) может быть решена анализом упрощенной интервальной модели этой функции. Если она немонотонна, то может быть применен квазиньютоновский метод [7], в котором в качестве целевой функции выступают выражения (6).

Предложенный метод был использован при исследовании РС-фильтра (рис. 2) с относительными параметрами: $R_0 = 1$, $R_1 = 100$, $R_2 = 1$, $C_1 = 0,0333$, $C_2 = 0,3$, $h_0 = 10$, $b_2 = 1$, $Q = 12$ [5]. Этим параметрам соответствует нормированная круговая частота полосы пропускания $\omega_0 = 1$. Анализ допусков проводился при отклонении параметров элементов схемы в пределах 5%. Отклонение полосы пропускания определялось по отклонению верхней точки АЧХ в полосе пропускания. В этом случае исследуемая функция имеет параболическую форму. При этом применение внутренней интерполяции нецелесообразно ввиду явно выраженной погрешности вычислений (большая кривизна). Поэтому предложенный метод анализа имеет огромные преимущества, в особенности для фильтров с очень узкими полосами пропускания. Результаты расчетов показывают, что при заданных отклонениях входных параметров максимальное отклонение центральной частоты составило: $\omega_{0_{\min}} = 0,7$; $\omega_{0_{\max}} = 1,564$.

В приведенном примере предлагаемый метод анализа является наиболее точным и быстрым методом для получения гарантированных допусков.

ВЫВОДЫ

Совместное применение метода анализа немонотонных выходных функций и интервального анализа позволяет разработать эффективный метод назначения гарантированных допусков и подходы для параметрического и допускового анализа и синтеза СВЧ устройств. Использование предложенного метода дает возможность решения задач с немонотонными выходными функциями и с нестандартными ограничениями. Направление дальнейших исследований связано с разработкой методов синтеза гарантированных допусков в условиях наличия локальных экстремумов выходных функций.

ПЕРЕЧЕНЬ ССЫЛОК

1. Конструкторско-технологические основы проектирования полосковых микросхем. / Под ред. И. П. Бушминского. – М.: Радио и связь, 1987. – 272 с.
2. Воробьев Е. А. Расчет производственных допусков устройств СВЧ. – Л.: Судостроение, 1980. – 147 с.
3. Шило Г. М. Формування інтервальных моделей для обчислення допусків // Радіоелектроніка. Інформатика. Управління. – 2002. – № 1. – С. 90–95.
4. Кришук В. М., Фарафонов О. Ю., Романенко С. М., Онищенко В. Ф., Шило Г. М. Врахування допусків і оптимізація параметрів мікросмужкових смугових фільтрів на зв'язаних лініях // Вісник Національного університету «Львівська політехніка» «Комп'ютерні системи проектування. Теорія і практика». – 2005. – № 548. – С. 83–90.
5. Знаменский А. Е., Теплюк И. Н. Активные РС-фильтры. – М.: «Связь», 1970. – 280 с.
6. Бертсекас Д. Условная оптимизация и методы множителей Лагранжа. – М: «Радио и связь», 1987. – 398 с.
7. Гапоненко Н. П., Вороний А. Ю. Анализ допусков для немонотонных выходных функций // Радиозелектроника и информатика. – 2005. – № 2. – С. 29–32.

Надійшла 11.12.06
 Після доробки 31.01.07

Запропоновано процедуру призначення допусків параметрів фільтрів з використанням методів математичного програмування для визначення глобального екстремуму АЧХ. Проведено дослідження впливу допусків на параметри пасивного мікросмужкового смугового фільтра на зв'язаних лініях з бічним зв'язком та активного смугового РС-фільтра з багатопетльовим зворотнім зв'язком. Показано переваги запропонованого методу в порів-

нянні з даними внутрішньої інтерполяції для інтервальних моделей.

The procedures designation of tolerances parameters filters with use of a method of mathematical programming for determination of global extreme amplitude frequency cha-

racteristic is offered. Probing research of tolerances on parameters of the passive microstrip band filter on coupled lines with side by side coupling and the active band RC-filter with a multiloop back coupling is carried out. It is shown advantages of the offered method in matching with the data of internal interpolation for interval models.

УДК 621.396.6.017:004.942

Н. П. Гапоненко, Е. В. Огренич

ПРОЕКТИРОВАНИЕ ПЛАСТИНЧАТЫХ РАДИАТОРОВ С ОПТИМАЛЬНЫМИ МАССОГАБАРИТНЫМИ ПАРАМЕТРАМИ

Анализируется тепловой режим пластинчатого радиатора при изменении его толщины. Показано, что при определенной толщине пластины обеспечивается минимальное значение массы или массогабаритных параметров радиатора. Даются рекомендации по выбору материалов.

ВВЕДЕНИЕ

В связи с ограниченным температурным диапазоном работы элементов электронных аппаратов при их проектировании возникает задача обеспечения заданного теплового режима. Эта задача обостряется требованиями к снижению габаритов радиоэлектронных устройств, что приводит к увеличению удельной плотности тепловыделения. Для ее решения используются теплоотводящие элементы и принудительные методы охлаждения [1–3]. Использование принудительных методов охлаждения связано с увеличением массы, энергопотребления и эксплуатационных затрат. Поэтому разработчики аппаратуры стараются в первую очередь обеспечить тепловой режим с помощью теплоотводящих устройств.

Самым популярным теплоотводящим устройством является радиатор. В настоящее время существует большое количество конструкций радиаторов. Методика их расчета предполагает предварительный выбор и проверочный расчет. Такой подход не позволяет выбирать оптимальную конструкцию радиатора, ограничивает число используемых критериев. Кроме того, известные методы расчета теплоотводящих устройств [4] не учитывают изменения коэффициентов теплоотдачи по радиатору. Не решена задача проектирования радиатора с минимальными массогабаритными критериями.

Целью работы является разработка методов проектирования пластинчатых радиаторов с минимальными массогабаритными показателями, разработка рекомендаций по выбору материалов и оценка влияния мест-

ных коэффициентов теплоотдачи на тепловой режим радиатора. Для решения поставленной задачи необходимо:

- разработать алгоритм проектирования радиатора с оптимальными массогабаритными показателями;
- разработать рекомендации по выбору материалов радиаторов;
- оценить погрешность, обусловленную неравным отводом тепла с поверхности радиатора.

1 ИССЛЕДОВАНИЕ СВОЙСТВ РАДИАТОРА

Пластинчатые радиаторы являются наиболее простыми и технологичными теплоотводящими устройствами. Они применяются при мощностях рассеивания не более 5 Вт. Но, несмотря на простоту конструкции, свойства таких радиаторов исследованы недостаточно.

Обычно толщина радиатора выбирается из условия минимальных перепадов температуры по радиатору. Однако это приводит к неоправданному увеличению его массы. С другой стороны, малая толщина увеличивает перепады температур и снижает среднюю температуру радиатора. Тепловой режим источника тепла тогда обеспечивается увеличением площади радиатора. Если уменьшение толщины значительно, то это может также привести к увеличению массы.

Различный вклад этих факторов обуславливает возможность применения различных критериев при проектировании. Для пластинчатых радиаторов могут использоваться критерии минимальной массы (M -критерий), минимальной площади (S -критерий) и массогабаритный критерий:

$$MS = m \cdot S, \quad (1)$$

где m и S – масса и площадь радиатора.

Этим критериям соответствуют M -, S - и MS -стратегии проектирования радиаторов.

Тепловая модель простейшего пластинчатого радиатора показана на рис. 1.

Если пренебречь отводом тепла в области источника, то процессы теплопередачи в такой модели описываются системой уравнений:

$$\left. \begin{aligned} \frac{d^2\vartheta}{dr^2} + \frac{1}{r} \frac{d\vartheta}{dr} - b^2\vartheta &= 0 \\ P &= -\lambda \frac{d\vartheta}{dr} S_1, (r = r_1) \\ \frac{d\vartheta}{dr} &= 0, (r = r_2) \end{aligned} \right\}, \quad (2)$$

где $\vartheta = t - t_c$ – перепад температур между точкой на радиаторе и температурой окружающей среды t_c ; $b^2 = \frac{a_1 + a_2}{\lambda \delta}$ – коэффициент, учитывающий теплоотдачу в окружающую среду; a_1 и a_2 – коэффициенты теплоотдачи в окружающую среду с нижней и верхней поверхностей радиатора; λ – коэффициент теплопроводности материала; δ – толщина радиатора; P – мощность источника тепла; $S_1 = 2\pi r_1 \delta$ – поверхность отвода тепла от источника в радиатор; r_1 и r_2 – радиусы источника тепла и радиатора.

Решение системы уравнений (1) имеет вид [1]:

$$\vartheta(r) = \frac{P}{\lambda b S_1} \frac{I_1(br_2)K_0(br) + I_0(br)K_1(br_2)}{I_1(br_2)K_1(br_1) - I_1(br_1)K_1(br_2)}, \quad (3)$$

где I_0 и I_1 – модифицированные функции Бесселя первого рода нулевого и первого порядков; K_0 и K_1 – модифицированные функции Бесселя второго рода нулевого и первого порядков.

При значительной толщине радиатора перепады температуры по его длине незначительны. Необходимый

тепловой режим тогда обеспечивается радиатором с радиусом:

$$r_2 = \sqrt{r_1^2 + \frac{P}{\pi(a_1 + a_2)(t_p - t_c)}}, \quad (4)$$

где t_p – среднеповерхностная температура радиатора.

Условие (4) соответствует минимальной площади теплоотводящей поверхности (S -критерий). Другие стратегии проектирования не приводят к явным выражениям для размеров радиатора. Возможность применения этих стратегий исследовалась на алюминиевом радиаторе с температурой источника тепла $t_{\max} = 60^\circ\text{C}$, температурой окружающей среды $t_c = 40^\circ\text{C}$, мощностью источника тепла $P = 4$ Вт и радиусом $r_1 = 5$ мм. Задавался коэффициент теплоотдачи $a_1 + a_2 = 20$ Вт/м²К.

Начальное значение радиуса r_2 определялось с помощью соотношения (4). Затем для каждой толщины δ подбирался радиус радиатора r_2 из условия обеспечения теплового режима:

$$\vartheta(r_1) + t_c = t_{\max}. \quad (5)$$

График зависимости массы радиатора от его толщины приведен на рис. 2. Рисунок показывает, что изменение толщины является эффективным средством уменьшения массы радиатора. Применение M -стратегии по сравнению с S -стратегией позволяет уменьшить массу радиатора почти в 3 раза.

Зависимость площади радиатора от его толщины является монотонно убывающей функцией (рис. 3). После толщины $\delta = 2$ мм площадь радиатора практически не изменяется. Радиус радиатора тогда может рассчитываться с помощью выражения (4).

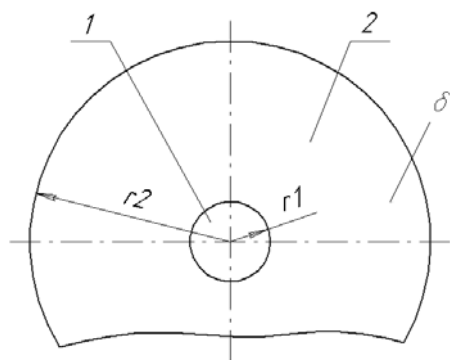


Рисунок 1 – Тепловая модель радиатора:
1 – область источника тепла; 2 – область теплоотвода

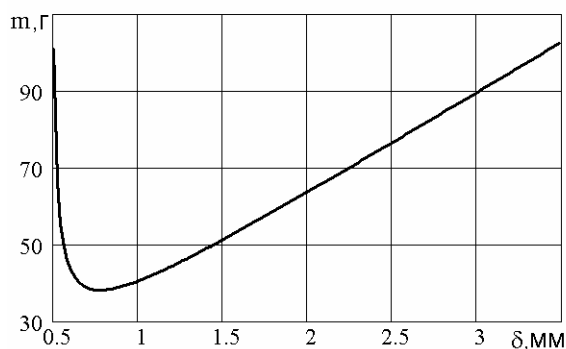


Рисунок 2 – Зависимость массы радиатора от его толщины

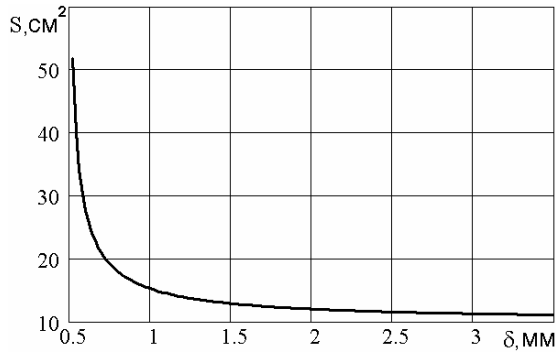


Рисунок 3 – Зависимость площади радиатора от его толщины

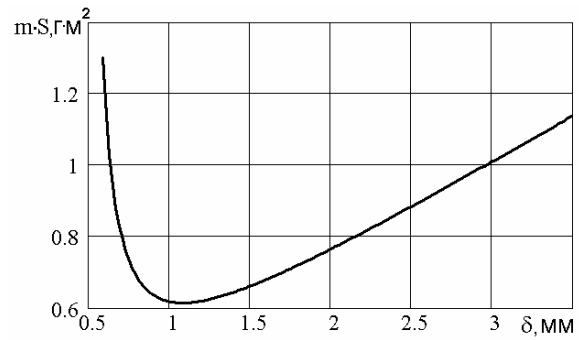


Рисунок 4 – Зависимость произведения площади на массу радиатора от его толщины

К промежуточным показателям приводит *MS*-стратегия проектирования (рис. 4). Она позволяет значительно уменьшить массогабаритные показатели. Сравнение параметров радиатора при *M*- и *MS*-стратегиях показывают, что *MS*-стратегия уменьшает площадь радиатора на 25 %, увеличивает массу на 15 %. Толщина радиатора при этом увеличивается почти в 1,5 раза, что позволяет увеличить жесткость пластины.

В графиках, показанных на рис. 2–4, не учтена нелинейная зависимость коэффициентов теплоотдачи от температуры радиатора. Эту особенность процессов теплопередачи необходимо учесть при разработке алгоритмов проектирования радиаторов.

2 ПРОЕКТИРОВАНИЕ РАДИАТОРА

При проектировании оказывается необходимым производить уточнение не только толщины радиатора, его радиуса, но и коэффициента теплоотдачи с поверхности. Кроме того, для расчета коэффициента теплоотдачи необходимо вычислять среднеповерхностную температуру радиатора:

$$t_m = t_c + \frac{2}{r_2^2 - r_1^2} \int_{r_1}^{r_2} \theta(r) r dr. \quad (6)$$

Нелинейные свойства коэффициента теплоотдачи учитываются с помощью итерационного алгоритма, который для *M*-стратегии проектирования радиатора имеет вид:

Шаг 1. Уточняется значение толщины радиатора $\delta^{(i)}$;

Шаг 2. Уточняется коэффициент теплоотдачи $a^{(i)} = a_1^{(i)} + a_2^{(i)}$;

Шаг 3. Уточняется $r_2^{(i)}$ из условия обеспечения теплового режима (5) с учетом (3);

Шаг 4. Определяется среднеповерхностная температура $t_m^{(i)}$ с помощью выражения (6);

Шаг 5. Проверяется условие:

$$\frac{t_m^i - t_m^{(i-1)}}{t_m^i} \leq \epsilon_t, \quad (7)$$

где ϵ_t – погрешность расчета среднеповерхностной температуры.

Если условие (7) не выполняется, то переход на шаг 2.

Шаг 6. Рассчитывается масса радиатора $m^{(i)}$;

Шаг 7. Проверяется условие:

$$m^{(i)} \rightarrow \min. \quad (8)$$

Если условие (8) не выполняется, то переход на шаг 1.

Шаг 8. Конец алгоритма.

В алгоритме коэффициент теплоотдачи рассчитывается с использованием критериальных уравнений [2]. Начальное значение коэффициента теплоотдачи выбирается на основе среднестатистических данных в зависимости от способа охлаждения и вида теплоносителя. Значение $r_2^{(0)}$ определяется из условия (4) для минимальной площади радиатора. Толщина $\delta^{(0)}$ может задаваться из условия равенства перепадов температур на цилиндрической стенке и перепадов температур источник-окружающая среда. Это приводит к начальному значению:

$$\delta^{(0)} = \frac{P}{2\pi\lambda(t_{\max} - t_c)} \ln \frac{r_2}{r_1}. \quad (9)$$

M – алгоритм использовался для сравнения эффективности различных материалов радиаторов. Результаты расчетов для рассмотренного ранее источника тепла и температуры окружающей среды при свободной конвекции и коэффициент черноты радиатора $\epsilon = 0,9$ приведены в табл. 1.

Таблица 1 – Эффективность использования материалов при *M*-стратегии

Материал	$d_{\text{опт}}$, мм	S , см ²	m , г	$m \cdot S$, г·см ²
Медь (М1)	0,38	172	58,4	1,004
Латунь (Л59)	0,59	173	90,1	1,556
Al сплавы (Ал2)	0,84	171	38,1	0,652
Mg сплавы (МА-1)	1,27	171	38,0	0,649
Сталь	2,94	171	388	6,652

Таблица 2 – Эффективность использования материалов при *MS*-стратегии

Материал	$d_{\text{опт}}$, мм	S , см ²	m , г	$m \cdot S$, г·см ²
Медь (М1)	0,54	134	64,5	0,863
Латунь (Л59)	0,85	133	100	1,332
Al сплавы (Ал2)	1,20	133	42,3	0,563
Mg сплавы (МА-1)	1,80	133	42,1	0,562
Сталь	4,17	134	430	5,756

Из табл. 1 видно, что наименьшую массу обеспечивают алюминиевые и магниевые сплавы. По массовым показателям радиаторы из магниевых сплавов не имеют преимуществ перед алюминиевыми, но имеют большую толщину. В связи с высокой стоимостью их применение в *M*-стратегии возможно при необходимости увеличения жесткости пластины. Следует также отметить, что при *M*-стратегии коэффициент теплопроводности не влияет на площадь радиатора.

Алгоритм проектирования радиатора на основе *MS*-стратегии отличается только шагами 6 и 7, которые принимают вид:

...

Шаг 6. Рассчитывается произведение массы радиатора на его площадь $m^{(i)}S^{(i)}$;

Шаг 7. Проверяется условие:

$$m^{(i)}S^{(i)} \rightarrow \min . \quad (10)$$

Если условие (9) не выполняется, то переход на шаг 1.

Результаты расчета радиатора с использованием *MS*-стратегии приведены в табл. 2.

Данные табл. 2 свидетельствуют, что свойства *MS*-стратегии аналогичны свойствам *M*-стратегии. По-прежнему наиболее эффективным материалом для радиаторов являются алюминиевые сплавы. По сравнению с *M*-стратегией достигается: увеличение толщины на 42 %, уменьшение площади на 28 %, увеличение массы на 11 %, уменьшение *MS*-критерия на 16 %.

3 ОЦЕНКА ПОГРЕШНОСТИ

Приведенные алгоритмы проектирования не учитывают неравномерность теплоотвода по радиатору, что может привести к погрешности определения оптимальных размеров. С целью определения погрешности бы-

ло проведено моделирование тепловых процессов в радиаторе методом конечных элементов. В результате моделирования установлено, что при горизонтальном расположении радиатора температуры источников тепла в обоих случаях не отличались. Это говорит о том, что предложенные алгоритмы могут без ограничений использоваться для проектирования горизонтально расположенных пластинчатых радиаторов.

Результаты моделирования для вертикально расположенного радиатора при *M*-стратегии показаны на рис. 5.

Увеличение интенсивности теплоотвода за счет увеличения скорости движения теплоносителя в этом случае привело к увеличению коэффициента теплоотдачи и снижению температуры источника на 1,6 градуса. Использование измененного коэффициента теплоотдачи в предложенных алгоритмах проектирования показало, что вертикальное расположение радиатора позволяет уменьшить его массу на 14 %, площадь на 20 % и увеличить толщину на 7 %. Аналогичные результаты получены и для *MS*-стратегии.

Таким образом, использование предложенных алгоритмов при вертикальном расположении радиатора

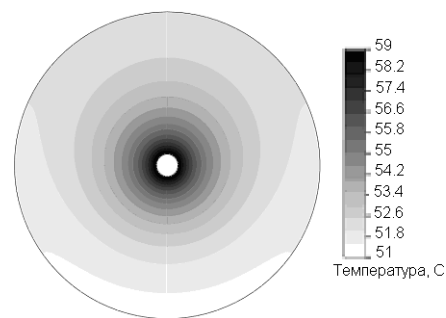


Рисунок 5 – Распределение температуры по вертикально расположенному радиатору

облегчает его тепловой режим, но не позволяет достичь оптимальных соотношений. Их применение оправдано в случаях, когда требования к массогабаритным показателям не слишком высоки. Иначе необходимо проводить моделирование процессов теплоотдачи методами конечных элементов и пересчитывать размеры конструкции в соответствии с измененным коэффициентом теплоотдачи.

ВЫВОДЫ

Анализ процессов теплопередачи в пластинчатых радиаторах показал, что при их проектировании возможно значительное уменьшение их массогабаритных параметров. Разработанный алгоритм позволяет автоматизировать определение оптимальных размеров дисковых радиаторов минимальной массы и массогабаритов. Учитываются нелинейные свойства коэффициентов теплоотдачи. По сравнению с *S*-стратегией *MS*-стратегия позволяет в несколько раз уменьшить массу радиатора. *MS*-стратегия позволяет увеличивать жесткость радиатора за счет увеличения толщины пластины при незначительном увеличении массы. Площадь радиатора при этом уменьшается до 30 %.

Исследование эффективности материалов показало, что наилучшие показатели обеспечивают алюминиевые и магниевые сплавы. Радиаторы из магниевых сплавов отличаются только толщиной радиатора, которая увеличивается до 50 %. В связи с ценовыми ограничениями магниевые сплавы могут использоваться при необходимости обеспечения жесткости пластины радиатора.

УДК 621.396.6.004:004.942

А. О. Намлинський

ІНТЕРВАЛЬНІ КОЕФІЦІЄНТИ ЗОВНІШНІХ ВПЛИВІВ В ЗАДАЧАХ ВИБОРУ ЕЛЕМЕНТІВ

Запропоновано метод призначення допусків на параметри елементів. Враховано вплив зовнішніх чинників на параметри елементів. Коефіцієнти зовнішніх впливів представлені інтервалами. Наведено реалізації стратегій максимального об'єму допусків і максимального об'єму коефіцієнтів зовнішніх впливів.

ВСТУП

Задачу вибору елементів можна подати як задачу призначення їх параметрів згідно деяких прийнятих критеріїв. Номінальне значення елемента призначається

© Намлинський А. О., 2007

Погрешність пропозованих алгоритмів перевірялась порівнянням з результатами моделювання процесів теплоотдачи методом кінцевих елементів. Результати моделювання показали, що при горизонтальному розташуванні пластини пропозовані алгоритми можуть використовуватися без обмежень. В випадку вертикального розташування пластини існує можливість подальшого удосконалення массогабаритних показателів. Но в цьому випадку необхідно уточнювати коефіцієнти теплоотдачи на основі результатів моделювання.

ПЕРЕЧЕНЬ ССЫЛОК

1. Дульнев Г. Н. Теплообмен в радиоэлектронных устройствах. – М. – Л.: Госэнергоиздат, 1963. – 288 с.
2. Дульнев Р. Н. Тепло- и массообмен в радиоэлектронной аппаратуре. – М.: Высш. шк., 1984. – 247 с.
3. Роткоп А. А., Спокойный Ю. Е. Обеспечение тепловых режимов при конструировании радиоэлектронной аппаратуры. – М.: Советское радио, 1976. – 232 с.
4. Ройзен А. И., Дулькин И. Н. Теловой расчет оребренных поверхностей. под ред. В. Г. Фастовского. – М.: Энергия, 1977. – 256 с.

Надійшла 15.08.06
Після доробки 21.09.06

Аналізується тепловий режим пластинчатого радіатора при зміні його товщини. Показано, що при деякій товщині пластини забезпечується мінімальне значення маси або массогабаритних параметрів радіатора. Приводяться рекомендації до вибору матеріалів

The thermal regime of the flanged radiator is analyzed at change of his thickness. It is shown, that at the certain thickness of a plate the minimal value of weight or weight clearance parameters of a radiator is provided. Recommendations at sampling materials are given.

на етапі схемотехнічного проектування. Існує велика кількість явищ, що обумовлюють відхилення виміряного значення основного параметра від його номіналу. Такі відхилення порушують нормальні режими роботи пристрою – впливають на його працездатність. Найбільший вклад у такі відхилення вносять технологічні процеси виробництва та вплив на пристрій зовнішніх чинників під час його експлуатації.

Більшість методик вибору елементів спрощують цю задачу внаслідок врахування лише технологічних відхилень виробництва, що характеризуються номіналь-

ними допусками елементів [1–3]. Сучасний рівень розвитку радіоелектронної промисловості ставить жорсткі умови до точності параметрів пристроїв. У такому випадку відхилення, що з'являються під впливом зовнішніх чинників є співрозмірними до номінального допуску елементу. Тому врахування впливу зовнішніх чинників у процесі обчислення параметрів елементу є важливим.

1 ІНТЕРВАЛЬНІ КОЕФІЦІЄНТИ ЗОВНІШНІХ ВПЛИВІВ

Відхилення під впливом зовнішніх чинників можна враховувати у моделях радіоелектронного пристрою. Однією з особливостей елементної бази є змінний характер коефіцієнта зовнішніх впливів (КЗВ) під час експлуатації (рис. 1).

У такому випадку його значення під час експлуатації може бути представлено інтервалом

$$\alpha_i = [\alpha_{mi}; \alpha_{pi}],$$

де α_m , α_p – межові значення КЗВ в нижньому та верхньому діапазонах зовнішніх впливів.

Нелінійна залежність параметра від зовнішнього чинника можна уявити у вигляді коефіцієнта пропорційності:

$$i = \alpha_{mi}; k_{i \text{ mi}}, \quad (1)$$

де $k_{\alpha i}$ – коефіцієнт пропорційності.

На значення КЗВ накладаються додаткові обмеження, що можуть формуватися властивостями матеріалу або іншими чинниками і бути вираженими нерівністю:

$$\alpha_{mi \text{ min}} \leq \alpha_{mi} \leq \alpha_{mi \text{ max}}. \quad (2)$$

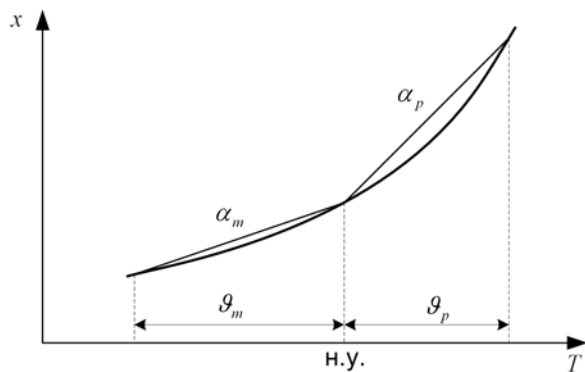


Рисунок 1 – Інтервальний КЗВ обумовлений нелінійністю параметра елемента

За допомогою спрощеної інтервальної моделі [4] обмеження на параметри елементів можуть бути записані у вигляді системи (3):

$$\begin{cases} a_0 + \sum_{i=1}^n a_i x_i (1 + \alpha_m \vartheta_m) \geq y_e; \\ a_0 + \sum_{i=1}^n a_i x_i (1 + \alpha_p \vartheta_p) \geq y_e; \\ \bar{a}_0 + \sum_{i=1}^n \bar{a}_i \bar{x}_i (1 + \alpha_m \vartheta_m) \leq \bar{y}_e; \\ \bar{a}_0 + \sum_{i=1}^n \bar{a}_i \bar{x}_i (1 + \alpha_p \vartheta_p) \leq \bar{y}_e, \end{cases} \quad (3)$$

де a_0 , \bar{a}_0 , a_i , \bar{a}_i – коефіцієнти моделі; x_i , \bar{x}_i – нижня та верхня межі номінального значення параметрів елементів; ϑ_m , ϑ_p – межові значення інтервалу зовнішніх впливів; y_e , \bar{y}_e – нижня та верхня експлуатаційні межі вихідної функції.

Неоднозначність рішення такої задачі призводить до встановлення критерію його оптимальності. У роботі використовуються критерії максимальних номінальних допусків (МНД) та максимальних коефіцієнтів зовнішніх впливів (МКЗВ) [5].

2 СТРАТЕГІЯ МНД

Критерієм оптимальності для стратегії МНД є [5]

$$V_\delta = \prod_{i=1}^n \delta_{xi}; \quad V_\delta \rightarrow \text{MAX},$$

де δ_{xi} – допуск вхідного параметра.

Декомпозиція задачі (3) для цієї стратегії дозволяє записати її у вигляді [5]:

$$\begin{cases} \sum_{i=1}^n a_i \delta_{xi} \leq \delta_{yr}; \\ \sum_{i=1}^n \bar{a}_i \bar{\delta}_{xi} \leq \bar{\delta}_{yr}, \end{cases} \quad \begin{cases} \sum_{i=1}^n s_{xi}^{\text{LM}} \leq s_y^{\text{LM}}; \\ \sum_{i=1}^n s_{xi}^{\text{LP}} \geq s_y^{\text{LP}}; \\ \sum_{i=1}^n s_{xi}^{\text{HM}} \geq s_y^{\text{HM}}; \\ \sum_{i=1}^n s_{xi}^{\text{HP}} \leq s_y^{\text{HP}}, \end{cases} \quad (4)$$

де \underline{s}_{xi} , \bar{s}_{xi} – нижній та верхній допуски вхідних параметрів; $\underline{\delta}_{yr}$, $\bar{\delta}_{yr}$ – номінальні нижній та верхній допуски вихідної функції; s_{xi}^{LM} , s_{xi}^{LP} , s_{xi}^{HM} , s_{xi}^{HP} – фактори компенсації вхідних параметрів. Верхній індекс: L, H – нижня та верхня межі вихідної функції, M, P – нижній та верхній діапазони зовнішніх впливів:

$$s_{xi}^{LM} = a_i \underline{x}_i \alpha_{mi}, \quad s_{xi}^{LP} = a_i \underline{x}_i \alpha_{pi}, \quad s_{xi}^{HM} = \bar{a}_i \bar{x}_i \alpha_{mi}, \\ s_{xi}^{HP} = \bar{a}_i \bar{x}_i \alpha_{pi};$$

s_y^{LM} , s_y^{LP} , s_y^{HM} , s_y^{HP} – фактори компенсації вихідної функції:

$$s_y^{LM} = \frac{(1-\underline{\lambda})(\underline{y}_e - y_r)}{\mathfrak{G}_m}, \quad s_y^{LP} = \frac{(1-\underline{\lambda})(\underline{y}_e - y_r)}{\mathfrak{G}_p}, \\ s_y^{HM} = \frac{(1-\bar{\lambda})(\bar{y}_e - y_r)}{\mathfrak{G}_m}, \quad s_y^{HP} = \frac{(1-\bar{\lambda})(\bar{y}_e - y_r)}{\mathfrak{G}_p},$$

де $\underline{\lambda}$, $\bar{\lambda}$ – рівні компенсації для нижньої та верхньої меж вихідної функції відповідно; y_r – номінальне значення вихідної функції.

Перша система вирішується згідно методик, запропонованих у [6]. Призначення КЗВ у випадку одного з обмежень наведено у [5]. У випадку системи обмежень область рішень для одного обмеження може не співпадати або частково співпадати з областями рішень для інших обмежень. Область рішення такої системи може бути знайдена як область пересічення інтервалів рішення всіх обмежень. Для цього необхідно обрати будь-яке обмеження і відобразити на область його рішення області рішення інших обмежень. Таке обране обмеження будемо називати базовим.

Таким чином область рішень може бути представлена у вигляді інтервалу на області базового обмеження. Оберемо за базове перше обмеження системи (4), що відповідає нижній межі вихідної функції при нижньому діапазоні зовнішніх впливів. Тоді ліву межу інтервалу рішень будуть формувати обмеження з протилежним знаком нерівності: друге та третє нерівняння системи. Праву межу – перше (базове) і четверте нерівняння.

Область рішення являє собою інтервал значень сум факторів компенсації. Внаслідок нетривіальності уявлення таких сум, відображення точки такого інтервалу на область рішень базової задачі є інтервалом. Межі цього інтервалу знаходяться наступним чином. Формується множина \mathbf{S}_x :

$$\mathbf{S}_x = \{s_{xi} | k_i > k_{i+1}, i = \overline{1, n}\}.$$

Тоді відображенням деякої суми σ на область базової задачі $base$ є інтервал σ^* :

$$base \sigma^*; \quad \sigma^* = \sigma_{min}^*; \sigma_{max}^*, \\ \sigma_{min}^* = \sum_{i=1}^n k_i s_{xi} + \sum_{i=tmin}^n k_i wid(\mathbf{s}_{xi}) + k_{tmin-1} \Delta s_{xtmin-1}; \\ \sigma_{max}^* = \sum_{i=1}^n k_i s_{xi} + \sum_{i=1}^{tmax} k_i wid(\mathbf{s}_{xi}) + k_{tmax+1} \Delta s_{xtmax+1}, \quad (5)$$

$$\Delta s_{xtmin-1} = \sigma - \sum_{i=1}^n s_{xi} - \sum_{i=tmin}^n wid(\mathbf{s}_{xi}), \\ \Delta s_{xtmax+1} = \sigma - \sum_{i=1}^n s_{xi} - \sum_{i=1}^{tmax} wid(\mathbf{s}_{xi}),$$

де k_i – коефіцієнт пропорційності параметра між обмеженням, з якого відбувається проекція, та базовим обмеженням; s_{xi} – нижня межа інтервалу фактора компенсації, що утворюється обмеженням (2):

$$\mathbf{s}_{xi} = [s_{xi}; \bar{s}_{xi}] = [a_i x_i \alpha_{imin}; a_i x_i \alpha_{imax}];$$

wid – функція, що повертає значення ширини інтервалу; $tmin$, $tmax$ – індекси множини \mathbf{S}_x , такі, що:

$$\sum_{i=1}^{tmax} wid(\mathbf{s}_{xi}) \leq \sigma - \sum_{i=1}^n s_{xi} \leq \sum_{i=1}^{tmax+1} wid(\mathbf{s}_{xi}), \\ \sum_{i=tmin}^n wid(\mathbf{s}_{xi}) \leq \sigma - \sum_{i=1}^n s_{xi} \leq \sum_{i=tmin-1}^n wid(\mathbf{s}_{xi}).$$

Значення факторів компенсації, що утворюють ліву межу інтервалу проекції:

$$s_{xi} = s_{xi}, \quad i = \overline{1, tmin-2}; \\ s_{xtmin-1} = s_{xtmin-1} + \Delta s_{xtmin-1}; \\ s_{xi} = \bar{s}_{xi}, \quad i = \overline{tmin, n},$$

праву:

$$s_{xi} = \bar{s}_{xi}, \quad i = \overline{1, tmax}; \\ s_{xtmax+1} = s_{xtmax+1} + \Delta s_{xtmax+1}; \\ s_{xi} = s_{xi}, \quad i = \overline{tmax+2, n}.$$

Задача знаходження лівої межі інтервалу рішень виглядає як

$$\sum_{i=1}^n s_{xi}^{LM} \rightarrow \text{MIN}; \quad \begin{cases} \sum_{i=1}^n s_{xi}^{HM} \geq s_y^{HM}; \\ \sum_{i=1}^n s_{xi}^{LP} \geq s_y^{LP}. \end{cases}$$

Для знаходження такої межі визначимо за допомогою (5) нижні межі проєкцій s_y^{HM} , s_{LM}^{HM*} і s_y^{LP} , s_{LM}^{LP*} . Визначимо додатково точки:

$$s_y^{HM'} = \sum_{i=1}^n \frac{k_{ai}}{\gamma_i} s_{xi}^{HM*}; \quad s_y^{LP'} = \sum_{i=1}^n \frac{\gamma_i}{k_{ai}} s_{xi}^{LP*},$$

де $s_y^{HM'}$, $s_y^{LP'}$ – фактори компенсації вихідної функції на областях LP і HM, що відповідають s_y^{HM*} і s_y^{LP*} ; γ_i – коефіцієнт пропорційності параметра між обмеженнями для різних меж вихідної функції і одного діапазону зовнішніх впливів:

$$\gamma_i = \bar{a}_i \bar{x}_i / \underline{a}_i \underline{x}_i;$$

s_{xi}^{HM*} , s_{xi}^{LP*} – значення факторів компенсації вхідних параметрів, що відповідають нижній межі проєкцій s_y^{HM} , s_{LM}^{HM*} і s_y^{LP} , s_{LM}^{LP*} , відповідно.

Існує три випадки взаємного розташування точок $s_y^{HM'}$, $s_y^{LP'}$ відносно інтервалів прийнятних значень для обмежень LP і HM відповідно:

1) $s_y^{HM'} \geq s_y^{LP}$, $s_y^{LP'} < s_y^{HM}$. Нижня межа інтервалу рішень системи (4) на області базової задачі:

$$\underline{R} = \sum_{i=1}^n s_{xi}^{HM*} / \gamma_i,$$

де \underline{R} – нижня межа інтервалу рішень системи.

2) $s_y^{LP'} \geq s_y^{HM}$, $s_y^{HM'} < s_y^{LP}$. Нижня межа інтервалу рішень системи (4) на області базової задачі:

$$\underline{R} = \sum_{i=1}^n s_{xi}^{LP*} / k_{ai}.$$

3) $s_y^{LP'} < s_y^{HM}$, $s_y^{HM'} < s_y^{LP}$. Нижня межа інтервалу рішень визначається перетином гіперплощин, що відповідають обмеженням 2-ї і 3-ї системи.

В останньому випадку можна скористатися наступною методикою. Сформуємо дві множини:

$$\Delta \mathbf{S}_G = \{\Delta s_{xgi} | 1/\gamma_i > 1/\gamma_{i+1}\}, \quad \Delta s_{xgi} = s_{xi}^{HM} - s_{xi}^{HM'};$$

$$\Delta \mathbf{S}_L = \{\Delta s_{xli} | k_{ai} > k_{ai+1}\}, \quad \Delta s_{xli} = s_{xi}^{HM} - s_{xi}^{HM'},$$

де $\Delta \mathbf{S}_G$, $\Delta \mathbf{S}_L$ – множини можливих відхилень факторів компенсації у сторону зменшення та збільшення відповідно; s_{xi}^{HM} – оптимальне рішення для обмеження HM.

Необхідно змінити значення s_{xi}^{HM} таким чином, щоб його проєкція на LP дорівнювала s_y^{LP} , а проєкція на LM була б мінімальна. Алгоритм такої зміни:

1. Ініціалізація: $i = 1$, $j = 1$, $d = s_y^{LP} - s_y^{HM'}$;
2. Якщо $\Delta s_{x1i} = 0$, то $i = i + 1$;
3. Якщо $\Delta s_{x2j} = 0$, то $j = j + 1$;
4. $\Delta s_{\min} = \min(\Delta s_{x1i}, \Delta s_{x2j})$,
 $\Delta d = \Delta s_{\min}(k_{ai}/\gamma_i - k_{aj}/\gamma_j)$;
5. Якщо $\Delta d < d$, $\Delta s_{x1i} = \Delta s_{x1i} - \Delta s_{\min}$, $\Delta s_{x2j} = \Delta s_{x2j} - \Delta s_{\min}$, перехід на крок 2;
6. $\Delta s_{\min} = d / (k_{ai}/\gamma_i - k_{aj}/\gamma_j)$,
 $\Delta s_{x1i} = \Delta s_{x1i} - \Delta s_{\min}$, $\Delta s_{x2j} = \Delta s_{x2j} - \Delta s_{\min}$;
7. Кінець.

Задача знаходження правої межі інтервалу рішення виглядає як

$$\sum_{i=1}^n s_{xi}^{LM} \rightarrow \text{MAX}; \quad \begin{cases} \sum_{i=1}^n s_{xi}^{LM} \leq s_y^{LM}; \\ \sum_{i=1}^n s_{xi}^{HP} \leq s_y^{HP}. \end{cases}$$

Для її вирішення знайдемо проєкцію s_y^{LM} , s_{HP}^{LM*} . Якщо інтервал проєкції і інтервал прийнятних значень області HP перетинаються:

$$s_y^{LM*} < s_y^{HP},$$

то права межа інтервалу рішень:

$$\bar{R} = s_y^{LM*}.$$

В іншому випадку необхідно знайти таке зміщення Δs_y^{LM} , що

$$s_y^{LM} + s_y^{LM} \cdot \sigma_x^{LM*}, \quad \sigma_x^{LM*} = s_y^{HP}; \quad \Delta s_y^{LM} \rightarrow \text{MAX}.$$

Сформуємо множину:

$$\mathbf{DS}_x^{LM} = \left\{ ds_{xi}^{LM} \mid k_{ai}\gamma_i > k_{ai+1}\gamma_{i+1} \right\},$$

$$ds_{xi}^{LM} = s_{xi}^{LM} - s_{xi}^{LM}.$$

Тоді задача знаходження зміщення Δs_y^{LM} може бути представлена як задача знаходження такого t , що

$$\sum_{i=1}^t ds_{xi} \cdot k_{ai}\gamma_i \leq \sigma_x^{LM*} - s_y^{HP} < \sum_{i=1}^{t+1} ds_{xi} \cdot k_{ai}\gamma_i.$$

Фактори компенсації s_{xi}^{LM} при цьому мають значення

$$s_{xi}^{LM} = \underline{s}_{xi}^{LM}, \quad i = \overline{1, t};$$

$$s_{xt+1}^{LM} = \underline{s}_{xt+1}^{LM} + ds_{xt+1} - \left(\sigma_x^{LM*} - s_y^{HP} - \sum_{i=1}^t ds_{xi} \cdot k_{ai} \gamma_i \right);$$

$$s_{xi}^{LM} = \bar{s}_{xi}^{LM} + ds_{xi}, \quad i = \overline{t+2, n}.$$

Якщо ширина інтервалу рішення є додатною $\text{wid}(\mathbf{R}) > 0$, то виникає задача пошуку таких значень факторів компенсації вхідних параметрів, що характеризуються максимальним об'ємом серед всіх комбінацій факторів, сума яких належить інтервалу рішень. Якщо оптимальне рішення для LM лежить у межах інтервалу рішення \mathbf{R} , то це рішення приймається як оптимальне і для системи. У іншому випадку рішенням обирається точка інтервалу \mathbf{R} , що лежить ближче до точки оптимального рішення для LM.

Якщо $\text{wid}(\mathbf{R}) < 0$, то це вказує на відсутність рішення. У цьому випадку необхідна корекція рівня компенсації вихідної функції. Така зміна відбувається шляхом визначення взаємовиключних пар обмежень і необхідних змін рівнів компенсації. Взаємовиключні пари обмежень можуть бути обмеженнями для однієї з меж вихідної функції. У такому випадку зміні належить лише одна з меж рівня компенсації. Якщо обмеження відносяться до різних меж вихідної функції, то змінюватися будуть обидві межі рівня компенсації. Відповідно до критерію стратегії, такі зміни мають бути мінімальні.

Результуючі зміни рівнів компенсації при цьому:

$$\Delta \underline{\lambda} = \max(\Delta \underline{\lambda}_{LP}^{LM}, \Delta \underline{\lambda}_{HM}^{LM}, \Delta \underline{\lambda}_{HP}^{LP});$$

$$\Delta \bar{\lambda} = \max(\Delta \bar{\lambda}_{HP}^{HM}, \Delta \bar{\lambda}_{LM}^{HM}, \Delta \bar{\lambda}_{LP}^{HP}),$$

де $\Delta \underline{\lambda}$, $\Delta \bar{\lambda}$ – нижня та верхня зміни рівня компенсації; $\Delta \underline{\lambda}_{LP}^{LM}$, $\Delta \underline{\lambda}_{HM}^{LM}$, $\Delta \underline{\lambda}_{HP}^{LP}$, $\Delta \bar{\lambda}_{HP}^{HM}$, $\Delta \bar{\lambda}_{LM}^{HM}$, $\Delta \bar{\lambda}_{LP}^{HP}$ – зміни меж рівнів компенсації обчислені для пар обмежень. Нижній та верхній індекси ідентифікують обмеження: L, H – нижня та верхня межі вихідної функції, M, P – нижній та верхній діапазони зовнішніх впливів.

Проекція фактора компенсації вихідної функції з одного обмеження пари на інше може вказувати на наявність їх взаємовиключення. Якщо інтервал проекції і інтервал допустимих значень базової задачі не перетинаються, то такі обмеження є взаємовиключні.

Якщо за базове брати обмеження типу $\sum_{i=1}^n s_{xi} < s_y$, то існує лише один вид взаємного розташування інтервалів допустимих значень і проекції фактора компенсації вихідної функції для взаємовиключних обмежень (рис. 2).

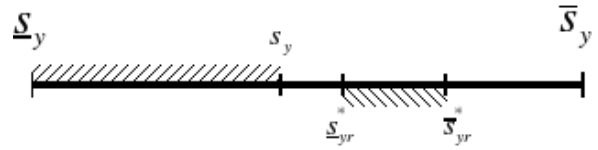


Рисунок 2 – Інтервали допустимих значень і проекції фактора компенсації вихідної функції

Розглянемо випадок для нижньої межі вихідної функції. Необхідна зміна рівня компенсації $\Delta \underline{\lambda}_{LP}^{LM}$ має бути такою, що

$$\sigma_x^{LP} = s_y^{LP} + \Delta \underline{\lambda}_{LP}^{LM} \cdot \delta_y / \vartheta_p; \quad \sigma_x^{LP} \leq \sigma_x^{LP*};$$

$$\sigma_x^{LP*} = s_y^{LM} + \Delta \underline{\lambda}_{LP}^{LM} \cdot \delta_y / \vartheta_m.$$

Сформуємо множину:

$$DS_x = \{ ds_{xi} | 1/k_{ai} > 1/k_{ai+1} \}, \quad ds_{xi} = s_{xi}^{LP} - \underline{s}_{xi}^{LP}.$$

Зі зменшенням Δs_y^{LP} на величину ds_{xi} відстань між інтервалами прийнятних значень і проекції змінюється як

$$\Delta d_i = ds_{xi} / k_{ai} - ds_{xi} \vartheta_p / \vartheta_m,$$

де Δd_i – часткові зміни відстані між проекцією та інтервалом прийнятних значень.

Тоді задачу пошуку зміни рівня компенсації можна записати як задачу знаходження t такого, що

$$\sum_{i=1}^t \Delta d_i \geq \underline{s}_y^{LP*} - s_y^{LM} \geq \sum_{i=1}^{t+1} \Delta d_i.$$

Зміна фактора компенсації для області LP і зміна рівня компенсації будуть дорівнювати

$$\Delta s_y^{LP} = \sum_{i=1}^{t+1} ds_{xi} - \frac{\sum_{i=1}^{t+1} \Delta d_i - (\underline{s}_y^{LP*} - s_y^{LM})}{1/k_{at+1} - \vartheta_p / \vartheta_m},$$

$$\Delta \underline{\lambda}_{LP}^{LM} = \frac{\Delta s_y^{LP} \vartheta_p}{\delta_y}.$$

Процедура обчислення зміни рівня компенсації для верхньої межі вихідної функції будується аналогічно.

У випадку, коли взаємовиключними є обмеження, що відповідають різним межах вихідної функції, зміна рівнів компенсації може відбуватися незалежно одна від одної. Зменшення рівня компенсації для однієї

межі, «послабляє» вимоги до іншої. Для стратегії МНД такі зміни пов'язані співвідношенням

$$\Delta \underline{\lambda} = \Delta \bar{\lambda},$$

що забезпечує максимальний об'єм допусків. Відповідно: $\Delta \underline{\lambda}_{\text{HM}}^{\text{LM}} = \Delta \bar{\lambda}_{\text{LM}}^{\text{HM}}$, $\Delta \underline{\lambda}_{\text{HP}}^{\text{LP}} = \Delta \bar{\lambda}_{\text{LP}}^{\text{HP}}$.

Розглянемо задачу знаходження $\Delta \underline{\lambda}_{\text{HM}}^{\text{LM}}$. Задача вирішується за методикою подібною до визначення $\Delta \underline{\lambda}_{\text{LP}}^{\text{LM}}$. Зміна рівня компенсації $\Delta \underline{\lambda}_{\text{HM}}^{\text{LM}}$ має бути такою, що

$$\begin{aligned} \sigma_x^{\text{HM}} &= s_y^{\text{HM}} - \Delta \underline{\lambda}_{\text{HM}}^{\text{LM}} \cdot \bar{\delta}_y / \vartheta_m; & x_{\text{LM}}^{\text{HM}} &= x_{\text{LM}}^{\text{HM}*}; \\ \sigma_x^{\text{HM}*} &= s_y^{\text{LM}} + \Delta \underline{\lambda}_{\text{HM}}^{\text{LM}} \cdot \bar{\delta}_y / \vartheta_m. \end{aligned}$$

Сформуємо множину:

$$\mathbf{DS}_{\mathbf{x}} = \{ds_{xi} | 1/\gamma_i > 1/\gamma_{i+1}\}, \quad ds_{xi} = s_{xi}^{\text{HM}} - s_{xi}^{\text{HM}*}.$$

Зі зменшенням Δs_y^{HM} на величину ds_{xi} відстань між інтервалами прийнятних значень і проекції змінюється як

$$\Delta d_i = ds_{xi} / \gamma_i + ds_{xi} \bar{\delta}_y / \bar{\delta}_y.$$

Тоді задачу пошуку зміни рівня компенсації можна записати як задачу знаходження t такого, що

$$\sum_{i=1}^t \Delta d_i \geq s_y^{\text{HM}*} - s_y^{\text{LM}} \geq \sum_{i=1}^{t+1} \Delta d_i.$$

Зміна фактора компенсації для області НМ і зміна рівня компенсації будуть дорівнювати

$$\begin{aligned} \Delta s_y^{\text{HM}} &= \sum_{i=1}^{t+1} ds_{xi} - \frac{\sum_{i=1}^{t+1} \Delta d_i - (s_y^{\text{HM}*} - s_y^{\text{LM}})}{1/\gamma_{t+1} + \bar{\delta}_y / \bar{\delta}_y}, \\ \Delta \underline{\lambda}_{\text{HM}}^{\text{LM}} &= \frac{\Delta s_y^{\text{HM}} \vartheta_m}{\bar{\delta}_y}. \end{aligned}$$

Визначення зміни рівню компенсації $\Delta \underline{\lambda}_{\text{HP}}^{\text{LP}}$ відбувається за подібною процедурою.

Весь процес призначення параметрів елементів може бути представлений у вигляді наступного ітераційного алгоритму:

1. Ініціалізація: $\underline{\lambda} = \bar{\lambda} = 1$;
2. Призначення допусків вхідних параметрів для заданих рівнів $\underline{\lambda}$, $\bar{\lambda}$;
3. Знаходження інтервалу рішення \mathbf{R} для факторів компенсації;

4. Перевірка інтервалу рішень: якщо рішення існує ($\text{wid}(\mathbf{R}) \geq 0$), то перехід на крок 8;

5. Рішення не існує ($\text{wid}(\mathbf{R}) < 0$): визначення змін рівнів компенсації $\Delta \underline{\lambda}$, $\Delta \bar{\lambda}$, $\underline{\lambda} = \underline{\lambda} - \Delta \underline{\lambda}$, $\bar{\lambda} = \bar{\lambda} - \Delta \bar{\lambda}$;

6. Перевірка рівнів компенсації: якщо $\underline{\lambda} < \underline{\lambda}_{\text{min}}$ або $\bar{\lambda} < \bar{\lambda}_{\text{min}}$, де $\underline{\lambda}_{\text{min}}$, $\bar{\lambda}_{\text{min}}$ – мінімально допустимі рівні компенсації, – то перехід на крок 10;

7. Перехід на крок 2;

8. Призначення факторів компенсації згідно критерію максимального об'єму і визначення КЗВ елементів;

9. Кінець;

10. Повідомлення про неможливість вирішення задачі при заданих $\underline{\lambda}_{\text{min}}$, $\bar{\lambda}_{\text{min}}$ і перехід на крок 9.

3 СТРАТЕГІЯ МКЗВ

Метою стратегії МКЗВ є призначення таких параметрів елементів, що характеризуються максимальним об'ємом КЗВ:

$$V_{\alpha} = \prod_{i=1}^n \frac{\alpha_{xi}}{\alpha_{xi \text{min}}}; \quad V_{\alpha} \rightarrow \text{MAX},$$

де V_{α} – нормований об'єм КЗВ.

Змінними в системі (3) у цьому випадку є значення меж вхідних параметрів. Їх призначення відбувається таким чином, щоб задовольнити призначенням КЗВ, а об'єм допусків вхідних параметрів був максимальним.

Призначення допусків для нижньої і верхньої меж вихідної функції є процедурами незалежними одна від одної. Внаслідок цього система обмежень (3) може бути представлена у вигляді двох систем з незалежними наборами змінних. Вводячи підстановки

$$\begin{aligned} \underline{\beta}_{mi} &= a_i(1 + \alpha_{mi} \vartheta_m), & \underline{\beta}_{pi} &= a_i(1 + \alpha_{mi} k_{ai} \vartheta_p); \\ \underline{x}_i &= x_{ri} - \bar{\delta}_{xi}, & \bar{x}_i &= x_{ri} - \bar{\delta}_{xi}, \end{aligned}$$

система (3) приймає вигляд

$$\begin{cases} \sum_{i=1}^n \underline{\beta}_{mi} \bar{\delta}_{xi} \leq a_0 + \sum_{i=1}^n \underline{\beta}_{mi} x_{ri} - y_e; \\ \sum_{i=1}^n \underline{\beta}_{pi} \bar{\delta}_{xi} \leq a_0 + \sum_{i=1}^n \underline{\beta}_{pi} x_{ri} - y_e; \\ \sum_{i=1}^n \bar{\beta}_{mi} \bar{\delta}_{xi} \leq \bar{y}_e - \bar{a}_0 - \sum_{i=1}^n \bar{\beta}_{mi} x_{ri}; \\ \sum_{i=1}^n \bar{\beta}_{pi} \bar{\delta}_{xi} \leq \bar{y}_e - \bar{a}_0 - \sum_{i=1}^n \bar{\beta}_{pi} x_{ri}. \end{cases} \quad (6)$$

Рішення, що характеризуються максимальним об'ємом, лежить на межі області прийнятних значень, що утворюється гіперплощинами обмежень системи. Таке рішення може бути сформовано одним або обома обмеженнями. Для визначення обмежень, що формують оптимальне рішення системи, необхідно визначити оптимальні рішення для обох обмежень. Якщо рішення для одного з обмежень вдовольняє іншому, то воно є і оптимальним рішенням системи. Визначення оптимальних рішень для кожного з обмежень відбувається за методикою [6]. Низьке значення λ , що характеризує рішення або відсутність розв'язку для одного з обмежень, вказує на необхідність зменшення об'єму КЗВ. Якщо рішення для першого обмеження є неприйнятним, то необхідно зменшити додатні КЗВ. Для другого обмеження – збільшити від'ємні КЗВ.

В іншому випадку оптимальне рішення системи лежить на перетині цих обмежень. Це дозволяє замінити нерівності системи на рівняння зі збереженням лівої та правої частин. Для вирішення задачі максимізації об'єму допусків із обмеженнями нового типу, може бути застосована наступна методика.

Підстановка оптимального рішення першого рівняння системи (6) до другого визначає різницю між існуючим і межовим значеннями для другого рівняння:

$$d = \beta_{pi} \delta_{xi}^M - \left(a_0 + \sum_{i=1}^n \beta_{pi} x_{ri} - y_e \right),$$

де d – відстань між існуючим і необхідним значеннями лівої частини другого обмеження.

Змінюючи значення рішення першого рівняння, можна скоротити таку відстань до нуля. Оптимальними змінами у цьому випадку будуть такі, що характеризуються $\partial d / \partial V_\delta$. При «ребалансуванні» рішення доцільно зменшувати параметри, для яких

$$h_i \geq 1, \quad h_i = \beta_{pi} / \beta_{mi},$$

і збільшувати ті, для яких $h_i < 1$. Зміна повинна бути пропорційна значенням h_i .

Для цього поділимо нормовані внутрішні параметри на дві множини:

$$\begin{aligned} \mathbf{U} &= \{u_i | \beta_{mi} \delta_{xi}\}; \\ \mathbf{U}_G &= \{u_{gi} | h_i \geq 1, i = \overline{1, n_G}\}, \\ \mathbf{U}_L &= \{u_{lj} | h_j \geq 1, j = \overline{1, n_L}\}, \end{aligned} \quad (7)$$

де \mathbf{U} , \mathbf{U}_G , \mathbf{U}_L – множини нормованих параметрів; n_G , n_L – кількість елементів що зменшуються та збільшуються, відповідно.

Зміни нормованих параметрів (7) мають бути такі, що

$$\begin{aligned} \sum_{i=1}^{n_G} \Delta u_{gi} h_i + \sum_{j=1}^{n_L} \Delta u_{lj} h_j &= -d; \\ \sum_{i=1}^{n_G} \Delta u_{gi} + \sum_{j=1}^{n_L} \Delta u_{lj} &= 0, \end{aligned} \quad (8)$$

де Δu_{gi} , Δu_{lj} – зміни нормованих параметрів.

Зменшення параметра має бути пропорційно $h_i / \sum_{i=1}^{n_G} h_i$, збільшення – $h_j / \sum_{j=1}^{n_L} 1/h_j$. Тоді (8) може бути записано як

$$du \cdot \frac{\sum_{i=1}^{n_G} h_i^2}{\sum_{i=1}^{n_G} h_i} - du \cdot \frac{n_L}{\sum_{j=1}^{n_L} 1/h_j} = -d,$$

де du – сумарна зміна нормованих параметрів, що зменшуються:

$$du = \sum_{i=1}^{n_G} \Delta u_{gi}.$$

Відповідно

$$du = d / \left(\frac{n_L}{\sum_{j=1}^{n_L} 1/h_j} - \frac{\sum_{i=1}^{n_G} h_i^2}{\sum_{i=1}^{n_G} h_i} \right).$$

Значення змін нормованих параметрів знаходять як

$$\Delta u_{gi} = du \cdot h_i / \sum_{i=1}^{n_G} h_i, \quad \Delta u_{lj} = - \frac{du/h_j}{\sum_{i=1}^{n_L} 1/h_j}.$$

Застосування стратегії МКЗВ для призначення параметрів елементів відбувається за алгоритмом:

1. Ініціалізація: $\alpha_{mi} = \alpha_{mi \min}$ if $\alpha_{mi} < 0$, $\alpha_{mi} = \alpha_{mi \max}$ if $\alpha_{mi} > 0$; $\underline{\lambda} = \bar{\lambda} = 1$;
2. Визначення допусків внутрішніх параметрів;
3. Якщо об'єм допусків $\underline{\lambda} \geq \underline{\lambda}_{\min}$ і $\bar{\lambda} \geq \bar{\lambda}_{\min}$, де $\underline{\lambda}_{\min}$, $\bar{\lambda}_{\min}$ – мінімально допустимі рівні компенсації, то перехід на крок 7;
4. Зміна значень КЗВ внутрішніх параметрів;

5. Якщо об'єм КЗВ $V_\alpha < V_{\alpha\min}$, де $V_{\alpha\min}$ – мінімально задовільний об'єм КЗВ, то рішення для встановлених умов не існує, перехід на крок 7;

6. Перехід на крок 2;

7. Кінець.

4 ПРИКЛАД ЗАСТОСУВАННЯ

Запропоновані методики було застосовано при призначенні параметрів елементів для ФНЧ (рис. 3). За вихідний параметр було обрано коефіцієнт загасання фільтра a (9). Вихідні дані для обчислення наведені у таблиці 1. Обчислення проводилися на частоті 12 кГц. Номінальне значення коефіцієнта загасання при цьому дорівнює $a = 6,16972$.

$$a(\omega) = \frac{R_2}{R_0} \times \sqrt{(1 - \omega^2 \cdot C_1 C_2 R_1 R_0)^2 + \left(\omega \cdot C_1 R_1 \left(1 + \frac{R_0}{R_1} + \frac{R_0}{R_2} \right) \right)^2}, \quad (9)$$

де ω – кутова частота; R_0, R_1, R_2, C_1, C_2 – параметри елементів.

Параметри моделі і інтервали вхідних параметрів для початкових значень рівня компенсації ($\underline{\lambda} = \bar{\lambda} = 1$) наведені у табл. 2.

Стратегія МНД. Інтервал рішень на області LM дорівнює $\mathbf{R} = [3, 7065; -3, 60723] \cdot 10^{-4}$. Ширина інтервалу рішень від'ємна – рішення при початковому рівні

Таблиця 1 – Вихідні дані

Параметр	Номінал	Межі α_{mi} , $10^6 \cdot 1/^\circ\text{C}$	$k_i = \alpha_{pi}/\alpha_{mi}$
R0	100,0 кОм	[55;1200]	1,3
R1	91,0 кОм	[55;1200]	1,1
R2	100,0 кОм	[55;1200]	1,1
C1	1,0 нФ	[-1750; -120]	1,4
C2	4,7 нФ	[-1750; -120]	1,5

Таблиця 3 – Результати обчислень за стратегією МНД

Параметр	α_m , $10^6 \cdot 1/^\circ\text{C}$	α_p , $10^6 \cdot 1/^\circ\text{C}$	$[\underline{\delta}_{xi}; \bar{\delta}_{xi}]$	$\delta_{xi}, \%$	$\bar{\delta}_{xi}, \%$
R0	1200	1560	[20,000;20,000] кОм	2,66 %	2,19 %
R1	170	187	[1,846;1,680] кОм	20,00 %	20,00 %
R2	55	61	[2,387;2,055] кОм	2,03 %	1,85 %
C1	-120	-168	[0,018;0,017] нФ	2,39 %	2,06 %
C2	-120	-180	[0,125;0,103] нФ	1,81 %	1,69 %

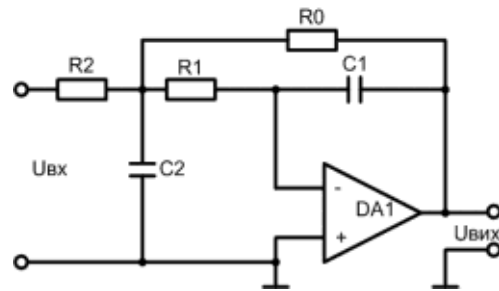


Рисунок 3 – Схема електрична принципова ФНЧ

компенсації не існує. Зміни рівня компенсації дорівнюють $\Delta\lambda_{LP}^{LM} = -1,53 \%$, $\Delta\lambda_{HP}^{HM} = 0 \%$, $\Delta\lambda_{HM}^{LM} = 0 \%$, $\Delta\lambda_{LP}^{HM} = 0 \%$. Прийнятні фактори компенсації знаходяться вже на другій ітерації. Значення КЗВ, що відповідають цим факторам, а також номінальні допуски елементів наведені у табл. 3.

Значення рівня компенсації при цьому становить $= 0,9847;1$, а допуск вихідного параметра $\mathbf{y}_r = [9,85;10,00] \%$. Нормований об'єм КЗВ складає $V_\alpha = 0,32$.

Стратегія МКЗВ. Вихідні дані для стратегій МКЗВ і МНД співпадають (табл. 1). Встановлено мінімальне значення рівня компенсації $= 0,25;0,25$. Результати стратегії наведено у табл. 4. Значення рівня компенсації при цьому становить $\lambda = 0,25;0,49$, а допуск вихідного параметра $\mathbf{y}_r = [2,50;4,93] \%$. Нормований об'єм КЗВ складає $V_\alpha = 7774,79$.

Таблиця 2 – Параметри моделі

Параметр	$[a_i; \bar{a}_i]$	$[x_i; \bar{x}_i]$
R0	$[2,17611; 2,33695] \cdot 10^{-6}$	[80,000;120,000] кОм
R1	$[6,57435; 7,34429] \cdot 10^{-5}$	[89,123;92,680] кОм
R2	$[5,08404; 6,00484] \cdot 10^{-5}$	[97,573;102,055] кОм
C1	$[6,69131; 7,30723] \cdot 10^9$	[0,982;1,017] нФ
C2	$[0,970635; 1,19676] \cdot 10^9$	[4,573;4,803] нФ

Таблиця 4 – Результати обчислень за стратегією МКЗВ

Параметр	α_m $10^6 \cdot 1/^\circ\text{C}$	α_p $10^6 \cdot 1/^\circ\text{C}$	$[\underline{\delta}_{xi}; \bar{\delta}_{xi}]$	$\delta_{xi}, \%$	$\bar{\delta}_{xi}, \%$
R0	1200	1560	[13,254;27,289] кОм	13,25 %	27,29 %
R1	1200	1320	[0,443;0,868] кОм	0,49 %	0,95 %
R2	1200	1320	[0,573;1,062] кОм	0,57 %	1,06 %
C1	-1514	-2120	[5,004;7,423] пФ	0,50 %	0,74 %
C2	-1514	-2271	[0,035;0,045] нФ	0,74 %	0,96 %

ВИСНОВКИ

Запропоновано метод призначення параметрів елементів. Враховано дію зовнішніх чинників за допомогою коефіцієнтів зовнішніх впливів. Коефіцієнти представлені керованими інтервалами, що дозволяє врахувати зміну фізичних властивостей елементної бази під час експлуатації і збільшити область адекватності методів призначення допусків. Розроблено стратегії максимальних номінальних допусків і максимальних коефіцієнтів зовнішніх впливів, що дозволяє знизити вимоги до точності виготовлення або стабільності елементної бази відповідно.

ПЕРЕЛІК ПОСИЛАНЬ

1. Михайлов А. В., Савин К. С. Точность радиоэлектронных устройств. – М.: Машиностроение, 1976. – 214 с.
2. Фридендер И. Г. Расчеты точности машин при проектировании. – Киев–Донецк: Вища школа, 1980. – 184 с.
3. Гехер К. Теория чувствительности и допусков электрических цепей. – М.: Сов. радио, 1973. – 199 с.

4. Шило Г. М. Формування інтервальних моделей для обчислення допусків // Радіоелектроніка. Інформатика. Управління. – 2002. – № 1. – С. 90–95.
5. Кришук В. М., Шило Г. М., Намлинський А. О., Гапоненко М. П. Вибір елементів при компенсації зовнішніх впливів // Радіоелектроніка. Інформатика. Управління. – 2004. – № 2 – С. 74–78.
6. Шило Г. М., Воропай О. Ю., Гапоненко М. П. Інтервальні методи призначення експлуатаційних допусків / Радіоелектроніка. Інформатика. Управління. – 2003. – № 2. – С. 78–82.

Надійшла 18.08.06
Після доробки 22.09.06

Предложен метод назначения допусков на параметры элементов. Учтено влияние внешних факторов на параметры элементов. Коэффициенты внешних воздействий представлены интервалами. Приведены реализации стратегий реализующих максимального объема допусков и максимального объема коэффициентов внешних воздействий.

Method of tolerance assignment is offered. Influence of external actions is taken into account. External action coefficients represented as intervals. Realization of strategies of maximal tolerance box and maximal external action coefficients are consider.

УДК 621.38

А. А. Никонова, Д. И. Левинзон, О. Ю. Небеснюк

ТЕХНОЛОГИЧЕСКИЕ ОСОБЕННОСТИ ИЗГОТОВЛЕНИЯ КОНТАКТНЫХ СИСТЕМ К ПОЛУПРОВОДНИКОВЫМ СТРУКТУРАМ

Разработана многослойная система металлизации, состоящая из слоя силицида, барьерного слоя и низкоомного верхнего контактного слоя для устранения взаимодействия с нижележащим слоем кремния в контактных «окнах».

ВВЕДЕНИЕ

При создании полупроводниковых приборов широко используется нанесение тонких металлических пленок,

как для формирования межсоединений, так и для изготовления омических контактов и выпрямляющих переходов металл-полупроводник.

Наибольший интерес представляют собой переходы металл-полупроводник типа переходов Шотки, которые могут обладать как омическими, так и выпрямляющими свойствами. Основные свойства контактов определяются барьером, возникающим на границе раздела этих материалов. Высота барьера, в случае иде-

© Никонова А. А., Левинзон Д. И., Небеснюк О. Ю., 2007

альной границы раздела, определяется разницей между работой выхода из металла и электронным сродством полупроводника. Таким образом, зная эти величины, характерные для каждого материала, можно заранее предсказать высоту барьера и параметры перехода. Переход металл-полупроводник начинает проявлять свойства омического контакта в случае увеличения концентрации примесей в полупроводнике, так как при этом обедненная область в полупроводнике становится настолько узкой, что возможно туннелирование электронов из металла.

Учитывая все вышесказанное, можно заранее предсказать будет ли металл на данном полупроводнике создавать омический или выпрямляющий контакт. Однако эти рассуждения не всегда справедливы. В частности, они могут быть непосредственно применимы к таким широко распространенным полупроводникам, как кремний, карбид кремния, германий или арсенид галлия. Это связано с тем, что на поверхности указанных полупроводников имеется высокая плотность поверхностных состояний, то есть создается дополнительный поверхностный потенциал, который может играть более существенную роль при образовании барьера на границе металл-полупроводник, чем уровень Ферми.

Необходимо учитывать, что в процессе создания контакта металл-полупроводник возможно взаимодействие этих двух материалов с увеличением или уменьшением концентрации доноров или акцепторов в полупроводнике. Такие же изменения могут происходить и в результате последующих термообработок. Существенное влияние на параметры перехода металл-полупроводник оказывает окисный слой, сформированный на границе этих двух материалов. Большинство полупроводников всегда имеет на поверхности «собственный» окисел. Например, кремний всегда покрыт пленкой окисла толщиной не менее 25 \AA , который вырастает на его поверхности на воздухе даже при комнатной температуре. Такой промежуточный окисел во-первых, уменьшает плотность поверхностных состояний, во-вторых, при увеличении толщины окисла возможен переход к МОП — структуре, в-третьих, дополнительный окисный слой снижает туннельные токи, как в прямом, так и в обратном направлении. И, наконец, если в случае идеального контакта Шоттки ток состоит из основных носителей, то при наличии промежуточного окисла возможна инжекция неосновных носителей, что соответственно изменяет общий механизм прохождения тока через переход. Наличие промежуточного окисла увеличивает сопротивление омического контакта. Омические контакты и барьеры Шоттки можно создавать в виде металлических пленок известными способами: термическое распыление, электронно-лучевое напыление, катодное распыление,

электрохимическое осаждение, метод получения пленок из газовой фазы.

ПОСТАНОВКА ПРОБЛЕМЫ

С уменьшением размеров элементов полупроводниковых приборов становятся все более жесткими требования к ширине тонкопленочных проводников, размерам контактных окон и свойствам самих контактов. Наиболее распространенным материалом, применяемым при изготовлении контактной металлизации и межсоединений в кремниевых ИС является алюминий, вследствие его низкого удельного сопротивления, возможности формирования низкоомных контактов к кремнию *p*- и *n*-типа, а также технологичности при последующих процессах сборки и герметизации. Такой серьезный недостаток однослойной металлизации, как взаимодействие с нижележащим слоем кремния в контактных «окнах», вызвал необходимость разработки многослойных систем металлизации. Подобная система обычно состоит из слоя силицида, барьерного слоя и низкоомного верхнего контактного слоя. Силицидный слой обеспечивает формирование надежного контакта к кремнию, а барьерный слой служит для предотвращения диффузии через силицид в кремний материала верхнего контактного слоя (обычно алюминия или золота). Известно, что в качестве материала барьерного слоя использовался сплав Ti–W, содержание титана в котором составило 10 % по массе [1]. Указанный сплав проявляет хорошие барьерные свойства по отношению к золоту, однако, он недостаточно устойчив при диффузии алюминия, особенно, если формирование пленок сплава протекает в хорошем вакууме при отсутствии остаточных газов.

Улучшение барьерных свойств по отношению к алюминию было достигнуто при использовании Ti–W–N пленок, полученных распылением в смеси N_2 –Ar. Присутствие других газовых примесей (например, кислорода) приводит к повышению удельного сопротивления барьерной пленки и к ухудшению ее адгезии к SiO и Au. В качестве барьерного слоя можно применять пленки нитрида какого-либо переходного металла [2].

Введение барьерного слоя нитрида титана между Ti и Pt в системе металлизации Ti–Pt–Au позволяет уменьшить взаимную диффузию Ti и Pt и повысить надежность системы в целом. Применение барьерных слоев TiN и TaN в соответствующих контактных системах (КС) Pt–TiN–Ni и Cr–TaN–Ni повышает стабильность КС при термообработках вплоть до температуры $600 \text{ }^\circ\text{C}$.

Однако при таких температурах нитридные барьерные слои оказываются неустойчивыми по отношению к алюминию [3]. Диффузия через нитридные барьерные слои в этом случае, по-видимому, происходит по микрокавернам в нитридной пленке, которые появля-

ются вследствие разложения нитрида и образования AlN. Отмечается, что отклонение нитридов от стехиометрии, а также микротрещины в КС, вызванные внутренними напряжениями, ухудшают барьерные свойства нитридных слоев.

ЦЕЛИ И ЗАДАЧИ ИССЛЕДОВАНИЯ

С целью предотвращения взаимодействия с нитридным барьерным слоем и, соответственно, с нижележащим слоем кремния, в качестве верхнего слоя авторами использовался не чистый алюминий, а его интерметаллическое соединение. В качестве материалов, образующих соединение с алюминием были выбраны Ti и Ta. Исследовались системы Pt–TiN–Ti–Al и Ni–TaN–Ta–Al. В качестве подложек в настоящей работе использовались кремниевые пластины с ориентацией (100) и удельным сопротивлением 5 Ом·см. На такие пластины в едином технологическом цикле распылением в триодной системе последовательно наносились слои Pt, TiN, Ti и Al или Ni, TaN, Ta и Al. Распыление производилось в смеси N₂–Ar.

Толщина слоев Al, Ti, Ta выбиралась такой, что в результате последующей термообработки формировались стабильные соединения Al₃Ti и Al₃Ta. Отжиг образцов производился в специальной вакуумной печи, откачиваемой ионным насосом до остаточного давления 6,6·10⁻⁶ тор.

Исследование указанных ИС до и после отжига проводилось на основе анализа спектров резерфордского обратного рассеяния ионов гелия. Такой анализ показал, что в результате отжига систем при температуре 600 °С в течение 10 мин. последовательно нанесенных на кремниевую подложку слоев Pt, TiN, Ti и Al платина реагирует с кремнием, а титан с алюминием. Диффузия алюминия через барьерный слой при этом не происходит. Дальнейший отжиг этих образцов при температуре 600 °С в течение 20 часов не приводит к дополнительной взаимной диффузии между слоями в КС.

Известно, что взаимная диффузия Ti и Al протекает уже при температуре 350 °С [4]. Если толщина пленки титана меньше, значения, рассчитанного из стехиометрии Al₃T, то избыток алюминия реагирует с нижележащими слоями. Это приводит к деградации КС, что не имеет места при избытке титана. Однако избыток титана также нежелателен, поскольку в этом случае наблюдается повышенное значение удельного сопротивления соединения.

Заметная взаимная диффузия Ta и Al наблюдалась при температурах выше 450 °С. Анализ спектров резерфордского обратного рассеяния ионов гелия для образцов многослойной системы Si–Ni–TaN–Ta–Al,

отожженных при температуре 600 °С, показал, что десятиминутный отжиг приводит лишь к образованию силицида никеля. Полное взаимодействие Ta и Al происходит после длительной (в течение 2,5 часов) термообработки. Последующий отжиг этих образцов в течение 20 часов не приводит к заметным изменениям в КС.

ВЫВОДЫ

Проведенные исследования показали, что контактная система металлизации к кремниевым полупроводниковым приборам, в которой в качестве диффузионного барьера используются слои нитрида переходного металла и интерметаллические соединения алюминия, проявляет высокую стабильность при воздействии высокотемпературных обработок. Важным технологическим преимуществом такой системы является возможность нанесения всех слоев в одном процессе распыления без разгерметизации. Наличие барьерного слоя необходимо также для предотвращения взаимной диффузии Pt и Ti, которая типична и для других бинарных систем металл-благородный металл.

Единственным недостатком интерметаллических соединений алюминия является повышенное (по сравнению с чистым алюминием) значение поверхностного сопротивления. Однако в случае Al₃Ti наблюдается увеличение поверхностного сопротивления лишь в два раза, что не оказывает существенного воздействия на характеристики контактов.

ПЕРЕЧЕНЬ ССЫЛОК

1. Bickermann M. Characterization of SiC crystals // Journal of Crystal Growth. – 2004. – № 269. – P. 432–442.
2. Енишерлова К. Л., Концевой Ю. А. Проблемы кремния в силовой электронике: монокристаллы, пластины, структуры // Известия ВУЗов. Материалы электронной техники. – 2000. – № 4. – С. 4–9.
3. Шульпина И. Л. Диагностика кремния методом рентгеновской топографии // Известия ВУЗов. Материалы электронной техники. – 2001. – № 4. – С. 73–76.
4. Шуе К.-К., Шуе Ч.-Ж., Хасегава Ю., Сакурай Т. Рост высококачественных пленок на поверхности карбида кремния. – М.: МИСиС, 2001. – 257 с.

Надійшла 05.10.06

Після доробки 10.01.07

Розроблено багатощарову систему металізації, яка складається з шару силіциду, бар'єрного шару та низькоомного верхнього контактної шару для усунення взаємодії з нижче лежачим шаром кремнію в контактних «вікнах».

Multilayer metallised system was developed. It consists of a silicic layer, a barrier layer and an upper contact part for the purpose of removing of interaction with a lower silicon part in contact holes.

Ю. К. Тараненко

ИСПОЛЬЗОВАНИЕ ПАРАМЕТРИЧЕСКОГО КОНТУРА РЕГУЛИРОВАНИЯ АМПЛИТУДЫ В АВТОКОЛЕБАТЕЛЬНЫХ СИСТЕМАХ ВИБРОЧАСТОТНЫХ ПЛОТНОМЕРОВ С МЕХАНИЧЕСКИМ РЕЗОНАТОРОМ

В статье рассмотрены вопросы повышения точности измерений путем использования параметрического контура регулировки амплитуды в автоколебательных системах виброчастотных измерителей плотности жидкости и газа с механическими резонаторами. Приведены структурные схемы систем возбуждения автоколебаний механического резонатора с регулировкой амплитуды колебаний и амплитуды тока в цепи возбудителя колебаний. Рассмотрены вопросы устойчивости автоколебаний в системах с параметрическими контурами.

1 АНАЛИЗ ТЕХНИЧЕСКОГО УРОВНЯ АНАЛИТИЧЕСКОЙ ИЗМЕРИТЕЛЬНОЙ ТЕХНИКИ, ОСНОВАННОЙ НА ВИБРОЧАСТОТНОМ АВТОГЕНЕРАТОРНОМ МЕТОДЕ ИЗМЕРЕНИЯ ПЛОТНОСТИ ЖИДКИХ И ГАЗООБРАЗНЫХ СРЕД

В условиях жесткой экономии энергоресурсов особое значение приобретает задача точного измерения плотности жидкостей и газов непосредственно в нефтегазопроводах и технологических аппаратах, в условиях резких изменений температуры, давления и скорости течения измеряемой среды. Известно, что все вещества реагируют между собой в заданных соотношениях – массовых долях. Массовый расход жидких и газообразных сред определяется как результат умножения объемного расхода на плотность, поэтому плотность веществ является одним из основных технологических параметров, который определяет качество исходного сырья и готовой продукции, а также затраты основных энергоносителей на производство этой продукции.

Известно, что повышение точности измерения массового расхода природного газа на границе Украины с Россией всего на 0,1 % позволяет Украине экономить более 10 млн. долл. в год [1]. Виброчастотные методы измерения плотности жидкости и газов, основанные на преобразовании измеряемой величины в частоту колебаний механической колебательной системы с распре-

деленными параметрами, способны обеспечить самую высокую точность измерения [2]. Применение компьютерных измерительных технологий для создания интеллектуальных частотно-цифровых измерительных систем, открывает дополнительные возможности по обработке выходных частотных сигналов виброчастотных измерительных преобразователей [3].

Наиболее распространенным резонатором виброчастотных измерительных преобразователей является камертон, на основе которого создано семейство вибрационных датчиков предельного уровня и плотности жидкости Liquiphant [4, 5] фирмы Endress+Hauser GmbH+Co KG, и, позднее, таких фирм, как Vega Grieshaber KG, Solartron Mobrey Ltd., Nivelco Process Control Co. Только в странах западной Европы серийный выпуск виброчастотных измерительных преобразователей достигает 260 тыс. ед. в год.

Для возбуждения колебаний камертона, в виброчастотных измерителях плотности жидкости используется автогенератор [6]. Автогенератор состоит из резонатора, систем возбуждения и устройства съема колебаний резонатора, включенных в цепь усилителя. Системы возбуждения колебаний резонатора могут быть электромагнитными, магнитоэлектрическими, пьезоэлектрическими. Для приема колебаний используются электромагнитные, емкостные, пьезоэлектрические системы. Основной погрешностью виброчастотных измерительных преобразователей является погрешность, связанная с отклонением частоты автоколебаний от собственной частоты колебаний механического резонатора. Основная погрешность зависит от многих факторов, среди которых: отклонения от условий возбуждения автоколебаний по амплитуде и фазе, изменение амплитуды автоколебаний вследствие изменения измеряемой величины. Таким образом, автогенератор является одним из основных источников погрешности виброчастотных измерительных преобразователей, что вызывает необходимость его тщательного исследования.

дующие обозначения: p – оператор дифференцирования, который для гармонических колебаний равняется $p = j \cdot \omega$, $j = \sqrt{-1}$; $K_{\text{СФ}}$, $T_{\text{СФ}}$ – статический коэффициент передачи и постоянная времени сглаживающего фильтра; $K_{\text{П}}$ – статический коэффициент передачи приемника колебаний резонатора; $T_{1\text{ТР}}$, $T_{2\text{ТР}}$ – постоянные времени передаточной функции трансформатора; $K_{\text{ТР}}$, $\bar{K}_{\text{ТР}}$ – статический и приведенный для условия $T_{2\text{ТР}} \ll T_{1\text{ТР}}$ коэффициенты передачи трансформатора; $K_{\text{В}}$, $T_{\text{В}}$ – статический коэффициент передачи и постоянная времени возбудителя колебаний резонатора.

В выражении для передаточной функции резонатора приняты следующие обозначения: ξ – нормированный коэффициент затухания колебаний резонатора; $T_{\text{р}}$ – период колебаний резонатора на собственной частоте; $K_{\text{р}}$ – статический коэффициент передачи резонатора.

Методом приведения распределенных механических параметров к сосредоточенным [8] получаем следующие соотношения для параметров резонатора:

$$\xi = r_{\text{Э}} / (2 \cdot m_{\text{Э}} \cdot \omega_{\text{р}}); \quad T_{\text{р}} = 1 / \omega_{\text{р}};$$

$$K_{\text{р}} = 1 / (m_{\text{Э}} \cdot \omega_{\text{р}}^2) = 1 / c_{\text{Э}},$$

где $\omega_{\text{р}}^2 = c_{\text{Э}} / (m_{\text{Э}}) = c_{\text{Э}} / (m_{\text{ЭР}} + m_{\text{ЭЖГ}})$ – квадрат круговой частоты собственных колебаний резонатора; $c_{\text{Э}}$ – эквивалентная приведенная жесткость резонатора; $m_{\text{Э}}$ – эквивалентная приведенная масса резонатора с контролируемой жидкостью или газом; $m_{\text{ЭР}}$ – эквивалентная приведенная масса резонатора; $m_{\text{ЭЖГ}}$ – эквивалентная приведенная масса контролируемой жидкости или газа в резонаторе или присоединенная к резонатору масса (для погружного плотномера); $r_{\text{Э}}$ – эквивалентный коэффициент затухания колебаний резонатора.

Следует отметить, что введение параметрического контура позволяет получить в основном контуре практически синусоидальные колебания, что обеспечивает работу основного контура на линейном участке. При этом уменьшается нестабильность частоты выходного сигнала, так как изменение степени насыщения усилителя, при изменении амплитуды, приводит к фазовым сдвигам.

Для структурной схемы возбуждения автоколебаний (рис. 1) уравнения, описывающие основной и параметрический контур регулировки амплитуды, с учетом функции преобразования выпрямителя, будут иметь вид:

$$\left. \begin{aligned} X - X \cdot Y \cdot K_{\text{ОК}} \cdot W_{\text{ОК}}(p) &= 0 \\ Y &= E_0 - K_{\text{СФ}} \cdot W_{\text{СФ}}(p) \cdot x_{\text{В}} \end{aligned} \right\} \quad (1)$$

где $K_{\text{ОК}} = K_{\text{р}} \cdot K_{\text{П}} \cdot K_{\text{УС1}} \cdot K_{\text{ТР}} \cdot K_{\text{В}}$ – общий коэффициент передачи основного контура; $W_{\text{ОК}}(p) = W_{\text{р}}(p) \times W_{\text{П}}(p) \cdot W_{\text{ТР}}(p) \cdot W_{\text{В}}(p)$ – передаточная функция основного контура; $x_{\text{В}} = \begin{cases} K_{\text{ВП}} \cdot K_{\text{УС2}} \cdot X; & \text{при } X \geq 0 \\ 0; & \text{при } X < 0 \end{cases}$; $K_{\text{ВП}}$ – статический коэффициент передачи выпрямителя.

Для структурной схемы возбуждения автоколебаний (рис. 2) уравнения, описывающие основной и параметрический контур регулировки амплитуды возбуждающего тока, с учетом функции преобразования выпрямителя, будут иметь вид:

$$\left. \begin{aligned} X - X \cdot Y \cdot K_{\text{ОК}} \cdot W_{\text{ОК}}(p) &= 0 \\ Y &= E_0 - K_{\text{СФ}} \cdot W_{\text{СФ}}(p) \cdot x_{\text{В}} \end{aligned} \right\} \quad (2)$$

где $x_{\text{В}} = \begin{cases} K_{\text{ВП}} \cdot Z; & \text{при } Z \geq 0 \\ 0; & \text{при } Z < 0 \end{cases}$; $Z = X \cdot \frac{K_{\text{УС1}}}{K_{\text{ОК}} \cdot W_{\text{ОК}}(p)}$.

На основе анализа приведенных систем уравнений (1), (2) можно сделать вывод, что исследуемые автоколебательные системы с параметрическим контуром регулировки амплитуды относятся к нелинейным системам третьего класса. Это подтверждается тем, что под знак нелинейности в виде множительного устройства входят две переменные X , Y , связанные между собой нелинейными дифференциальными уравнениями.

4 ПРИМЕНЕНИЕ МЕТОДА РАЗДЕЛЬНОЙ ГАРМОНИЧЕСКОЙ ЛИНЕАРИЗАЦИИ ДЛЯ АНАЛИЗА ПАРАМЕТРОВ АВТОКОЛЕБАТЕЛЬНЫХ СИСТЕМ ВИБРОЧАСТОТНЫХ ПЛОТНОМЕРОВ

Для определения параметров установившихся автоколебаний воспользуемся методом раздельной гармонической линеаризации [8], который заключается в том, что гармоническая линеаризация основного контура производится по первой гармонике колебаний, а гармоническая линеаризация нелинейных уравнений, описывающих параметрический контур автоматической регулировки амплитуды, по постоянной составляющей – нулевой гармонике. Условием применимости метода раздельной гармонической линеаризации является наличие в параметрическом контуре достаточно хорошего фильтра нижних частот (сглаживающего фильтра) и постоянной составляющей на выходе нелинейности. Для рассматриваемых схем приведенные условия соблюдаются.

Проведем гармоническую линеаризацию нелинейности основного контура по первой гармонике координаты X , рассматривая Y как некий параметр. Характеристическое уравнение основного контура относи-

тельно координаты X после линеаризации запишется в виде:

$$1 - Y \cdot K_{\text{OK}} \cdot W_{\text{OK}}(p) = 0. \quad (3)$$

Подставляя в уравнение (3) $p = j \cdot \omega$, получим следующие два уравнения для амплитуды и частоты автоколебаний:

$$Y = \frac{1}{K_{\text{OK}} \cdot |W_{\text{OK}}(p)|} = \frac{1}{\bar{K}_{\text{OK}}}, \quad (4)$$

$$\arctg[W_{\text{OK}}(p)] = 2 \cdot \pi \cdot k, \quad (5)$$

где $\bar{K}_{\text{OK}} = K_{\text{OK}} \cdot |W_{\text{OK}}(p)|$ – коэффициент усиления основного контура на частоте автоколебаний; k – целое число, принимающее значения 1, 2, 3 ... и так далее.

Из уравнений (4), (5) следует, что частота автоколебаний в первом приближении не зависит от параметров параметрического контура, а полностью определяется фазовой характеристикой основного контура. Учитывая, что резонатор является колебательным звеном и создает фазовый сдвиг $\pi/2$, частота автоколебаний будет равна собственной частоте колебаний резонатора, $\omega = \omega_p$, если цепь обратной связи будет являться дифференцирующим либо интегрирующим звеном, в зависимости от типа усилителя в основном контуре.

Линеаризация второго уравнения системы (1) по постоянной составляющей для установившегося режима автоколебаний дает следующее соотношение между параметрами основного и параметрического контуров:

$$\frac{1}{\bar{K}_{\text{OK}}} = E_0 - K_{\text{УС2}} \cdot K_{\text{ВП}} \cdot K_{\text{СФ}} \cdot \frac{1}{\pi} \cdot a_X, \quad (6)$$

где a_X – коэффициент гармонической линеаризации параметрического контура с регулировкой амплитуды колебаний резонатора.

Из уравнения (6) можно определить коэффициент гармонической линеаризации

$$a_X = \frac{\pi}{K_{\text{УС2}} \cdot K_{\text{ВП}} \cdot K_{\text{СФ}}} \cdot \left(E_0 - \frac{1}{\bar{K}_{\text{OK}}} \right).$$

Линеаризация второго уравнения системы (2) по постоянной составляющей, для установившегося режима автоколебаний, дает следующее соотношение между параметрами основного и параметрического контуров:

$$\frac{1}{\bar{K}_{\text{OK}}} = E_0 - K_{\text{ВП}} \cdot K_{\text{СФ}} \cdot \frac{1}{\pi} \cdot a_Z, \quad (7)$$

где a_Z – коэффициент гармонической линеаризации параметрического контура с регулировкой амплитуды возбуждающего тока.

Из уравнения (7) можно определить коэффициент гармонической линеаризации

$$a_Z = \frac{\pi}{K_{\text{ВП}} \cdot K_{\text{СФ}}} \cdot \left(E_0 - \frac{1}{\bar{K}_{\text{OK}}} \right).$$

На основе анализа уравнений (6), (7) можно сделать вывод, что действие параметрических контуров для обеспечения заданной амплитуды колебаний резонатора и амплитуды возбуждающего тока одинаково, а структурные схемы, приведенные на рис. 1, 2, динамически эквивалентны.

5 АНАЛИЗ УСТОЙЧИВОСТИ АВТОГЕНЕРАТОРА С ПАРАМЕТРИЧЕСКИМ КОНТУРОМ АВТОМАТИЧЕСКОЙ РЕГУЛИРОВАНИЯ АМПЛИТУДЫ В УСЛОВИЯХ НЕЗАВИСИМОСТИ КОЭФФИЦИЕНТА ПЕРЕДАЧИ ОСНОВНОГО КОНТУРА ОТ АМПЛИТУДЫ

Условия устойчивости автоколебаний резонаторов виброчастотных плотномеров имеют очень важное практическое значение. Это объясняется тем, что смена нагрузки на резонатор, например, при попадании воздушных включений в контролируруемую жидкость или пылевой твердой фракции в контролируемый газ, может привести к срыву автоколебаний и последующей потери частотного информационного сигнала. Поэтому рассмотрим вопрос устойчивости автоколебаний для схем, приведенных на рис. 1, 2, в условиях независимости коэффициента передачи основного контура от амплитуды.

При медленно меняющемся значении Y дифференциальное уравнение, описывающее движение основного контура, имеет вид:

$$X - X \cdot Y \cdot K_{\text{OK}} \cdot W_{\text{OK}}(p) = 0 \quad (8)$$

Дифференциальное уравнение в операторном виде относительно амплитуды колебаний $A_X(t)$ с учетом изменения координаты X по закону $X(t) = A_X(t) \times \sin(\omega \cdot t + \varphi)$, имеет вид:

$$A_X(t) - [A_X(t) \cdot Y] \cdot [K_{\text{OK}} \cdot W_{\text{OK}}(p)] = 0, \quad (9)$$

где $[W_{\text{OK}}(p)]$ – передаточная функция по огибающей для линейной части основного контура.

Под $[W_{\text{OK}}(p)]$ будем понимать отношение изображений по Лапласу огибающей выходного гармонического сигнала к огибающей входного гармонического сигнала, взятых при нулевых начальных условиях [9].

Передаточные функции линейной части основного контура имеют вид:

$$K_{\text{окд}} \cdot W_{\text{окд}}(p) = \frac{K_{\text{окд}} \cdot p}{T_{\text{р}}^2 \cdot p^2 + 2 \cdot \xi \cdot T_{\text{р}} \cdot p + 1} - \text{передаточная}$$

функция основного контура при дифференцирующей обратной связи резонатора;

$$K_{\text{оки}} \cdot W_{\text{оки}}(p) = \frac{K_{\text{оки}}}{p \cdot (T_{\text{р}}^2 \cdot p^2 + 2 \cdot \xi \cdot T_{\text{р}} \cdot p + 1)} - \text{пе-}$$

редаточная функция основного контура при интегрирующей обратной связи резонатора.

Для дифференцирующей и интегрирующей обратной связи колебательного звена применим следующие выражения передаточных функций по огибающей, которые принято обозначать в квадратных скобках [10]:

$$[K_{\text{окд}} \cdot W_{\text{окд}}(p)] = \frac{K_{\text{окд}}}{2 \cdot \xi \cdot T_{\text{р}}} \cdot \frac{-\tau \cdot p}{T \cdot p + 1}, \quad (10)$$

$$[K_{\text{оки}} \cdot W_{\text{оки}}(p)] = \frac{K_{\text{оки}} \cdot T_{\text{р}}}{2 \cdot \xi} \cdot \frac{-\tau \cdot p}{T \cdot p + 1}, \quad (11)$$

где $T = (\xi \cdot \omega_{\text{р}})^{-1}$; $\tau = \frac{\xi^2}{Y} \cdot T$.

Обобщая (10) и (11), передаточную функцию по огибающей для линейной части основного контура, можно записать в виде:

$$[K_{\text{ок}} \cdot W_{\text{ок}}(p)] = \bar{K}_{\text{ок}} \cdot \frac{-\tau \cdot p}{T \cdot p + 1}, \quad (12)$$

где $\bar{K}_{\text{ок}}$ – обобщенный коэффициент передачи линейной части основного контура.

Координата Y связана с огибающей уравнением:

$$Y = E_0 - K_{\text{д}} \cdot A_X(p) \cdot W_{\text{сф}}(p), \quad (13)$$

где $K_{\text{д}} = \frac{1}{\pi} \cdot K_{\text{вп}} \cdot K_{\text{сф}} \cdot K_{\text{ус2}}$.

Подставляя уравнение (13) в уравнение (9), получим нелинейное дифференциальное уравнение относительно $A_X(p)$

$$A_X(p) - [K_{\text{ок}} \cdot W_{\text{ок}}(p)] \times \{A_X(p) [E_0 - K_{\text{д}} \cdot A_X(p) \cdot W_{\text{сф}}(p)]\} = 0. \quad (14)$$

Примем передаточную функцию сглаживающего фильтра в виде апериодического звена второго порядка.

$$W_{\text{сф}}(p) = \frac{1}{(T_{1\text{сф}} \cdot p + 1)(T_{2\text{сф}} \cdot p + 1)},$$

где $T_{1\text{сф}}$, $T_{2\text{сф}}$ – постоянные времени фазосдвигающего фильтра второго порядка.

Уравнение (14) линеаризуем относительно приращения $\Delta A_X(p)$ и, учитывая, что $A_X(p) = A_{X0} + \Delta A_X(p)$, получим

$$\Delta A_X(p) - [K_{\text{ок}} \cdot W_{\text{ок}}(p)] \times \{ \Delta A_X(p) \cdot [E_0 - K_{\text{д}} \cdot A_{X0}] - K_{\text{д}} \cdot A_{X0} \cdot W_{\text{сф}}(p) \cdot \Delta A_X(p) \} = 0, \quad (15)$$

где A_{X0} – некоторое значение амплитуды установившихся автоколебаний.

Характеристическое уравнение относительно приращения $\Delta A_X(p)$, с учетом (15), будет иметь вид:

$$1 - [K_{\text{ок}} \cdot W_{\text{ок}}(p)] \times \{ E_0 - K_{\text{д}} \cdot A_{X0} \cdot [1 + W_{\text{сф}}(p)] \} = 0. \quad (16)$$

Подставим в (16) соотношение $[K_{\text{ок}} \cdot W_{\text{ок}}(p)]$ из (12), и с учетом $W_{\text{сф}}(p)$, получим:

$$(T + \tau) \cdot T_{1\text{сф}} \cdot T_{2\text{сф}} \cdot p^3 + (T + \tau) \cdot (T_{1\text{сф}} \cdot T_{2\text{сф}}) \cdot p^2 + [T - \tau \cdot (\bar{K}_{\text{сум}} - 2)] \cdot p + (\bar{K}_{\text{сум}} - 1), \quad (17)$$

где $\bar{K}_{\text{сум}} = \bar{K}_{\text{ок}} \cdot E_0$; $\bar{K}_{\text{ок}} \cdot (E_0 - K_{\text{д}} \cdot A_{X0}) = 1$, что следует из условия существования автоколебаний.

Из характеристического уравнения (17) получим, используя критерий Гурвица, следующие условия устойчивости автоколебаний по амплитуде:

$$2 + \frac{4}{\xi^2} > \bar{K}_{\text{сум}} > 1, \quad (18)$$

$$[T - \tau \cdot (\bar{K}_{\text{сум}} - 1)] \cdot (T_{1\text{сф}} \cdot T_{2\text{сф}}) > (\bar{K}_{\text{сум}} - 1) \cdot T_{1\text{сф}} \cdot T_{2\text{сф}}. \quad (19)$$

При выполнении неравенства (18), а также неравенства $T_{1\text{сф}} < [T - \tau \cdot (\bar{K}_{\text{сум}} - 1)] / (\bar{K}_{\text{сум}} - 1)$, условие (19) выполняется при любых рациональных значениях входящих в него величин. Следовательно, в параметрическом контуре можно использовать сглаживающий фильтр в виде апериодического звена первого порядка, так как возможно выполнение равенства $T_{2\text{сф}} = 0$.

При выполнении неравенства (18), а также неравенства $T_{1\text{сф}} > [T - \tau \cdot (\bar{K}_{\text{сум}} - 1)] / (\bar{K}_{\text{сум}} - 1)$ для обеспечения устойчивости автоколебаний по амплитуде,

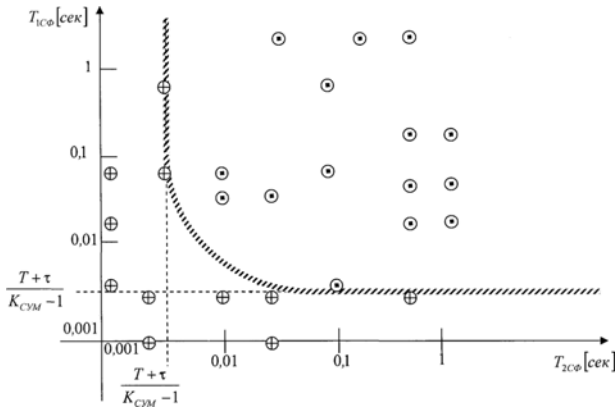


Рисунок 3 – Граница устойчивости по амплитуде относительно параметров T_{1CF} , T_{2CF} в логарифмическом масштабе:

крестики – полученные экспериментально значение параметров T_{1CF} , T_{2CF} для устойчивых по амплитуде автоколебаний; точки – полученные экспериментально значения параметров T_{1CF} , T_{2CF} для неустойчивых по амплитуде автоколебаний

с учетом (19), необходимо выполнение следующего условия:

$$T_{2CF} < \frac{T_{1CF} \cdot [T - \tau \cdot (\bar{K}_{СУМ} - 1)]}{T_{1CF} \cdot (\bar{K}_{СУМ} - 1) - [T - \tau \cdot (\bar{K}_{СУМ} - 1)]}. \quad (20)$$

На рис. 3 в логарифмическом масштабе построена граница устойчивости автоколебательной системы относительно параметров T_{1CF} , T_{2CF} . Точками и крестиками нанесены экспериментальные значения, соответственно, для неустойчивых и устойчивых по амплитуде автоколебаний.

Для автоколебательных систем с параметрическим контуром стабилизации амплитуды колебаний резонатора и с параметрическим контуром стабилизации тока возбуждения, характеристические уравнения относительно приращения амплитуды $\Delta A_Z(p)$ одинаковые и имеют вид (17). Соответственно, условия устойчивости автоколебаний одинаковы для указанных параметрических контуров.

6 АНАЛИЗ УСТОЙЧИВОСТИ АВТОГЕНЕРАТОРА С ПАРАМЕТРИЧЕСКИМ КОНТУРОМ АВТОМАТИЧЕСКОЙ РЕГУЛИРОВАНИЯ АМПЛИТУДЫ В УСЛОВИЯХ ЗАВИСИМОСТИ КОЭФФИЦИЕНТА ПЕРЕДАЧИ ОСНОВНОГО КОНТУРА ОТ АМПЛИТУДЫ

Параметрический контур вводится, для того, чтобы обеспечить работу всех усилительных элементов основного контура на линейных участках. Однако, в силу различных причин, возможен уход режима усилительных элементов за линейную зону. При этом будет наблюдаться зависимость коэффициента передачи основного контура от амплитуды колебаний

$$\bar{K}_{ОК} = \bar{K}_{ОК}(A_X). \quad (21)$$

Найдем условия устойчивости автоколебательной системы при зависимости коэффициента передачи основного контура от амплитуды и для сглаживающего фильтра в виде апериодического звена первого порядка. Для этого перепишем характеристическое уравнение (17) относительно $A_X(p)$ с учетом (21) в виде:

$$1 - \frac{\tau \cdot p + 1}{T \cdot p + 1} \times \left[\bar{K}_{ОК}(A_{X0}) \cdot (E_0 - K_D \cdot A_{X0}) - \bar{K}_{ОК}(A_{X0}) \cdot K_D \cdot A_{X0} \cdot \frac{1}{T_{CF} + p} + \frac{\partial \bar{K}_{ОК}(A_{X0})}{\partial A_{X0}} \cdot A_{X0} \cdot (E_0 - K_D \cdot A_{X0}) \right] = 0. \quad (22)$$

После очевидных преобразований (22) получим характеристическое уравнение вида:

$$T_{CF} \cdot \left[T + \tau \cdot \left(1 + \frac{\partial \bar{K}_{ОК}(A_X)}{\partial A_X} \cdot A_{X0} \cdot Y_0 \right) \right] \cdot p^2 + \left[T + \tau \cdot \left(1 + \frac{\partial \bar{K}_{ОК}(A_X)}{\partial A_X} \cdot A_{X0} \cdot Y_0 \right) - (\bar{K}_{СУМ} - 1) \cdot \tau - T_{CF} \cdot \frac{\partial \bar{K}_{ОК}(A_X)}{\partial A_X} \cdot A_{X0} \cdot Y_0 \right] \cdot p + \left[\bar{K}_{СУМ} - 1 - \frac{\partial \bar{K}_{ОК}(A_X)}{\partial A_X} \cdot A_{X0} \cdot Y_0 \right] = 0, \quad (23)$$

где Y_0 – некоторое значение управляющей координаты Y при установившихся автоколебаниях.

Для характеристического уравнения (23) получим следующие критерии устойчивости

$$\left. \begin{aligned} T_{\text{СФ}} \cdot \left[T + \tau \cdot \left(1 + \frac{\partial \bar{K}_{\text{ОК}}(A_X)}{\partial A_X} \cdot A_{X0} \cdot Y_0 \right) \right] > 0 \\ \left[T + \tau \cdot \left(1 + \frac{\partial \bar{K}_{\text{ОК}}(A_X)}{\partial A_X} \cdot A_{X0} \cdot Y_0 \right) - (\bar{K}_{\text{СУМ}} - 1) \cdot \tau - T_{\text{СФ}} \cdot \frac{\partial \bar{K}_{\text{ОК}}(A_X)}{\partial A_X} \cdot A_{X0} \cdot Y_0 \right] > 0, \\ \left[\bar{K}_{\text{СУМ}} - 1 - \frac{\partial \bar{K}_{\text{ОК}}(A_X)}{\partial A_X} \cdot A_{X0} \cdot Y_0 \right] > 0 \end{aligned} \right\} \quad (24)$$

Анализ (24) показывает, что потеря устойчивости возможна только в одном случае, если частная производная от коэффициента передачи основного контура по амплитуде имеет отрицательный знак и выполняется следующее неравенство:

$$\left| \frac{\partial \bar{K}_{\text{ОК}}(A_X)}{\partial A_X} \cdot A_{X0} \cdot Y_0 \right| > \left| 1 + \frac{T}{\tau} \right|. \quad (25)$$

Полученные соотношения использовались автором при проектировании дифференциальных датчиков виброчастотных плотномеров, что позволило уменьшить погрешность измерения плотности жидкостей и газов в среднем на 10 % [11] только за счет применения параметрических контуров.

На основании выше изложенного можно сделать следующие выводы.

Автоколебательные системы с параметрическим контуром регулирования амплитуды относятся к нелинейным системам третьего класса. Параметрические контуры регулирования амплитуды колебаний резонатора и регулирования амплитуды возбуждающего тока для магнитоэлектрической системы возбуждения колебаний резонатора динамически эквивалентны. Получены условия устойчивости автоколебаний при независимости от амплитуды коэффициента передачи основного контура: для сглаживающего фильтра в виде апериодического звена первого порядка – (18); для сглаживающего фильтра в виде апериодического звена второго порядка – (18), (19), (20). Когда коэффициент передачи основного контура зависит от амплитуды колебаний, для сглаживающего фильтра в виде апериодического звена первого порядка устойчивые автоколебания будут существовать при условии (24), а потеря устойчивости наступает при условии (25). Полученные соотношения могут быть использованы при проектировании, настройке и эксплуатации автоколебательных систем виброчастотных измерителей плотности жидкости и газов.

ПЕРЕЧЕНЬ ССЫЛОК

1. Коллак Б. Д. Методы підвищення точності автоматизованих систем обліку енергоносіїв // Вимірвальна та

- обчислювальна техніка в технологічних процесах. – 1999. – № 4. – С. 91–93.
2. Эткин Л. Г. Виброчастотные датчики. Теория и практика. – М.: Изд-во МГТУ им. Н. Э. Баумана, 2004. – 408 с.
3. Красивская М. И., Простомолотов А. С., Скачко Н. Ю. Мобильные персональные компьютеры в виртуальных частотно-цифровых средствах измерений // Сб. трудов. РНТОРЭС им. А. С. Попова. Науч. сессия, посв. Дню радио. – 2005. – Вып. IX-1. – С. 259–260.
4. Apparatus for determining and/or monitoring a predetermined contents level in a container: Пат. 5191316 США/ Dreyer V. (США); Опубл. 2.03.1993.
5. Apparatus for establishing and/or monitoring a predetermined filling level in a container. Пат. 6236322 США/ Lopatin S. (Россия), Pfeiffer H. (Германия), Mueller A. (Германия), Dreyer V. (США), Brutschin W. (США); Опубл. 22.05.2001.
6. Тараненко Ю. К. Автоколебательная система виброотно-частотных преобразователей камертонного типа // Метрология. – 1983 – № 2. – С. 39–44.
7. Козлов Ю. М., Юсупов Р. М. Беспойсковые самонастраивающиеся системы. М.: Наука, 1969. – 410 с.
8. Вибрации в технике: Справочник в 6 томах / Под. ред. В. Н. Челомея. – М.:Машиностроение, 1978. – Т.1. – 420 с.
9. Ядыгин И. Б. О периодической функции по огибающей // Автоматика и телемеханика. – 1966 – № 8. С. 25–34.
10. Ядыгин И. Б. Передаточные функции по огибающей для типовых систем // Автоматика и телемеханика. – 1971. – № 3. – С. 36–44.
11. Тараненко Ю. К. Методика проектування оптимальних диференціальних датчиків щільності рідини та концентрації дисперсних систем // Вопр. химии и хим. технологии. – 2006. – № 4. – С. 211–214.

Надійшла 18.10.06
Після доробки 7.02.07

In clause the questions of increase of accuracy of measurements are considered by use of a parametrical contour of adjustment of amplitude in auto oscillatory systems of vibrating gauges of density of liquids and gases with mechanical resonators. The block diagrams of systems of excitation of auto fluctuations of the mechanical resonator with adjustment of amplitude of fluctuations and amplitude of a current in a circuit of the activator of fluctuations are given. The questions of stability of auto fluctuations in systems with parametrical contours are considered.

У статті розглянуті питання підвищення точності вимірів шляхом використання параметричного контуру регулювання амплітуди у автоколивальних системах віброчастотних вимірювачів щільності рідини та газу з механічними резонаторами. Наведені структурні схеми систем збудження автоколивальних механічного резонатора з регулюванням амплітуди коливань та амплітуди струму у ланцюгу збудника коливань. Розглянуті питання стійкості автоколивальних у системах з параметричними контурами.

Г. М. Шило, Д. А. Коваленко

ОСОБЛИВОСТІ КОМПЕНСАЦІЇ ЗОВНІШНІХ ВПЛИВІВ ПРИ ПРИЗНАЧЕННІ НОРМАЛЬНИХ ДОПУСКІВ

Досліджується вплив зовнішніх чинників на вихідні функції електронних апаратів при нормальному законі розподілу параметрів елементів. Встановлені співвідношення між параметрами меж області працездатності та еліпсоїдної допускової області. Оцінено можливість використання спрощених компенсаційних моделей нормальних та інтервальних допусків. Досліджуються компенсаційні властивості для нелінійних вихідних функцій.

ВСТУП

Однією з важливих задач, які вирішуються при проектуванні радіоелектронної апаратури, є забезпечення необхідної точності вихідних характеристик впродовж терміну експлуатації. При виготовленні апаратури номінальні відхилення параметрів контролюються при нормальних умовах навколишнього середовища. В процесі експлуатації апаратури температурні впливи, зміна вологості, старіння матеріалів та інші чинники спричиняють подальші відхилення параметрів, що може викликати суттєву зміну вихідних характеристик. Підтримка необхідної точності має забезпечуватися вибором відповідних параметрів елементів, в тому числі допустимих відхилень та коефіцієнтів зовнішніх впливів.

Існуючі методи забезпечення точності вирішувались переважно призначенням експлуатаційних допусків, які відповідають максимально можливим відхиленням в процесі експлуатації [1–5]. Номінальні допуски тоді обчислювались за експлуатаційними допусками з урахуванням коефіцієнтів зовнішніх впливів. Можлива взаємна компенсація зміни параметрів елементів при цьому не враховується, що призводить до підвищених вимог до параметрів елементів та збільшення вартості апаратури.

Задача призначення допусків з урахуванням можливості компенсації впливу зміни параметрів елементів в процесі експлуатації почала вирішуватись для інтервальних (гарантованих) допусків, які призначаються переважно у дрібносерійному виробництві [6, 7]. Запропоновані алгоритми враховують дворівневі інтервальні моделі вихідних функцій, що забезпечує високу точність процедур вибору елементів.

Метою роботи є дослідження можливості досягнення компенсації впливу на вихідні характеристики при зміні параметрів елементів в процесі експлуатації для

нормального закону розподілу параметрів, яким відповідають нормальні допуски. Це дозволить збільшити допустимі відхилення параметрів для апаратури в умовах серійного виробництва. Для вирішення поставленої задачі необхідно:

- сформулювати математичні моделі вихідних функцій з урахуванням зовнішніх впливів при нормальному законі розподілу параметрів;
- провести дослідження компенсаційних властивостей вихідних функцій з використанням отриманих співвідношень.

1 КОМПЕНСАЦІЙНІ МОДЕЛІ НОРМАЛЬНИХ ДОПУСКІВ

Вплив на значення параметрів елементів зовнішніх факторів (температури, старіння, вологості і т. і.) можна врахувати співвідношенням:

$$x = x_r(1 + \alpha \cdot \nu), \quad (1)$$

де x_r – значення параметра при нормальних умовах навколишнього середовища; α – приведений коефіцієнт зовнішніх впливів; ν – значення зовнішніх впливів.

Під впливом зовнішніх чинників відбувається зміна параметрів функції нормального закону розподілу, що відображено на рис. 1, де m_r , m_m та m_p – математичні сподівання вхідних параметрів при нормальних умовах навколишнього середовища, нижньому та верхньому значенні зовнішніх впливів, ν_r , ν_m та ν_p – значення зовнішніх чинників при нормальних умовах навколишнього середовища, їх верхні та нижні межові значення, y – густина розподілу величини x .

Залежність параметрів нормального розподілу від зовнішніх впливів збігається з виразом (1), що відповідає основним властивостям математичного сподівання та дисперсії [8]:

$$m = m_r(1 + \alpha\nu), \quad \sigma = \sigma_r(1 + \alpha\nu), \quad (2)$$

де σ – середньоквадратичні відхилення параметрів елементів; σ_r – середньоквадратичні відхилення пара-

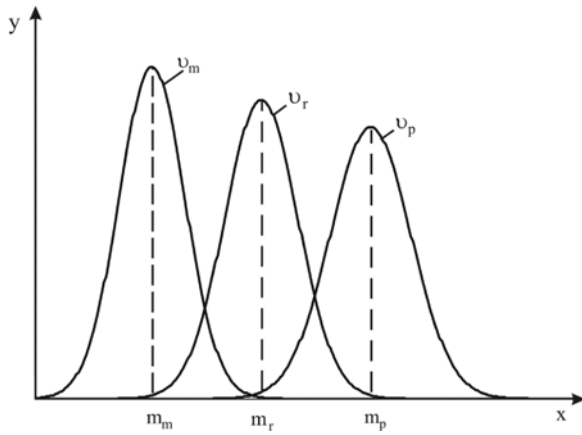


Рисунок 1 – Зміна нормального закону розподілу при впливі зовнішніх чинників

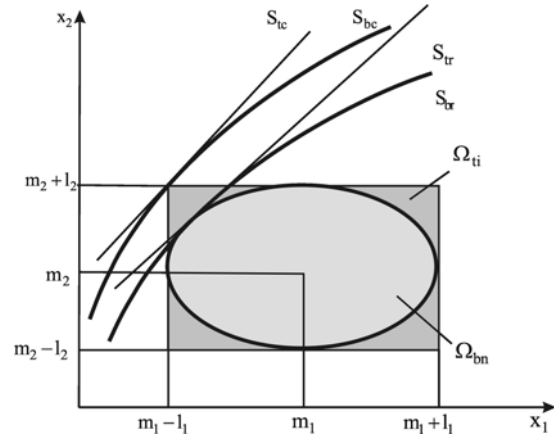


Рисунок 2 – Конфігурація допускової області та дотичних гіперповерхонь

метрів елементів при нормальних умовах навколишнього середовища.

При багатовимірному нормальному законі розподілу параметрів формується еліпсоїдна допускова область [9, 10]. Параметри цієї області пов'язані з параметрами дотичних гіперплощин області працездатності співвідношенням:

$$\sum_{i=1}^n a_i^2 l_i^2 = b^2, \quad (3)$$

де $a_i = \left. \frac{dy}{dx_i} \right|_{X_b}$ – лінійні коефіцієнти моделі вихідної

функції в точці $X_b = (x_{b1}, \dots, x_{bn})$; x_{bi} – точка дотику межевої граничної гіперповерхні області працездатності та еліпсоїдної допускової області; $l_i = \sigma_i \gamma$ – напіввісі еліпсоїду; γ – коефіцієнт поля розсіювання параметрів;

$$b = y(x_1, \dots, x_n) - a_0 - \sum_{i=1}^n a_i m_i;$$

$y(x_1, \dots, x_n)$ – межеве значення вихідної функції.

Конфігурація допускової області та розташування дотичних гіперповерхонь до області працездатності показана на рис. 2, де Ω_{tr} та Ω_{tl} – еліпсоїдна та гіперпаралелепіпедна допускові області, що відповідають нормальним та інтервальним допускам; S_{br} , S_{bl} – межові гіперповерхні області працездатності, що відповідають нормальному та інтервальному законам розподілу; S_{tr} , S_{tl} – дотичні гіперплощини до еліпсоїдної та гіперпаралелепіпедної допускової області відповідно.

Підстановка співвідношень (2) перетворює рівняння (3) до вигляду:

$$\begin{aligned} \left(y_r(1 + \alpha_y v) + c + \sum_{i=1}^n a_i m_{ri}(1 + \alpha_i v) \right)^2 &= \\ &= \sum_{i=1}^n a_i^2 l_{ri}^2 (1 + \alpha_i v)^2, \end{aligned} \quad (4)$$

де y_r та l_{ri} – межові значення вихідної функції та напіввісі еліпсоїду при нормальних умовах експлуатації;

$$c = \sum_{i=1}^n a_i x_{bi} - y(x_{b1}, \dots, x_{bn}).$$

Рівняння (4) розкладається на два рівняння, перше з яких відповідає нормальним умовам експлуатації, а друге – дії зовнішніх впливів:

$$\left(y_r + c + \sum_{i=1}^n a_i m_{ri} \right)^2 = \sum_{i=1}^n a_i^2 l_{ri}^2, \quad (5)$$

$$\begin{aligned} y_r \alpha_y \left(2b_r + y_r \alpha_y v - 2 \sum_{i=1}^n a_i m_{ri} \alpha_i v \right) &= \\ = \sum_{i=1}^n \left(a_i^2 l_{ri}^2 \alpha_i (2 + \alpha_i v) + a_i m_{ri} \alpha_i \left(2b_r - \sum_{j=1}^n a_j m_{rj} \alpha_j v \right) \right). \end{aligned} \quad (6)$$

Із рівняння (6) витікає, що при нормальному законі розподілу параметрів досягти компенсації зовнішніх впливів можливо тільки при одному зі значень зовнішніх факторів. Якщо знехтувати елементами впливу другого порядку, то рівняння (6) спрощується:

$$y_r \alpha_y b_r = \sum_{i=1}^n a_i \alpha_i (a_i l_{ri}^2 + m_{ri} b_r). \quad (7)$$

Рівняння (7) можна перетворити до вигляду:

$$s_y = \sum_{i=1}^n s_{xi}, \quad (8)$$

де $s_y = y_r \alpha_y b_r$ – фактор компенсації вихідної функції;
 $s_{xi} = a_i \alpha_i (a_i l_{ri}^2 + m_{ri} b_r)$ – фактор компенсації елемента.

Отримане рівняння має вигляд подібний рівнянням, які використовувались при компенсації зовнішніх впливів для обчислення інтервальних допусків [6, 7]. Але особливості факторів компенсації можуть змінити поведінку допусків областей. Тому необхідно провести аналіз особливостей дрейфу допусків областей нормальних допусків при зовнішніх впливах і оцінити можливість використання різних співвідношень для компенсації зовнішніх впливів.

2 ДОСЛІДЖЕННЯ КОМПЕНСАЦІЙНИХ ВЛАСТИВОСТЕЙ

Можливість використання співвідношень (7) для забезпечення компенсації зовнішніх впливів оцінювалась порівнянням отриманих параметрів елементів з параметрами, які обчислювались за допомогою співвідношення (6). Крім цього визначались параметри елементів при компенсації зовнішніх впливів для інтервальних допусків областей, які утворюються півосями еліпсоїда:

$$y_r \alpha_y = \sum_{i=1}^n a_i \alpha_i (m_i \pm l_i). \quad (9)$$

Знак в (9) залежить від розташування точок дотику межових гіперповерхонь області працездатності та інтервальних допусків областей.

Дослідження проводилось на тестовій функції

$$y(x_1, x_2) = \frac{x_1}{x_2}$$

при математичних сподіваннях елементів $m_i = 1$ та відносних відхиленнях вхідних параметрів при нормальних умовах навколишнього середовища $\delta_1 = 0,5$ та $\delta_2 = 0,2$. Нижнє та верхнє значення зовнішніх впливів задавались $v_m = -50$ °C та $v_p = 50$ °C. Для елемента x_1 значення коефіцієнта зовнішніх впливів задавалось $\alpha_1 = 0,01$ 1/К, а для другого елемента розраховувалось з умов компенсації зовнішніх чинників за допомогою співвідношень (6), (7) та (9).

При виконанні умов компенсації (6) по верхній межі області працездатності коефіцієнти зовнішніх впливів склали $\alpha_{2m} = 0,00492$ 1/К та $\alpha_{2p} = 0,00503$ 1/К для нижнього та верхнього значення зовнішніх впливів відповідно. Компенсація по нижньому рівню області працездатності приводила до коефіцієнтів $\alpha_{2m} = 0,00507$ 1/К та $\alpha_{2p} = 0,00498$ 1/К.

Порівняння результатів, отриманих з використанням виразів (7) та (9), наведені в таблиці 1 та таблиці 2, де δ_p та δ_m – відхилення коефіцієнтів зовнішніх впливів від значень, отриманих за співвідношенням (6) при верхньому та нижньому значенні зовнішніх факторів.

Таблиця 1 – Похибки використання лінійних компенсаційних моделей для верхньої межі області працездатності

$\delta, \%$	Модель компенсації	
	інтервальна	нормальна лінійна
δ_p	1,551	1,434
δ_m	0,544	0,658

Таблиця 2 – Похибки використання лінійних компенсаційних моделей для верхньої межі області працездатності

$\delta, \%$	Модель компенсації	
	інтервальна	нормальна лінійна
δ_p	1,374	1,328
δ_m	0,467	0,514

Із таблиць видно, що навіть при великій зміні параметрів елементів, які досягали 50 %, інтервальна та спрощена нормальна компенсаційні моделі дають незначну похибку у визначенні коефіцієнтів зовнішніх впливів. Але лінійна нормальна компенсаційна модель дає менший розкид параметрів. Це пов'язано з використанням параметрів межових гіперплощин дотичних до еліпсоїдної допусквої області.

Під дією зовнішніх впливів відбувається переміщення допусків областей, що показано на рис. 3, де умови компенсації виконано для верхньої межі ви-

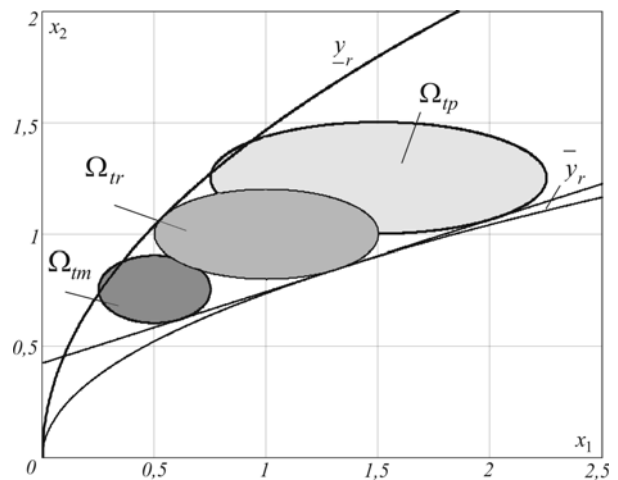


Рисунок 3 – Дрейф допусків областей при зовнішніх впливах

хідної функції. Нормальним умовам експлуатації відповідає область Ω_{tr} , а нижнім та верхнім значенням зовнішніх чинників – області Ω_{tm} та Ω_{tp} ; \underline{y}_r та \overline{y}_r – верхня та нижня межові гіперповерхні області працездатності.

Із рис. 3 видно, що допускові області при зовнішніх впливах не виходять за увігнуту межову гіперповерхню області працездатності, яка відповідає верхньому значенню вихідної функції. Опукла межова гіперповерхня області працездатності перетинається допусковими областями при межових значеннях зовнішніх впливів.

Якщо компенсація зовнішніх впливів проводиться для нижньої межі вихідної функції, то допускова область починає перетинати обидві межові гіперповерхні області працездатності.

ВИСНОВКИ

Встановлено співвідношення, що пов'язують параметри дотичних гіперплощин до області працездатності та еліпсоїдної допускової області при дії зовнішніх чинників. Виявлено, що при нормальному законі розподілу параметрів елементів компенсація дії зовнішніх впливів досягається тільки при одному зі значень зовнішніх факторів.

Використання лінійних інтервальних та спрощених нормальних моделей компенсації дає незначні похибки навіть при великих відхиленнях параметрів під дією зовнішніх чинників. Менший розкид коефіцієнтів зовнішніх впливів забезпечує лінійна нормальна модель компенсації.

Встановлено, що для нелінійних вихідних функцій допускова область не виходить з увігнуту межову гіперповерхню області працездатності, якщо компенсація проводиться для цієї межі. В інших випадках допускова область перетинає обидві межові гіперплощини.

ПЕРЕЛІК ПОСИЛАНЬ

1. Михайлов А. В., Савин К. С. Точность радиоэлектронных устройств. – М.: Машиностроение, 1976. – 214 с.
2. Фомин А. В., Борисов В. Ф., Чермошевский В. В. Допуски в радиоэлектронной аппаратуре. – М.: Сов. радио, 1973. – 108 с.
3. Цветков А. Ф. Методы расчета допусков в радиоэлектронной аппаратуре. – Рязань: РРТИ, 1970. – 131 с.
4. Фридендер И. Г. Расчеты точности машин при проектировании. – Киев–Донецк: Высшая школа, 1980. – 184 с.
5. Шило Г. М., Воропай О. Ю., Гапоненко М. П. Интервальные методы назначения эксплуатационных допусков // Радиоэлектроника. Информатика. Управление. – 2003. – № 2. – С. 78–82.
6. Кришук В. М., Шило Г. М., Намлинский А. О., Гапоненко М. П. Выбор элементов при компенсации внешних влияний // Радиоэлектроника. Информатика. Управление. – 2004. – №2. – С. 36–41.
7. Шило Г. Н., Намлинский А. А., Гапоненко Н. П. Компенсация и оптимизация при назначении номинальных допусков // Радиоэлектроника и информатика. – 2004. – №2. – С. 39–43.
8. Вентцель Е. С. Теория вероятностей. – М.: Высшая школа, 2001. – 575 с.
9. Шило Г. Н., Коваленко Д. А., Гапоненко Н. П. Расчет допусков методом касательных при корреляционной связи // Радиоэлектроника. Информатика. Управление. – 2006. – №1 – С. 29–32.
10. Шило Г. Н., Воропай А. Ю., Гапоненко Н. П. Расчет и назначение допусков методом касательных // Известия вузов Радиоэлектроника. – 2006. – № 2 – С. 43–53.

Надійшла 13.10.06
Після доробки 4.12.06

Исследуется влияние внешних факторов на выходные функции электронных аппаратов при нормальном законе распределения параметров элементов. Установлены соотношения между параметрами границ области работоспособности и эллипсоидной допусковой области. Оценена возможность использования упрощенных компенсационных моделей нормальных и интервальных допусков. Исследуются компенсационные свойства для нелинейных выходных функций.

The external factors influence on output functions of electronic devices under the normal distribution law of the parameters of elements is being investigated. The ratios between parameters of the operability region's limits and the ellipsoidal tolerance region. The application capability of simplified normal tolerances and interval tolerances models was estimated. The compensation behavior is investigated for nonlinear output functions.

МАТЕМАТИЧНЕ ТА КОМП'ЮТЕРНЕ МОДЕЛЮВАННЯ

МАТЕМАТИЧЕСКОЕ И КОМПЬЮТЕРНОЕ МОДЕЛИРОВАНИЕ

MATHEMATICAL AND COMPUTER MODELLING

УДК 519.2+519.86

А. И. Вершина, Т. Н. Семерюк, Б. Т. Солдатов

ПРИМЕНЕНИЕ ЦЕПЕЙ МАРКОВА ПРИ МОДЕЛИРОВАНИИ УЧЕБНОГО ПРОЦЕССА

Рассмотрены условия, при которых обучение можно описать как марковский процесс. Постоянство вероятностей переходов между состояниями обучения определяется законом распределения времени на обучение и дисциплиной планирования. Элементы фундаментальной матрицы тесно связаны с затратами времени на обучение.

ВВЕДЕНИЕ

Математическая модель процесса обучения [1–3] объединяет в себе представление его структуры в виде взаимосвязи этапов обучения и параметров, характеризующих свойства этих этапов и процесса обучения в целом. В качестве параметров этапов предлагается использовать вероятность качественного усвоения и качественной проверки знаний, а также затраты на выполнение этих работ. Вводимые вероятности с одной стороны отражают качество обучения, а с другой – представляют вероятности переходов от одного этапа обучения к другому. При использовании вероятностных характеристик важным моментом является их постоянство, что позволяет описывать обучение цепями

Маркова. Постоянство вероятностей переходов связано с проблемой нормирования времени, выделяемого на обучение.

ПОСТАНОВКА ЗАДАЧИ

Описание процесса обучения цепями Маркова предполагает наличие состояний процесса и вероятностей переходов от состояния к состоянию. Итерационный характер процесса обучения, предполагающий повторение некоторых состояний, требует исследования дополнительных условий, обеспечивающих постоянство вероятностей переходов. Наличие состояний, которыми завершается процесс обучения, предполагает использование поглощающих цепей Маркова.

Структура матрицы переходов для поглощающей цепи Маркова имеет вид [4]:

$$P = \begin{bmatrix} Q & R \\ O & E \end{bmatrix}, \quad (1)$$

где Q – подматрица, описывающая поведение процесса до попадания в поглощающее состояние; R – под-

© Вершина А. И., Семерюк Т. Н., Солдатов Б. Т., 2007

матрица переходов в поглощающие состояния; O , E – нулевая и единичная подматрицы.

Использование фундаментальной матрицы $N = (E - Q)^{-1}$ позволяет получить ряд важнейших характеристик исследуемого процесса. Элемент $n_{i,j}$ матрицы N дает ожидаемое количество моментов времени, которое проводит процесс в состоянии j до попадания в поглощающее состояние при условии, что он начался в состоянии i .

Вероятность завершения процесса в том или ином поглощающем состоянии в зависимости от того, какое состояние является исходным, определяется элементами матрицы $B = NR$.

Таким образом, постоянство вероятностных характеристик этапов процесса обучения является основой создания модели обучения. Это привело к необходимости исследования условий, при которых выполняется это постоянство.

ИССЛЕДОВАНИЕ ПРОБЛЕМЫ

В простейшем случае можно считать, что вероятность $\Delta k(t)$ усвоения знаний в бесконечно малом промежутке времени Δt пропорциональна величине этого промежутка, то есть,

$$\Delta k(t) = k(t + \Delta t) - k(t) = [1 - k(t)]\lambda \Delta t, \quad (2)$$

где $k(t)$ – вероятность того, что за время t знания усвоены; λ – коэффициент пропорциональности, который отражает интенсивность усвоения знаний.

Это приводит к экспоненциальному закону распределения времени на усвоение знаний:

$$p_k(t) = \frac{dk(t)}{dt} = \lambda e^{-\lambda t}.$$

При выделении времени на обучение T вероятность усвоения знаний определяется выражением

$$K(T) = \int_0^T p(t) dt = \lambda \int_0^T e^{-\lambda t} dt = 1 - e^{-\lambda T}.$$

Ожидаемое время \bar{T} на усвоение знаний равно

$$\bar{T} = \int_0^{\infty} t p(t) dt = \lambda \int_0^{\infty} t e^{-\lambda t} dt = \frac{1}{\lambda}.$$

Обычно время на обучение T_0 ограничено и выделяется пропорционально ожидаемому времени на усвоение знаний \bar{T} :

$$T_0 = \mu \bar{T} = \mu \frac{1}{\lambda},$$

где μ – коэффициент пропорциональности.

В этом случае имеем

$$K(T_0) = \int_0^{\frac{1}{\lambda}} p(t) dt = \lambda \int_0^{\frac{1}{\lambda}} e^{-\lambda t} dt.$$

После внесения λ под знак интеграла, замены переменной $z = \lambda * t$ и изменения пределов интегрирования получим

$$K(T_0) = \int_0^{\mu} e^{-z} dz = K(\mu).$$

Из этого следует, что вероятность усвоения знаний в случае выделения времени на обучение пропорционально ожидаемому времени \bar{T} не зависит от значения λ .

Для оценки затрат на обучение необходимо знать закономерности их изменения в итерационном процессе. Экспоненциальному распределению соответствует процесс без последствия, и распределение оставшегося времени на обучение не должно зависеть от того, сколько времени было ранее потрачено. Следовательно, ожидаемое дополнительное время также равно \bar{T} .

Обозначим через d вероятность необходимости дополнительного времени на обучение, тогда общее количество моментов времени

$$n_{\Sigma} = 1 \cdot (1-d) + 2 \cdot d(1-d) + 3 \cdot d^2(1-d) + \dots = (1-d) \cdot (1 + 2d + 3d^2 + \dots). \quad (3)$$

Выражение во вторых скобках представляет производную суммы ряда геометрической прогрессии:

$$s = d + d^2 + d^3 + \dots = \frac{d}{1-d} : s' = \frac{1}{(1-d)^2}, \quad (4)$$

отсюда

$$n_{\Sigma} = \frac{1}{1-d}. \quad (5)$$

Суммарные затраты времени равны

$$T_{\Sigma} = \frac{\mu \bar{T}}{1-d} = \mu n \bar{T}, \quad (6)$$

где $n = n_{\Sigma}$ – значение соответствующего элемента фундаментальной матрицы.

Описанный подход дает хорошие результаты для небольших изучаемых объемов и пригоден для моделирования нижних уровней процесса обучения. При переходе на вышестоящие уровни объем дополнительно выделяемого времени должен уменьшаться. Для описания вышестоящих уровней обучения будем считать, что вероятность в бесконечно малом промежутке времени пропорциональна величине промежутка усвоения элементов знаний, то есть экспоненциальным законом будет описываться время на усвоение только части материала от общего объема изучаемого предмета.

Изучение предмета объединяет в себе совокупность усвоения α элементов знаний, входящих в этот предмет. В этом случае плотность распределения времени на изучение предмета представляет собой композицию экспоненциальных законов распределения, что приводит к гамма-распределению:

$$p(t) = \frac{\lambda^\alpha}{\Gamma(\alpha)} t^{\alpha-1} e^{-\lambda t}, \quad (7)$$

где $\Gamma(\alpha) = \int_0^\infty t^{\alpha-1} e^{-t} dt$ – гамма-функция Эйлера.

Для произвольного времени T , затраченного на усвоение знаний, вероятность $K(T)$ усвоения знаний определяется выражением

$$K(T) = \int_0^T p(t) dt = \frac{\lambda^\alpha}{\Gamma(\alpha)} \int_0^T t^{\alpha-1} e^{-\lambda t} dt.$$

Ожидаемое время на усвоение знаний равно

$$\bar{T} = \int_0^\infty t p(t) dt = \frac{\lambda^\alpha}{\Gamma(\alpha)} \int_0^\infty t^\alpha e^{-\lambda t} dt = \frac{\alpha}{\lambda}.$$

Если время на обучение T_0 ограничено и пропорционально ожидаемому времени на усвоение знаний \bar{T} , то можно записать выражение:

$$T_0 = \mu \bar{T} = \mu \frac{\alpha}{\lambda}.$$

В этом случае имеем

$$K(T_0) = \int_0^{\mu \frac{\alpha}{\lambda}} p(t) dt = \frac{\lambda^\alpha}{\Gamma(\alpha)} \int_0^{\mu \frac{\alpha}{\lambda}} t^{\alpha-1} e^{-\lambda t} dt.$$

После внесения λ^α под знак интеграла, замены переменной $z = \lambda t$ и изменения пределов интегрирования получим

$$K(T_0) = \frac{1}{\Gamma(\alpha)} \int_0^{\mu \alpha} z^{\alpha-1} e^{-z} dz = K(\mu, \alpha). \quad (8)$$

Из этого следует, что вероятность усвоения знаний, в случае выделения времени на обучение пропорционально ожидаемому времени \bar{T} , не зависит от значения λ .

Дополнительное время для значительных объемов изучаемого материала, естественно, должно уменьшаться. Для того чтобы вероятность перехода оставалась неизменной, будем считать, что процесс усвоения знаний в одинаковой мере затрагивает каждый элемент знания. В результате общее количество элементов α останется прежним, и должно изменяться значение λ . Тогда, в соответствии с полученным выражением (8), вероятность не изменится. Дополнительное время, вы-

деляемое на усвоение определенного объема знаний, будем по-прежнему брать пропорционально ожидаемому времени.

Если считать, что элементы знаний независимы, то вероятность усвоения определенного объема знаний равна произведению вероятностей усвоения каждого элемента знаний. Если элементы знаний считать равнозначными, то $K(T_0)$ можно представить как степень α вероятности k усвоения элемента знаний

$$K(T_0) = k^\alpha,$$

отсюда

$$k = \sqrt[\alpha]{K(T_0)}.$$

Вероятность k_β того, что будет усвоено β элементов из общего количества α , равна

$$k_\beta = C_\alpha^\beta k^\beta (1-k)^{\alpha-\beta},$$

то есть величина β подчиняется биномиальному распределению [5].

Для этого распределения ожидаемое количество усвоенных элементов знаний будет равно

$$n = \alpha \cdot k = \alpha \cdot \sqrt[\alpha]{K(T_0)}.$$

Для гамма-распределения будем иметь выделяемое дополнительное время $T_{\text{доп}}$, равное

$$T_{\text{доп}} = \mu \frac{n}{\lambda} = \mu \frac{\alpha k}{\lambda} = \mu \frac{\alpha}{\lambda} \sqrt[\alpha]{K(T_0)}.$$

Для того, чтобы вероятность $K(T_0)$ осталась постоянной необходимо оставить постоянным значение α , а изменять значение λ . Выражение (6) для общих затрат на обучение приобретает вид

$$T_\Sigma = \mu \alpha \sqrt[\alpha]{K(T_0)} n \bar{T}$$

или

$$T_\Sigma = \mu \cdot n \cdot \frac{\alpha}{\lambda} \sqrt[\alpha]{K\left(\mu \frac{\alpha}{\lambda}\right)}.$$

Если все элементы получения знаний различаются по сложности и имеют различные значения λ , то выражение для композиции экспоненциальных законов распределения имеет довольно сложный вид. В этом случае воспользуемся условием Линдберга, которое должно выполняться для любого $\tau > 0$:

$$\lim_{\alpha \rightarrow \infty} \left(1 / \sum_{i=1}^{\alpha} D_i \right) \sum_{i=1}^{\alpha} \int_{|x-m_i| > \tau \sqrt{\sum_{i=1}^{\alpha} D_i}} (x-m_i)^2 f_i(x) dx = 0,$$

где $m_i, D_i, f_i(x)$ и α – математическое ожидание, дисперсия и плотность распределения i -й случайной величины x .

При выполнении этого условия закон распределения суммы независимых случайных величин неограниченно приближается к нормальному закону распределения:

$$f(x) = \frac{1}{\sigma\sqrt{2\pi}} e^{-\frac{(x-m)^2}{2\sigma^2}}. \quad (9)$$

Для произвольного времени T , затраченного на усвоение знаний, вероятность $K(T)$ усвоения знаний определяется выражением

$$K(T) = \frac{1}{\sigma\sqrt{2\pi}} \int_{-\infty}^T e^{-\frac{(z-m)^2}{2\sigma^2}} dz = \frac{1}{2} + \Phi\left(\frac{T-m}{\sigma}\right),$$

где $\Phi(x) = \frac{1}{\sqrt{2\pi}} \int_0^x e^{-\frac{t^2}{2}} dt$ – функция Лапласа.

Если выделять время на обучение, исходя из выражения

$$T_0 = m + \mu\sigma, \quad (10)$$

то

$$K(T_0) = \frac{1}{2} + \Phi(\mu),$$

то есть зависит только от коэффициента μ . Его постоянство обеспечит постоянство переходов между состояниями процесса обучения.

Если выделять время на обучение пропорционально ожидаемому времени, то есть

$$T_0 = \mu m,$$

то значение $K(T_0)$ будет иметь вид

$$K(T_0) = \frac{1}{2} + \Phi\left(\frac{(\mu-1)m}{\sigma}\right). \quad (11)$$

В этом случае $K(T_0)$ зависит от трех величин μ , m и σ .

Будем считать, что случайные значения времени для различных элементов обучения некоррелированы. В соответствии с теоремой сложения математических ожиданий и дисперсии суммы случайных некоррелированных величин, имеем

$$m = \sum_{i=1}^{\alpha} m_i \text{ и } D = \sigma^2 = \sum_{i=1}^{\alpha} D_i.$$

Для равнозначных элементов обучения имеем

$$m = \sum_{i=1}^{\alpha} \frac{1}{\lambda} = \frac{\alpha}{\lambda} \text{ и } D = \sigma^2 = \sum_{i=1}^{\alpha} \frac{1}{\lambda^2} = \frac{\alpha}{\lambda^2}. \quad (12)$$

После подстановки (12) в (11) получим

$$K(T_0) = \frac{1}{2} + \Phi((\mu-1)\sqrt{\alpha}).$$

Этот результат согласуется с выводами, которые следовали из выражения (8), то есть $K(T_0)$ зависит от величин μ и α . Дальнейшие рассуждения о суммарных затратах аналогичны (3–6).

Для неравнозначных элементов обучения имеем

$$m = \sum_{i=1}^{\alpha} \frac{1}{\lambda_i} \text{ и } D = \sigma^2 = \sum_{i=1}^{\alpha} \frac{1}{\lambda_i^2}$$

и выражение (11) приобретает вид

$$K(T_0) = \frac{1}{2} + \Phi\left(\frac{(\mu-1) \sum_{i=1}^{\alpha} \frac{1}{\lambda_i}}{\sqrt{\sum_{i=1}^{\alpha} \frac{1}{\lambda_i^2}}}\right).$$

Таким образом, условие выделения времени на обучение может обеспечить условия постоянства вероятностных характеристик. Наиболее удобным является алгоритм, определяемый выражением (10), который делает эти характеристики зависимыми только от параметра μ . Причем, если μ в выражении (10) принять равным нулю, то это будет соответствовать выделению времени на обучение, равному ожидаемому \bar{T} . Значение вероятности $K(T_0)$ будет равным 0,5.

ОСНОВНЫЕ ПОДХОДЫ К ВЫДЕЛЕНИЮ ВРЕМЕНИ НА ОБУЧЕНИЕ

Выделение времени на изучение элементов обучения.

Время на обучение выделяется пропорционально ожидаемому времени на усвоение элементов знаний. Эта величина остается постоянной для выделения дополнительного времени в соответствии с отсутствием последствия.

Выделение времени на изучение совокупности равнозначных элементов обучения.

Время на обучение выделяется пропорционально количеству элементов объема изучаемого материала и обратно пропорционально интенсивности изучения элементов обучения. Эта величина изменяется пропорционально ожидаемому объему изучаемого материала.

Выделение времени на изучение совокупности неравнозначных элементов обучения.

Время на обучение выделяется пропорционально ожидаемому объему времени на изучение материала с учетом отклонений. Эта величина также изменяется пропорционально ожидаемому объему изучаемого материала и вычисляется как и в предыдущем случае.

ВЫВОДЫ

Получены выражения для законов распределения времени на усвоение элементов знаний и некоторого объема знаний как совокупности этих элементов, которые описываются соответственно экспоненциальным распределением (2), гамма-распределением (7), как предельный случай, нормальным распределением (9).

Для каждого распределения приведена процедура выделения времени на обучение, которая обеспечивает постоянство вероятностных характеристик.

Процедура планирования выделения времени на обучение позволяет создать условия постоянства вероятностных характеристик, что позволяет описывать обучение как марковский процесс (1).

Вероятность получения элемента знаний в малом промежутке времени пропорциональна величине этого промежутка, приводит к экспоненциальному закону вероятностей распределения времени на усвоение элемента знаний.

Композиция экспоненциальных законов при постоянстве интенсивности обучения приводит к гамма-распределению времени на усвоение некоторого объема знаний как совокупности элементов знаний.

Условие Линдеберга позволяет рассматривать в качестве предельного распределения нормальный закон. Выделение времени на обучение определяет постоянство вероятностных характеристик, что позволяет при моделировании процесса обучения использовать цепи Маркова.

Оценка общих затрат на обучение должна учитывать их изменение в условиях итерационного характера процесса обучения. При этом для нижних уровней обучения, когда объектом обучения являются элементы знаний, в соответствии со свойством экспоненциального закона, заключающегося в отсутствии последей-

ствия, объемы дополнительного времени на обучение совпадают с исходным объемом выделяемого времени.

Для вышестоящих уровней необходимо учитывать сокращение времени. Использование биномиального закона распределения позволяет оценить это сокращение. В то же время уменьшение числа изучаемых элементов может привести к изменению вероятностей переходов, поэтому предложена процедура изменения интенсивности обучения, эквивалентная сокращению количества элементов.

ПЕРЕЧЕНЬ ССЫЛОК

1. Вершина А. И., Солдатов Б. Т. Моделирование процесса обучения // Радиоелектроніка, інформатика, управління. – Запоріжжя: ЗДТУ. – 2003. – № 1. – С. 65–72.
2. Вершина А. И., Солдатов Б. Т., Ермоленко А. А. Учебный процесс как иерархическая система // Радиоелектроніка. Інформатика. Управління. – 2004. – № 1. – С. 54–62.
3. Вершина А. И., Киричек Г. Г., Пиза Д. М. Модель информационного обеспечения учебного процесса университета // Радиоелектроніка. Інформатика. Управління. – 2004. – № 2. – С. 64–68.
4. Кемени Дж., Снел Дж. Конечные цепи Маркова. – М.: Наука, 1970. – 272 с.
5. Вентцель Е. С. Теория вероятностей: Учебник для вузов. – 7-е изд., стер. – М.: Высш. шк., 2001. – 575 с.

Надійшла 26.05.06
Після доробки 2.11.06

Розглянуто умови, при яких навчання можна описати як марківський процес. Незмінність імовірності переходів між станами навчання визначається законом розподілу часу на навчання та дисципліною планування. Елементи фундаментальної матриці тісно пов'язані з витратами часу на навчання.

The conditions are considered, at which training is possible to describe as process of Markov. The law of distribution of time on training and discipline of planning determine constancy of probabilities of transitions between states of training. The elements of a fundamental matrix are closely connected to expenses time for training.

УДК 519.2:368.01

С. Н. Герасин, Е. В. Слипченко

МЕТОДЫ ОЦЕНКИ РЕНТАБЕЛЬНОСТИ ФИНАНСОВОЙ ДЕЯТЕЛЬНОСТИ ИЕРАРХИЧЕСКИХ БИЗНЕС-СТРУКТУР

Изучена иерархическая структура бизнес-процессов. Выделен класс типовых и элементарных бизнес-процессов. Для них сформулированы типичные задачи оптимального управления оборотными средствами и предложены конкретные алгоритмы их решения.

ВВЕДЕНИЕ

Одной из главных проблем подавляющего большинства предприятий является улучшение их финансово-экономического состояния в условиях дефицита обо-

© Герасин С. Н., Слипченко Е. В., 2007

ротных средств [1, 2]. Это связано и с высокими темпами внедрения инноваций, с осторожной кредитной политикой банков и обострением конкуренции. В таком случае весьма актуальной является задача разработки эффективных механизмов управления оборотными средствами в различных бизнес-структурах. В данной статье на примере трехуровневого бизнес-процесса будут формализованы типичные задачи оптимального управления финансовыми потоками с целью максимизации прибыли и предложены конкретные алгоритмы решения поставленных задач.

ПОСТАНОВКА ЗАДАЧИ

Рассмотрим простую трехуровневую структуру бизнес-процессов (БП). На верхнем уровне находится направление бизнеса, его реализует организационное подразделение – бизнес-единица (БЕ), выполняющая несколько основных типовых БП. Каждый типовой БП (ТБП) компонуется (конструируется) из элементарных БП (ЭБП) нижнего уровня. Вначале рассмотрению подлежит частный случай бизнес-единицы, реализующей один типовой БП.

Бизнес-единица создается для обеспечения максимального зарабатывания денег на порученном ей направлении бизнеса и не имеет накладных расходов (только прямые, которые нужны для реализации БП и исчезают при его прекращении), критерием финансового результата является наличие денег (Cash) на конец планового периода.

Бизнес строится на производстве и продаже определенной группы продукции (товаров или услуг). Чтобы произвести и продать одну единицу продукции (изделие, единичный заказ и т. п.) нужно выполнить заданный комплекс работ, составляющий операцию, которую будем называть – элементарным бизнес-процессом (ЭБП).

Поскольку существуют примеры фирм, созданных для реализации только одного основного бизнес-процесса (как правило, это малые предприятия), то на уровне одной бизнес-единицы решается практически полный комплекс задач управления финансами. Причем, при решении задачи 1 (формулировка цели), как правило, основная цель и критерии формулируются следующим образом: «Обеспечение прибыльного и финансово стабильного функционирования и развития фирмы».

Коммерческие предпринимательские организации чаще всего предпочитают в качестве основной характеристики бизнеса критерий максимизация величины денежной массы (Cash). Ниже, если не оговорено специально, имеется в виду этот критерий.

Рассмотрим сначала критерий максимизации наличия (Cash) за период T . В этом случае естественно основным инструментом для моделирования финан-

совых аспектов бизнес-процесса, подготовки и принятия финансовых решений являются методы анализа финансовых потоков (Cash flow analysis).

С точки зрения финансовых потоков за период T любой ТБП и реализующая его бизнес-единица (БЕ) описывается:

- 1) потоком поступлений денег $S = \{S(t_k)\}$ в момент времени t_k (приток финансов);
- 2) потоком платежей $C = \{C(t_l)\}$ в моменты времени t_l , (отток финансов);
- 3) внутренними для ТБП (горизонтальными) потоками между ЭБП, образующими этот ТБП;
- 4) внешними для ТБП потоками (вертикальными) между ТБП и БП:
 - приток финансов от БЕ на ТБП $_i$ $\Delta_{БЕ,i}$ (например, стартовых вложений C_0);
 - отток финансов от ТБП $_i$ к БЕ $\Delta_{i,БЕ}$ (например, поступления от БП S).

Аналогично для БЕ имеются внешние потоки финансов: приток $\Delta_{0,БЕ}$ и отток $\Delta_{БЕ,0}$.

Тогда основная задача максимизации заработанного за период суммы денег ΔF формулируется следующим образом.

А. На уровне элементарного БП:

Из множества возможных ЭБП выбрать такой, который максимизирует величину

$$\Delta F_{ЭБП}(\tau) = \Delta_{ЭБП,0} - \Delta_{0,ЭБП} = S(\tau) - C_0.$$

Б. На уровне бизнес-единицы:

$$\Delta F_{БЕ}(T) = \sum_{i=1}^n \Delta F_i(T) - \sum_{t=0}^T (\Delta_{БЕ,0}(t) - \Delta_{0,БЕ}(t)) - F_{БЕ}(0),$$

где $F_{БЕ}(0)$ – начальный объем собственных оборотных средств бизнес-единицы в момент времени $t_0 = 0$.

ЭЛЕМЕНТАРНЫЙ БИЗНЕС-ПРОЦЕСС. ОСНОВНЫЕ ХАРАКТЕРИСТИКИ И ЗАДАЧИ ПРИНЯТИЯ РЕШЕНИЙ

Простейший вариант ЭБП – это одна операция продолжительностью τ , в начале которой вкладывается сумма $C(0)$, а через время τ возвращается большая сумма $S(\tau)$. В этом случае все финансовые характеристики ЭБП совпадают с точностью до интерпретаций с классическими характеристиками и формулами для любой простейшей финансовой сделки: однократное предоставление в долг суммы $C(0)$ с условием, что через время τ будет возвращена сумма $S(\tau)$.

Однако известные публикации [3, 4] в основном посвящены анализу одного частного случая БП – финансовых операций, связанных с работой на рынке

капитала (в основном с выдачей в долг и возвратом денег или ценных бумаг).

В данной работе нами рассматриваются активные предпринимательские операции (бизнес-процессы), ориентированные на производство товаров и услуг, в результате вложений в которые создается добавленная стоимость и зарабатываются деньги. Это, в частности, дает возможность, рассматривая с единой методической точки зрения пару операций (бизнес-процесс и финансовую операцию, например, по выдаче кредита), определить взаимовыгодные условия взятия заемных средств для реализации БП.

Ниже рассмотрены следующие два варианта ЭБП:

1. ЭБП-1: фиксирована продолжительность бизнес-цикла τ , но объем продаж S и вложения C могут непрерывно изменяться в интервалах от нуля до максимально возможного: $0 \leq S(\tau) \leq S^{\max}$, $0 \leq C_0 \leq C_0^{\max}$. Далее для ЭБП-1 используется условное название «непрерывный». Пример ЭБП-1 — партия товаров, непрерывное производство, когда $S^{\max} = \min\{S_{\text{рын}}^{\max}, S_{\text{пр}}^{\max}\}$, $S_{\text{рын}}^{\max}$ — ограничение по возможностям производства.

В некоторых случаях возможен непрерывный ЭБП-1А, у которого продолжительность τ может изменяться при изменении выделяемого на их выполнение ресурса C , $\tau = f(C)$. Пример ЭБП-1А: проекты: создание крупных изделий (например, суда, турбины) стройки, другие комплексы работ, C — выделяемые объемы ресурсов типа мощностей (например, людей), W — объем работ. Тогда, например, в простейшем варианте $\tau = f(\tau) = \frac{W_0}{C(0)}$.

2. ЭБП-2: фиксированы все характеристики ЭБП: C_0 , τ и $F(\tau)$. Далее используется условное название «дискретный» ЭБП. Пример: поток изделий или проектов средней величины с фиксированной технологией.

Наиболее распространен на практике ЭБП-1, к которому приближается и описание БП из потока дискретных ЭБП-2 при увеличении масштаба (числа дискретных изделий). В связи с этим в работе основное внимание уделяется непрерывному ЭБП-1.

ПЕРВИЧНЫЕ ДАННЫЕ И ПОКАЗАТЕЛИ ЭФФЕКТИВНОСТИ ЭБП

Первичные данные:

τ — продолжительность реализации ЭБП (с момента вложения денег $C(0)$, до момента получения заработанных денег $S(\tau)$);

$S(\tau)$ — выручка от продажи товара (услуги), сумма поступлений денежных средств в конце ЭБП;

$C(0)$ — стартовые вложения, соответствующие прямым переменным затратам, минимально необходимым для выполнения только ЭБП (без косвенных затрат), приносящего за время τ выручку $S(\tau)$.

Основной характеристикой ЭБП является его выгодность или невыгодность с финансовой точки зрения, которая определяется величиной дохода (прибыли) $F(\tau) = S(\tau) - C(0)$.

Возможны только три случая:

- 1) Если $F(\tau) > 0$, то ЭБП является прибыльным (доходным) делом.
- 2) Если $F(\tau) = 0$, то ЭБП является безубыточным.
- 3) Если $F(\tau) < 0$, то ЭБП является убыточным.

Случай 3, когда элементарный БП (производство одного изделия, услуги) убыточен, т. е. прямые переменные затраты на изделие выше цены его продажи (выручки), означает, что чем больше мы произведем и продадим таких изделий, тем больший убыток мы получим.

Случай 2, когда ЭБП является неприбыльным, но в то же время безубыточным в ряде ситуаций может оказаться выгодным. Например, имеются избыточные мощности, обеспечивающие службы и их достаточно для увеличения на ΔS объема производства и продаж, в состав затрат $C(0)$ входит и зарплата (а иногда и амортизация используемого оборудования), при высоком уровне налогообложения прибыли. Тогда условие $S(\tau) = C(0)$ достаточно для воспроизводства, а в результате дополнительной реализации ЭБП организация получает увеличение фонда зарплаты и загрузки мощностей, что снижает долю коэффициента постоянных расходов в себестоимости.

В связи с этим далее в качестве ЭБП рассматриваются только первые два случая (безубыточный и прибыльный), обладающие свойством

$$F(\tau) \geq 0 \quad (1)$$

(необходимое условие существования бизнеса).

Условие $F(\tau) = 0$ может обеспечить при многократном повторении ЭБП только воспроизводственный процесс (с сохранением уровня объема производства и продаж). При условии $F(\tau) > 0$ возможен рост, развитие наращивание бизнеса. Тогда появляются показатели, характеризующие эффективность (финансовую динамику БП в зависимости от числа циклов (ЭБП) и времени, а также характеристики отдачи на единицу вложений).

ОСНОВНЫЕ ХАРАКТЕРИСТИКИ ЭФФЕКТИВНОСТИ

Первый показатель эффективности характеризует прибыльность по отношению к начальной сумме (затратам):

$$r(\tau) = \frac{F(\tau)}{C(0)} = \frac{S(\tau) - C(0)}{C(0)}. \quad (2)$$

Этот показатель эффективности (далее будем обозначать r) называют в зависимости от контекста нормой прибыли, доходностью, рентабельностью, относительным ростом, процентной ставкой (при умножении на 100 %) или ставкой процента (английский эквивалент return, interest rate).

В данной работе применительно к БП будем использовать термины доходность (бизнеса) и рентабельность. Причем для рассматриваемых ниже задач управления оборотными средствами (тогда $C(0)$ – это оборотные средства, $S(\tau)$ – выручка), адекватным для r является термин «рентабельность по оборотным средствам» или просто рентабельность. По определению ЭБП в $C(0)$ учитываются только прямые переменные затраты, тогда $F(\tau)$ – это маржинальная прибыль, и r – это маржинальная рентабельность ЭБП, т. е. удельная маржинальная рентабельность на единицу изделия (товара, услуги). Из определения r следует связь начальных вложений и выручки [5]

$$S(\tau) = C(0)(1 + r), \quad (3)$$

величина $\frac{S(\tau)}{C(0)} = k_H = (1 + r)$ называется коэффициентом наращивания бизнеса.

Второй классический показатель – прибыльность по отношению к конечной сумме (выручке).

$$d_\tau = \frac{F(\tau)}{S(\tau)} = \frac{S(\tau) - C(0)}{S(\tau)} = d. \quad (4)$$

Этот показатель (далее обозначение d) в различных ситуациях называют рентабельностью по выручке (по обороту) или дисконтом (discount rate). Применительно к БП d показывает, какую долю (процент) от выручки составляет прибыль, т. е. какой доход, прибыль мы получаем с каждой гривны поступлений (выручки). Простые преобразования показывают, что

$$d = 1 - \frac{C(0)}{S(\tau)} = 1 - \alpha,$$

где α – доля прямых переменных затрат в выручке (цене для одного изделия). Тогда d – это доля выручки, остающаяся после вычета доли прямых переменных затрат в выручке [6].

ЗАДАЧИ И АЛГОРИТМЫ ПРИНЯТИЯ РЕШЕНИЙ ПО ЭБП

Наиболее выгодным вариантом ЭБП является тот, при котором при тех же затратах собственных средств достигается максимум $F(\tau)$ (прибыли, дохода).

Задача 1. Имеется только один фиксированный вариант ЭБП. Прямые переменные затраты на единицу продукции $C(0)$ вкладываются до начала БП, выручка $S(\tau)$ получается по окончании ЭБП через время τ . $C(0)$ и $S(\tau)$ в сопоставимых ценах (с учетом инфляции). Требуется принять следующее решение: исключить данный ЭБП из дальнейшего рассмотрения или нет.

Правило (алгоритм 1) принятия решения:

1. Определяются первичные показатели $C(0)$, $S(\tau)$, τ .
2. Вычисляется значение дохода $F(\tau) = S(\tau) - C(0)$.
3. В случае если $F(\tau) < 0$ ЭБП исключается из рассмотрения, если $F(\tau) \geq 0$ оставляется для дальнейшего анализа.

Задача 2. Условие задачи 1 в условиях дефицита средств. На начало периода $[0, \tau]$ имеется наличие (собственные средства) $F(0)$, $0 \leq F(0) < C(0)$, но есть возможность привлечения заемных средств, которые должны быть возвращены в момент τ , например, со ставкой процента i (далее кредитной ставкой). Требуется принять решение: брать заемные средства или нет.

Алгоритм 2.

1. Выполняется алгоритм 1. Если рассмотрение ЭБП нецелесообразно, то и кредит брать нецелесообразно. Если целесообразно, то шаг 2.

2. Задача 2 приводится к задаче 1: находится новое значение

$$\tilde{S}(\tau) = S(\tau) - i(C(0) - F(0)).$$

3. Применяется алгоритм 1: если $\Delta \tilde{F}(\tau) = \tilde{S}(\tau) - C(0) < 0$, то брать заемные средства нецелесообразно, в противном случае целесообразно.

Задача 3. Имеется n_1 альтернативных вариантов вложения имеющейся суммы $F(0)$ в ЭБП одинаковой продолжительности τ . Требуется выбрать один наиболее выгодный вариант ЭБП.

Правило (алгоритм 3) принятия решений:

Предварительный шаг. Используется алгоритм 1, после применения которого остается n альтернатив.

1. Для всех n вариантов ЭБП, $j = 1, 2, \dots, n$ вычисляется значение рентабельности

$$r_{j\tau} = \frac{S_j(\tau) - C_j(0)}{C_j(0)}.$$

2. Выбирается ЭБП i^* с максимальной рентабельностью

$$j^* = \arg \max_{1 \leq j \leq n} \{r_j\}.$$

Вопросы близкие по содержанию были рассмотрены в работах [7–10].

ВЫВОДЫ

В статье рассмотрены бизнес-структуры иерархического типа, их аналогом являются матричные проектные структуры, которые успешно использовались в СССР, при программно-целевом управлении, например развитии оборонной промышленности. Разработаны и обоснованы эффективные механизмы управления оборотными средствами в производственно-торговых фирмах, ориентированных на бизнес-процессы и имеющих бизнес-единицы, в условиях дефицита оборотного капитала. Основное внимание уделено задачам подготовки и принятия выгодных финансовых решений, являющихся основой эффективного процесса управления.

ПЕРЕЧЕНЬ ССЫЛОК

1. Арзякова О. Н., Агарков Г. А., Кормышев В. М. Управление финансовыми ресурсами государственного вуза в рыночных условиях (информационное и математическое моделирование) // Университетское управление: практика и анализ. – 1998. – № 4(7). – С. 49–51.
2. Башарин Г. П. Начала финансовой математики. – М.: Инфра-М, 1997. – 160 с.
3. Вендров А. М. Методы и средства моделирования бизнес-процессов // Информационный бюллетень «Jet Info». – 2004. – № 10(137). – С. 1–32.
4. Данич В. Денежно-финансовые и товарные потоки // Бизнес-информ. 1999. – № 13–14. – С. 43–48.
5. Минко Е. П., Слипенко Е. В. Математические модели финансовых потоков, возникающих в процессе дея-

- тельности многоуровневых бизнес-структур // Радиоэлектроника и информатика. – 2003. – № 1. – С. 140–143.
6. Пуртов В. Ф., Слипенко Е. В. Управление финансовой деятельностью иерархических бизнес-структур // Теория и техника передачи, приема и обработки информации: Материалы 9-й Международной конференции (19–22 сентября). – Туапсе-Харьков: Харьковский национальный университет радиоэлектроники, 2003. – С. 441–442.
 7. Слипенко Е. В. Исследование влияния экономических факторов на рентабельность деятельности бизнес-структур // Автоматизированные системы управления и приборы автоматизации. – 2002. – Вып. 119. – С. 16–21.
 8. Слипенко Е. В. Оптимальное управление оборотными средствами с учетом специфики банка // Радиоэлектроника и информатика. – 2004. – № 2 – С. 140–144.
 9. Слипенко Е. В. Статистический анализ влияния экономических факторов на рентабельность деятельности бизнес-структур в Харьковской области // Радиоэлектроника и информатика. – 2002. – № 2. – С. 119–122.
 10. Слипенко Е. В., Пуртов В. Ф. Управление финансовыми потоками иерархических бизнес-структур // Теория и техника передачи, приема и обработки информации: Материалы 10-й Международной конференции (1–4 октября). – Туапсе-Харьков: Харьковский национальный университет радиоэлектроники, 2004. – С. 315–316.

Надійшла 22.05.06

Вивчена ієрархічна структура бізнес-процесів. Виділений клас типових і елементарних бізнес-процесів. Для них сформульовані типові завдання оптимального управління оборотними коштами і запропоновані конкретні алгоритми їх рішення.

The hierarchical structure of business-processes is studied. The class of business-processes of models and elementary is selected. For them the typical tasks of optimum management circulating assets are formulated and the concrete algorithms of their decision are offered.

УДК 004.65:330.47

М. К. Демин

МЕТОДЫ ПРЕДСТАВЛЕНИЯ ИЕРАРХИЧЕСКОЙ ИНФОРМАЦИИ В РЕЛЯЦИОННЫХ БАЗАХ ДАННЫХ ИНФОРМАЦИОННО- УПРАВЛЕНЧЕСКИХ АРХИТЕКТУР

В статье анализируются возможные методы представления иерархической информации в базах данных информационно-управленческих архитектур, приведены результаты тестирования эффективности этих методов в реальных задачах, выбран лучший метод, даны общие рекомендации по применению рассмотренных методов в реляционных базах данных.

ВВЕДЕНИЕ

База данных информационно-управленческих архитектур (ИУА) должна хранить и обрабатывать большое количество иерархической информации. Традици-

онно обработка данных, носящих иерархический характер, организуется при помощи рекурсивных процедур. Многие современные реляционные СУБД обеспечивают возможность использования рекурсии в хранимых процедурах (InterBase, Oracle). Следовательно, в таких СУБД возможна обработка иерархической информации стандартными методами. Но производительность в таких случаях может оказаться неудовлетворительной. Исследование эффективности различных методов представления иерархической информации в задачах характерных для БД ИУА ранее не проводилось.

© Демин М. К., 2007

С одним методом представления информации можно использовать различные методы обработки. Но для рассматриваемых методов можно отметить, что относительно определенного представления эффективный метод обработки будет только один. В дальнейшем в статье под методом представления подразумевается, собственно, метод представления и обусловленный им метод обработки иерархической информации.

Наиболее распространенным является метод представления иерархической информации, основанный на хранении матрицы смежности в виде таблицы:

NODE_ID	PARENT_ID

где NODE_ID – идентификатор (код) узла иерархии, PARENT_ID – код родительского узла.

В [1] описано применение данного метода и отмечена возможность использования рекурсивных хранимых процедур для извлечения иерархической информации. В [2] описана структура реляционной БД ИУА, в которой для представления иерархической информации использовался именно этот метод. Его использование возможно только в том случае, если СУБД поддерживает рекурсивные хранимые процедуры. Иначе для обработки иерархической информации потребуется отдельное клиентское приложение, которое будет использовать временные таблицы или курсоры [3].

Метод вложенных множеств был предложен Джо Селко в [4]. Его проще всего объяснить при помощи рис. 1.

Таблица, содержащая информацию о дереве, будет следующей:

NODE_ID	LEFT_	RIGHT_

Дерево обходится в прямом порядке. Считается, что каждый узел посещается дважды (на спуске и на возврате). Начинается обход при значении счетчика 1. При каждом посещении узла счетчик увеличивается на 1. При первом посещении узла значение счетчика помещается в столбец LEFT_, а при втором в столбец RIGHT_. Возможность анализа иерархической информации обеспечивает, то что для всех дочерних узлов некоторого элемента дерева выполняются следующие условия: LEFT_дочернего > LEFT_родительского и RIGHT_дочернего < RIGHT_родительского.

Строковый метод представления иерархий описан в [5]. Таблица, содержащая информацию о дереве для данного метода:

NODE_ID	HR_POS

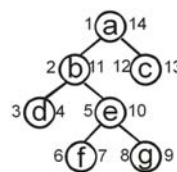


Рисунок 1 – Метод вложенных множеств представления иерархической информации

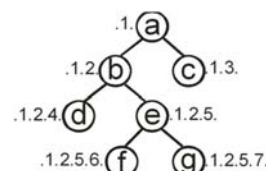


Рисунок 2 – Строковый метод представления иерархической информации

Здесь значением атрибута HR_POS является строка, по которой можно восстановить позицию элемента в иерархии. Эта строка состоит из идентификаторов родительских узлов и идентификатора самого элемента, разделенных точками. Для дерева, представленного на рис. 1, соответствующие значения строки HR_POS указаны на рис. 2.

Здесь полагается, что код узла a – 1, b – 2, и т. д.

Так как точки являются разделителями идентификаторов, то целесообразно их использовать исключительно в этом качестве. Т. е. точки в начале и в конце строки можно опустить (например, вместо значения «.1.2.5.6.» использовать «1.2.5.6»). Тем более что от этого реализация методов работы с иерархической информацией не усложняется.

Аналогичный метод применялся в [6], но там вместо идентификаторов узлов в строке, определяющей позицию в иерархии, использовалась локальная нумерация в пределах родительского узла. Следует отметить, что в этом случае гораздо сложнее обеспечить автоматический контроль целостности данных.

ПОСТАНОВКА ЗАДАЧИ

Необходимо провести сравнение эффективности различных методов представления иерархической информации в реляционных БД, а также выбрать метод для реляционной БД ИУА.

МЕТОДЫ И АЛГОРИТМЫ РЕШЕНИЯ

Вначале остановимся на таких немаловажных факторах, влияющих на выбор метода, как сложность

реализации, возможность осуществления автоматического контроля целостности данных и требования, предъявляемые этими методами к СУБД.

Наиболее простым в реализации является метод хранения матрицы смежности. Автоматический контроль целостности данных в этом случае легко реализуется при помощи триггеров. Так, при вставке элемента можно проверять присутствует ли в таблице родительский элемент. В приведенном нами примере организации данных для этого метода того же самого можно добиться использованием внешнего ключа. Но ввиду многослойности ИУА в реальной БД родительский элемент может присутствовать в таблице, определяющей подчинение, но относиться к другой иерархии. В таком случае необходимо использование триггера. При удалении элемента можно удалять все дочерние узлы, т. е. узлы, у которых PARENT_ID равно NODE_ID, удаляемого элемента. Для таких элементов вызовется тот же триггер, и будет произведено рекурсивное удаление поддерева.

При использовании этого метода СУБД должна поддерживать рекурсивные хранимые процедуры.

Следует отметить, что производительность выборки иерархий, осуществляемой при помощи только рекурсивных хранимых процедур, в случае, если хранятся большие объемы иерархической информации, оказывается недостаточной даже для простого извлечения иерархий и передачи их клиенту. Поэтому метод был оптимизирован следующим образом (на примере БД ИУА, описанной в [1]).

Таблица SUBJECTION, в которой храниться информация о подчинении узлов, была модифицирована и приняла вид:

NODE_ID	PARENT_ID	TH_ID	ENT_ID

Т. е. был добавлен столбец ENT_ID, что позволило переносить (копировать) информацию об извлекаемой иерархии во вспомогательную таблицу перед началом работы рекурсивной хранимой процедуры. После извлечения информации данные из вспомогательной таблицы удаляются. Это повысило производительность, так как рекурсивная часть алгоритма стала работать только с необходимыми данными из вспомогательной таблицы. Все эти действия оформлены в виде хранимой процедуры и выполняются в рамках отдельной транзакции. Поскольку изменения, вносимые в пределах транзакции, не видны другим транзакциям до ее подтверждения, и данные, вносимые во вспомогательную таблицу удаляются до подтверждения транзакции, то метод можно использовать и в многопользовательских системах.

При извлечении иерархий целиком механизм рекурсии можно и не использовать, но в таком случае ус-

ложнится восстановление иерархической информации на стороне клиента БД. Кроме того, при извлечении либо анализе только части иерархии рекурсия все равно понадобится.

В дальнейшем будет рассматриваться именно указанная модификация метода.

При использовании метода вложенных множеств не предъявляется каких-либо особых требований к СУБД, но он является наиболее сложным в реализации. Приведем для примера хранимую процедуру, осуществляющую переподчинение узла (реализация для СУБД FireBird). Замечание: автор метода в своих статьях [2] не приводит процедуру переподчинения узла. Приведенный пример является разработкой автора данной статьи.

```
CREATE PROCEDURE RESUBJECT_TREE (
SUB_TREE_ROOT INTEGER/*переподчиняемый узел*/,
NEW_PARENT INTEGER/*его новый родитель*/)
AS
DECLARE VARIABLE SUB_TREE_LEFT INTEGER;
DECLARE VARIABLE SUB_TREE_RIGHT INTEGER;
DECLARE VARIABLE NEW_PARENT_RIGHT INTEGER;
DECLARE VARIABLE SUB_TREE_NODES_COUNT INTEGER;
DECLARE VARIABLE RIGHT_DIFF INTEGER;
begin
select left_, right_ from «Personnel»
where «Personnel».name = :sub_tree_root
into :sub_tree_left, :sub_tree_right;

select count(*) from «Personnel»
where left_ >= :sub_tree_left and right_ <= :sub_tree_right
into :sub_tree_nodes_count;
/*шаг 1*/
update «Personnel»
set left_ = left_ + 1000,
right_ = right_ + 1000
where left_ >= :sub_tree_left and right_ <= :sub_tree_right;
/*шаг 2*/
update «Personnel»
set right_ = right_ - 2 * :sub_tree_nodes_count
where right_ > :sub_tree_right and left_ < 1000;

update «Personnel»
set left_ = left_ - 2 * :sub_tree_nodes_count
where left_ > :sub_tree_right and left_ < 1000;

/*шаг 3*/
select right_ from «Personnel»
where «Personnel».name = :new_parent
into :new_parent_right;

update «Personnel»
set left_ = left_+2 * :sub_tree_nodes_count
where left_ > :new_parent_right and left_ < 1000;

update «Personnel»
set right_ = right_ + 2 * :sub_tree_nodes_count
where right_ >= :new_parent_right and right_ < 1000;
/*шаг 4*/
right_diff = sub_tree_right + 1001 - new_parent_right - 2 *
```

```

sub_tree_nodes_count;
update «Personnel»
set left_ = left_ — :right_diff,
right_ = right_ — :right_diff
where left_ >= :sub_tree_left + 1000 and right_ <=
:sub_tree_right + 1000;
end
    
```

На первом шаге переподчиняемая часть дерева исключается из рассмотрения путем прибавления к значениям LEFT_ и RIGHT_ большого числа (здесь 1000, число должно быть больше чем максимальное количество элементов в иерархии умноженное на 2). Вся иерархическая информация переподчиняемой части дерева сохраняется.

На втором шаге происходит удаление поддерева из старой позиции. Для этого обновляются значения LEFT_ и RIGHT_, которые в прямом порядке обхода дерева находились после корня перемещаемой части дерева.

На третьем шаге освобождается позиция для вставки, путем изменения соответствующих значений LEFT_ и RIGHT_.

На четвертом шаге происходит приведение значений LEFT_ и RIGHT_ элементов перемещаемого поддерева в соответствие с новой структурой иерархии. Делается это путем вычитания из LEFT_ и RIGHT_ специального значения, зависящего от количества элементов поддерева, а также его старой и новой позиций.

Дополнительную ясность можно внести при помощи схемы (рис. 3). Здесь элемент b ставится в подчинение элементу c.

Организация контроля целостности данных также возможна при помощи триггеров. Например, при удалении узла требуется удалять также все элементы, для которых выполняется условие $LEFT_{\text{элемента}} > LEFT_{\text{удаляемого}}$ и $RIGHT_{\text{элемента}} < RIGHT_{\text{удаляемого}}$. В случае органи-

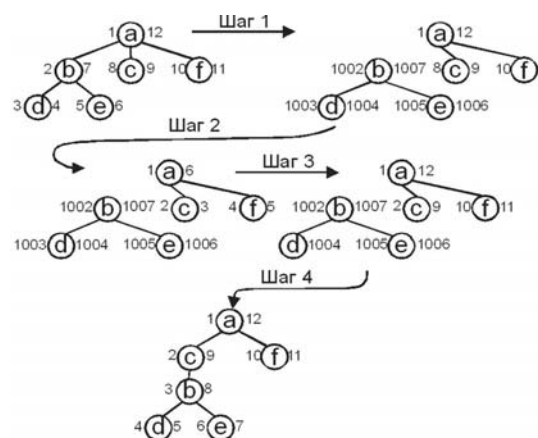


Рисунок 3 – Переподчинение узла при использовании метода вложенных множеств

зации такого механизма в виде триггеров, триггер будет вызываться и для каждого дочернего элемента, что является излишним, так как условие истинно для всех дочерних элементов независимо от уровня. Альтернативным подходом может быть запрет пользователю удалять данные из таблицы напрямую, а для удаления использовать хранимую процедуру.

Отметим возможность использование для столбцов LEFT_ и RIGHT_ вещественных типов данных. В этом случае вставка элемента и переподчинение узлов может осуществляться без обновления значений LEFT_ и RIGHT_ всего дерева. Но так как число значимых цифр при машинном представлении вещественных чисел является конечным, то в этом случае структура дерева не может быть произвольной. Чтобы сохранить преимущества использования вещественных типов данных необходимо вводить ограничения на максимальное количество дочерних элементов и максимальную глубину дерева. Для БД ИУА такие ограничения получаются слишком жесткими, поэтому описанная модификация метода в дальнейшем не рассматривается.

Строковый метод представления иерархической информации по сложности реализации занимает промежуточное положение между двумя предыдущими. Для использования этого метода необходимо, чтобы СУБД предоставляла развитые возможности работы со строками. Необходимость наличия таких возможностей продемонстрируем примером. Функция переподчинения элемента при использовании строкового метода (реализация для СУБД FireBird):

```

CREATE PROCEDURE RESUBJECT_TREE_V2 (
SUB_TREE_ROOT INTEGER /*переподчиняемый узел*/,
NEW_PARENT INTEGER /*его новый родитель*/ )
AS
DECLARE VARIABLE NODE_HR_POS VARCHAR(100);
DECLARE VARIABLE NODE_HR_POS_LEN VARCHAR(100);
DECLARE VARIABLE NEW_PARENT_HR_POS VARCHAR(100);
begin
select hr_pos from hierarchy where name = :SUB_TREE_ROOT
into :node_hr_pos;
select hr_pos from hierarchy where name = :NEW_PARENT
into :new_parent_hr_pos;

new_parent_hr_pos = rtrim(new_parent_hr_pos);
node_hr_pos = rtrim(node_hr_pos);
node_hr_pos_len = strlen(node_hr_pos);

update hierarchy
set hr_pos = :new_parent_hr_pos || '.' ||
substr(hr_pos, :node_hr_pos_len, strlen(hr_pos))
where hr_pos like :node_hr_pos || '%';
end
    
```

Здесь strlen – определяет длину строки, substr – извлекает подстроку, rtrim – удаляет завершающие пробелы, кроме того используется операция конкатенации строк.

Замечания относительно организации контроля целостности данных аналогичны таковым по предыдущему методу.

Из представленных трех методов этот наиболее требователен к памяти. Так для описания одной дуги в БД ИУА при использовании первого метода необходимо 12 байт, при использовании второго – 16, а для третьего метода – более 100. Для первичных ключей чаще всего используется четырехбайтовое целое, его максимальное значение: 2 147 483 647. В строковом представлении такое значение может занимать до 10 байт. Прибавим 1 байт на разделитель. Получится, что максимальный объем используемой для описания одной дуги памяти равен максимальной глубине иерархии, умноженной на 11. Объем используемой памяти можно сократить, заменив для столбца HR_POS тип char на varchar, но это зачастую приводит к снижению производительности, так как каждая строка таблицы будет занимать различный объем памяти, что замедляет извлечение информации.

Теперь перейдем к сравнению эффективности предложенных методов в реальных задачах. Был проведен эксперимент по тестированию предложенных методов. Далее приводится описание этого эксперимента и его результаты.

Для анализа эффективности была создана база данных, структура ее аналогична описанной в [2], но иерархическая информация представляется сразу тремя методами. БД была заполнена тестовой информацией, включающей 4000 иерархий. Размеры иерархий варьируются от 2 до 300 элементов, средний размер иерархий – 26 элементов, максимальная глубина – 8, максимальное количество дочерних элементов – 9 (среднее – 5). Тестовая БД реализована с использованием СУБД FireBird 1.5. Конфигурация тестовой платформы: Athlon 64 3000+, 512 М ОЗУ, HDD Samsung SpinPoint 2504c 7200 SATA II.

РЕЗУЛЬТАТЫ

Приведем результаты определения времени извлечения одной иерархии. Эти тесты выполнялись 100 раз, в таблице 1 указаны средние значения.

Данные извлекались в виде, удобном для восстановления иерархии на стороне клиента. То есть после информации об элементе иерархии сразу же должны следовать данные описывающие его дочерние элементы. Для метода хранения матрицы смежности это обеспечивается самим механизмом работы рекурсивной хранимой процедуры, а для других методов сортировки. В случае метода вложенных множеств сортировка выполняется по значению столбца LEFT_, а для строкового – HR_POS.

При извлечении иерархий небольшого размера (до 8 элементов) методы показывают сходную производи-

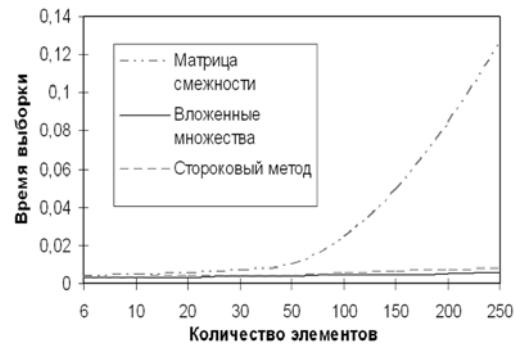


Рисунок 4 – График зависимости времени выборки от размера иерархии

тельность. Но чем больше иерархия, тем сильнее отстает метод хранения матрицы смежности. При размере иерархии 250 элементов он в 25 раз медленнее метода вложенных множеств. В тоже время строковый метод показывает лишь немногим меньшую эффективность. Из графика (рис. 4) видно, что метод хранения матрицы смежности непригоден для работы с большими иерархиями.

Таблица 1 – Время выборки иерархий в зависимости от количества элементов

Количество элементов в иерархии	Время выборки, мсек		
	Метод хранения матрицы смежности	Метод вложенных множеств	Строковый метод
6	3,83	3,27	3,28
10	4,29	3,29	3,29
20	5,34	3,43	3,49
50	10,05	3,71	3,82
100	24,35	4,28	4,63
150	48,89	4,94	5,46
200	83,93	5,46	6,18
250	126,43	5,87	6,81

С целью проверки эффективности обработки больших массивов иерархической информации на стороне сервера был выполнен еще один тест на выборку данных. Время работы хранимой процедуры, выполняющей извлечение 1000 иерархий представлены в следующей таблице 2.

Таблица 2 – Время выборки иерархий 1000 иерархий

	Время выборки, сек		
	Метод хранения матрицы смежности	Метод вложенных множеств	Строковый метод
При первом запуске	232,277	50,955	54,585
При последующих запусках	221,306	2,774	2,912

При первом запуске время работы значительно больше. Это может быть связано с тем, что в этот момент выполняются операции поиска информации на диске и ее чтение, а при последующих запусках система пользуется кэшируемыми данными. Обратим внимание на весьма незначительную разницу в производительности между методом вложенных множеств и строковым, несмотря на значительно больший объем памяти, используемый последним. Можно предположить, что это связано с тем, что операции линейного чтения выполняются очень быстро и основную часть времени работы занимают операции поиска.

Здесь же можно отметить влияние оптимизации на производительность метода хранения матрицы смежности. Так в случае использования оптимизированного метода время выполнения этого запроса составило 232,277 сек, а неоптимизированного 1081,784 сек.

Далее приведем результаты тестов переподчинения и удаления поддерев (табл. 3 и 4).

Таблица 3 – Время переподчинения 1000 поддеревьев

Время, сек		
Метод хранения матрицы смежности	Метод вложенных множеств	Строковый метод
2,704	19,317	7,068

Таблица 4 – Время удаления 1000 поддеревьев

Время, сек		
Метод хранения матрицы смежности	Метод вложенных множеств	Строковый метод
288,387	12,199	4,909

В задаче переподчинения наиболее быстрым оказался метод хранения матрицы смежности, т. к. в этом случае для выполнения операции необходима модификация лишь одного значения. В случае строкового метода должны быть изменены данные о положении в иерархии всего переподчиняемого поддерева, а в случае метода вложенных множеств требуется модифицировать информацию о значительной части узлов, входящих в иерархию в целом, часто требуется модифицировать информацию о большинстве узлов.

Метод хранения матрицы смежности показывает результаты значительно худшие, чем два других метода. Опять сказывается рекурсивный характер используемых алгоритмов. Строковый метод показывает наилучшую производительность, так как при удалении поддерева не требуется модификация информации о других элементах иерархии, в отличие от метода вложенных множеств.

ВЫВОДЫ

В статье были проанализированы возможные методы представления иерархической информации в реля-

ционных базах данных, на примере их использования в БД ИУА. Получил дальнейшее развитие метод вложенных множеств. Был проведен эксперимент по сравнительному тестированию указанных методов.

Если в проектируемой БД объем иерархической информации будет незначительным и СУБД поддерживает рекурсивные хранимые процедуры, то можно порекомендовать метод хранения матрицы смежности как наиболее просто реализуемый. В современных условиях, когда в основном не требуется жесткой экономии памяти, оправданным будет использование строкового метода, по соображениям простоты реализации и более высокой производительности в задачах модификации данных. Метод вложенных множеств можно порекомендовать использовать в случае, если требуется максимальная скорость выборки данных либо при необходимости экономии памяти.

В БД ИУА операции извлечения будут выполняться гораздо чаще операций модификации данных, поэтому наиболее важной для этой задачи является производительность при выборке иерархической информации. Но в тоже время не исключено и достаточно широкое использование различных операций модификации данных в задачах определения типовых ИУА. Поэтому для БД ИУА предпочтительным является строковый метод представления иерархической информации.

Направлениями дальнейших исследований могут являться анализ статистических характеристик производительности предложенных методов, влияние методов представления иерархической информации на эффективность БД ИУА в целом. Кроме того, при заполнении БД тестовой информацией учитывался тот факт, что число подчиненных не должно быть больше девяти, а оптимальный охват контролем – пять подчиненных [7]. Необходимо уточнить эту характеристику иерархий на основании результатов мониторинга ИУА.

ПЕРЕЧЕНЬ ССЫЛОК

1. Ковязин А. Мир InterBase: Архитектура, администрирование и разработка приложений баз данных в InterBase / Firebird / Yaffil / А. Ковязин, С. Востриков. – СПб.: Питер, 2005. – 496 с.
2. Демин М. К. Реляционные модели и базы данных многоуровневых информационно-управленческих архитектур // Информационные технологии в научных исследованиях и учебном процессе. Спецвыпуск: Сб. науч. тр. – Алчевск: ДонГТУ, 2005. – С. 38–50.
3. SQL Server Books online. Expanding Hierarchies [Electronic resource] / Microsoft. – Electronic data. – [USA], [199?]. – Access order: [http://msdn2.microsoft.com/en-us/library/aa172799\(SQL.80\).aspx](http://msdn2.microsoft.com/en-us/library/aa172799(SQL.80).aspx), free. – Title from screen.
4. Celko J. A Look at SQL Trees / Joe Celko // DBMS online [Electronic resource] / Miller Freeman, Inc. – Electronic mag. – [USA]: DBMS and Internet Systems, 1996. – March. – Access order: <http://www.dbmsmag.com/9603d06.html>, free. – Title from screen.
5. Ben-Gan I. Advanced Transact-SQL for SQL Server 2000 / Itzik Ben-Gan, Tom Moreau. – Berkeley: Apress, 2000. – 864 p.

6. Данич В. Н. Реляционные модели и базы данных информационно-управленческих архитектур / В. Н. Данич, Н. А. Митина, И. С. Хрестина // Вестник ВГУ, Луганск: Изд. ВГУ, 2003. – № 7 (65). – С. 119–131.
7. Байнев В. Ф. Основы теории организаций. Минск: БГУ, 2002. – 87 с.

обраний кращій метод, надані загальні рекомендації щодо застосування розглянутих методів у реляційних базах даних.

Надійшла 18.12.06
Після доробки 30.01.07

У статті аналізуються можливі методи представлення ієрархічної інформації в базах даних інформаційно-управлінських архітектур, наведені результати тестування ефективності цих методів у реальних задачах,

Possible methods of representation of the hierarchical information in databases of Information-Administrative Architectures are analyzed in the article. Results of testing of these methods in real tasks are listed. Best method was chosen. General recommendations of using considered methods in relational databases are given.

УДК 004.75

М. Б. Ильяшенко

АЛГОРИТМ НАХОЖДЕНИЯ ГРАФ-ПОДГРАФ ИЗОМОРФИЗМА ДЛЯ ВЗВЕЩЕННЫХ ГРАФОВ И ЕГО ПРИМЕНЕНИЕ

В работе предлагается подход к решению проблемы распределения задач в вычислительных сетях, обладающих стабильными характеристиками использования ресурсов. Приводится описание и исследование характеристик переборного алгоритма точного решения задачи. Предлагается способ построения систем планирования распределения ресурсов, обеспечивающий заданное качество обслуживания задач.

ВВЕДЕНИЕ

Задачи распределения нагрузки – это средство более эффективного использования ресурсов вычислительной системы за счет распределения решаемых задач среди вычислительных узлов системы. Существует множество подходов к распределению и балансировке нагрузки, применяемых в зависимости от требований, предъявляемых к вычислительной среде. Основными являются графо-аналитический подход [1–5], подход математического программирования [6, 7] и эвристический подход [8, 9]. Применяются также гибридные методы [10].

В данной работе рассматриваются вопросы применения графо-аналитических методов для размещения задач среди узлов вычислительной сети, обладающей достаточно стабильными характеристиками, такими как, средняя интенсивность обмена данными между узлами сети и средняя загрузка процессоров вычислительных узлов.

Построение систем планирования размещения задач в сети позволит производить статическую балансировку нагрузки на уровне планирования загрузки вычислительной сети в целом, в которой каждой пользовательской задаче выделяются (резервируются) ресурсы,

гарантирующие заданное качество обслуживания. Такие системы необходимы компаниям и провайдерам, предоставляющим услуги продажи или аренды вычислительных мощностей. Большинство динамических систем балансировки нагрузки ставит первоочередной целью ускорение вычислений и не используется в коммерческих целях обеспечения гарантированного уровня обслуживания пользовательских задач.

Ввиду того, что характеристики среды меняются редко и незначительно, нет необходимости постоянно реагировать на изменения интенсивности использования ресурсов и применения методов балансировки нагрузки в реальном времени таких, как применяются при параллельных вычислениях.

Основными особенностями динамических методов балансировки нагрузки:

- постоянно часть вычислительных ресурсов выделена на процесс, отслеживающий изменения в загрузке узлов вычислительной среды и планирующий перераспределение нагрузки между узлами сети;
- перемещение исполняемых процессов между узлами сети обычно производится на системном уровне, что не всегда возможно или затруднено, например, случае, если сам процесс работает на системном уровне (как в системах контроля и наблюдения процессы, обрабатывающие сигналы с датчиков);
- если распределенная вычислительная среда работает под управлением различных операционных систем, применение единой системы балансировки нагрузки часто бывает затруднено;
- техническая реализация систем балансировки нагрузки в реальном времени более сложна и их работа менее надежна, что является препятствием их примене-

© Ильяшенко М. Б., 2007

ния в системах с повышенными требованиями к надежности (таких как бортовые системы, системы контроля и управления в химической промышленности, ядерной энергетике);

– во многих системах отсутствуют средства обеспечения качества обслуживания пользовательских задач.

В случае редкого или однократного перераспределения нагрузки между вычислительными узлами среды имеется возможность применения точных, более эффективных алгоритмов распределения нагрузки, отсутствующая при динамической балансировке, так как подобные алгоритмы, как правило, имеют высокую вычислительную сложность. Но применение таких алгоритмов при однократной оптимизации может быть оправдано и экономически целесообразно ввиду последующей длительной эксплуатации сконфигурированной вычислительной среды.

Таким образом, существует область применения алгоритмов однократного распределения нагрузки за счет спланированного выделения ресурсов задачам в системах, характеризующихся относительной стабильностью параметров вычислительной среды.

1 СБОР СТАТИСТИЧЕСКОЙ ИНФОРМАЦИИ ОБ ИСПОЛЬЗОВАНИИ РЕСУРСОВ ВЫЧИСЛИТЕЛЬНОЙ СИСТЕМЫ

Для решения задачи спланированного выделения ресурсов в вычислительных сетях необходимо иметь информацию о конфигурации и характеристиках вычислительной среды и выполняемых задач. На основании этой информации строится два взвешенных графа: характеристический граф вычислительной среды и характеристический граф задачи. Предложенные далее алгоритмы оптимизации и распределения ресурсов будут работать с этими двумя характеристическими графами в качестве исходных данных.

При решении задач планирования нагрузки принимаются во внимание множество параметров характеризующих вычислительную среду, но наиболее важными из них являются: загрузка процессоров вычислительных узлов и интенсивность обмена данными между каждой парой узлов сети.

На основе собранных статистических данных будет сформирован взвешенный по ребрам и вершинам характеристический граф задачи.

Весы, приписанные вершинам графа, отражают интенсивность вычислений выполняемых каждым узлом системы. Ввиду того, что процессоры различных узлов могут иметь различную производительность, представляется удобным использовать единую шкалу для оценки этого значения, например MFLOPS или MIPS в зависимости от характеристик задачи.

Весы, приписанные ребрам графа, отражают среднюю (или в некоторых случаях пиковую) интенсив-

ность использования каналов связи между двумя узлами сети. Если между узлами не существует прямой связи, то предполагается, что выбран некий маршрут, по которому передается информация через промежуточные узлы и интенсивность транзитного трафика так же учитывается при оценке интенсивности обмена между узлами. Эту величину можно измерять в абсолютных значениях (Mbps).

Для получения статистической информации о задаче могут применяться как системные, так и встроенные в программу средства. Системными средствами операционных систем линейки Windows являются системные мониторы, в Unix-системах можно использовать утилиту SAR (System Activity Report).

Встроенные методы могут варьироваться в зависимости от инструментальных средств, используемых при написании программ, например, для MPI (Message Passing Interface), представляется целесообразным написание функций обмена данными, основанных на стандартных функциях MPI, снабженных средствами накопления статистической информации об интенсивности обмена данными между процессорами и вычислительной нагрузке узлов. Применение такого подхода позволит с минимальными изменениями в прикладных программах приспособить их для работы совместно с системой распределения ресурсов.

Для характеристического графа вычислительной среды веса, приписанные вершинам, отражают производительность узлов сети (в MFLOPS или MIPS), а веса, приписанные ребрам – пропускную способность каналов связи. Эти данные могут быть внесены разработчиком вручную (если размеры вычислительной среды не слишком велики, или велась ее поэтапная разработка) или собираться автоматически с помощью различных утилит.

Характеристический граф среды представляет собой набор ограничений «сверху», накладываемых на систему при распределении задач между узлами сети.

2 АЛГОРИТМ РАСПРЕДЕЛЕНИЯ ПРОЦЕССОВ МЕЖДУ УЗЛАМИ СЕТИ

Так как в работе не рассматривается влияние простоев при передаче данных по сети на загрузку и производительность узлов, то загруженность каналов связи и вычислительных узлов не должна превышать их максимальных значений.

На сегодняшний день системы, обладающие необходимым набором предпосылок, встречаются достаточно часто. Вычислительные возможности процессоров возросли, как и пропускная способность каналов передачи данных и часто встречается ситуация, когда избыточное количество компьютеров работает при неполной загрузке. При этом представляет интерес размещение распределенных задач, состоящих из отдельных неза-

висимых, но взаимодействующих между собой процессов, среди вычислительных узлов компьютерной сети.

Проблема размещения распределенных задач среди узлов компьютерной сети формируется следующим образом. Имеется компьютерная сеть, характеризующаяся производительностью вычислительных узлов и пропускной способностью каналов связи между ними. Необходимо разместить среди узлов этой сети распределенную задачу, состоящую из независимых процессов, с известными потребностями в вычислительной производительности процессора, взаимодействующие между собой согласно заданной топологии. При этом для каждой пары взаимодействующих процессов задана необходимая пропускная способность канала связи, которая должна быть зарезервирована для этого взаимодействия.

В терминах теории графов задача формулируется как задача нахождения изоморфной подстановки двух взвешенных по ребрам и вершинам графов.

На входе алгоритма имеется два графа: граф $G_A = (V_A, E_A, N_A, R_A)$ – характеристический граф вычислительной среды и граф $G_B = (V_B, E_B, N_B, R_B)$ – характеристический граф задачи, где N_i – веса, приписанные вершинам графов, а R_i – веса, приписанные ребрам графов. Решением задачи является подстановка $\phi: V_B \rightarrow V_A$, такая, что для каждой пары вершин $v_i, v_j \in V_B$, если $(v_i, v_j) \in E_B$, то $(\phi(v_i), \phi(v_j)) \in E_A$, $(n_i, n_j) \in N_B \leq (\phi(n_i), \phi(n_j)) \in N_A$ и для всех $r_i \in R_B$ выполняется $r_i \in R_B \leq \phi(r_i) \in R_A$.

Как следует из формулировки, задача очень близка к задаче граф-подграф изоморфизма для взвешенных графов. Поэтому предлагается использовать для ее точного решения алгоритм построенный на основе алгоритма нахождения граф-подграф изоморфизма [11–14]. Веса ребер и вершин используются как набор ограничений при поиске подстановки. Кроме того, в подстановке каждому ребру графа G_B должно соответствовать ребро графа G_A , но обратное не обязательно (т. е. в подстановке могут быть ребра графа G_A , которым нет соответствия среди ребер графа G_B).

Алгоритм нахождения подстановки удобно описывать в терминах поиска в пространстве состояний. Каждое состояние s процесса совмещения вершин соответствует частичной подстановке $\phi(s)$, содержащей лишь часть вершин полной подстановки. Каждому состоянию так же соответствуют подграфы $G_A(s)$ и $G_B(s)$, полученные из вершин графов G_A и G_B , вошедших в частичную подстановку $\phi(s)$, и ребер, соединяющих эти вершины. В дальнейшем обозначим через $\phi_A(s)$ и $\phi_B(s)$ проекции подстановки $\phi(s)$ на V_A и V_B .

Алгоритм состоит из предварительной и основной части. В предварительной части выполняются операции упорядочивания вершин графов и производятся

однократные, по ходу алгоритма, операций, призванные сократить область поиска основной, переборной части алгоритма.

2.1 Предварительная часть алгоритма

Основные действия, выполняемые в предварительной части алгоритма – сортировка вершин графов и формирование матрицы возможных совмещений.

Матрица возможных совмещений $M_{i,j}$ – это бинарная таблица размером $|V_A| \times |V_B|$. Каждому элементу таблицы соответствует пара вершин исходных графов $V_{A,i}$ и $V_{B,j}$. Значения матрицы формируются следующим образом: $M_{i,j} = 0$, если на основании предварительных проверок вершины $V_{A,i}$ и $V_{B,j}$ совместить нельзя; $M_{i,j} = 1$, в противном случае.

Смысл матрицы возможных совмещений в том, чтобы выполнить однократно в рамках предварительной части алгоритма все проверки, не основанные на информации, полученной в процессе совмещения вершин, тем самым, ускорить обработку соответствующих ограничений, сведя ее к одной операции сравнения.

В программе реализованы следующие предварительные проверки:

1. $M_{i,j} = 0$, если $|V_{A,i}| < |V_{B,j}|$, где $|V_{X,Y}|$ – степень вершины Y графа X ;
2. $M_{i,j} = 0$, если $|N_{A,i}| < |N_{B,j}|$, где $|N_{X,Y}|$ – вес приписанный вершине Y графа X ;
3. $M_{i,j} = 0$, если $|V_{A,i}^{in}| < |V_{B,j}^{in}|$, где $|V_{X,Y}^{in}|$ – число входящих ребер вершины Y графа X ;
4. $M_{i,j} = 0$, если $|\sum R_{A,i}^{in}| < |\sum R_{B,j}^{in}|$, где $|\sum R_{X,Y}^{in}|$ – сумма весов входящих ребер вершины Y графа X ;
5. $M_{i,j} = 0$, если $|V_{A,i}^{out}| < |V_{B,j}^{out}|$, где $|V_{X,Y}^{out}|$ – число исходящих ребер вершины Y графа X ;
6. $M_{i,j} = 0$, если $|\sum R_{A,i}^{out}| < |\sum R_{B,j}^{out}|$, где $|\sum R_{X,Y}^{out}|$ – сумма весов исходящих ребер вершины Y графа X ;
7. $M_{i,j} = 0$, если $|W_{A,i}^{vertex}| < |W_{B,j}^{vertex}|$, где $|W_{X,Y}^{vertex}|$ – число вершин в волновом разложении графа X начиная с вершины Y ;
8. $M_{i,j} = 0$, если $\sum_{l=1}^k W_{A,i,l}^{vertex} < \sum_{l=1}^k W_{B,j,l}^{vertex}$, $k = 1 \dots |W_{B,j}^{vertex}|$, где $|W_{X,Y,l}^{vertex}|$ – число вершин в l -й волне волнового разложения графа X начиная с вершины Y ;
9. $M_{i,j} = 0$, если $|W_{A,i}^{ribes}| < |W_{B,j}^{ribes}|$, где $|W_{X,Y}^{ribes}|$ – число ребер в волновом разложении графа X начиная с вершины Y ;

10. $M_{i,j} = 0$, если $\sum_{l=1}^k W_{A,i,l}^{ribes} < \sum_{l=1}^k W_{B,j,l}^{ribes}$, $k = 1 \dots \lfloor W_{B,j}^{ribes} \rfloor$, где $\lfloor W_{X,Y,l}^{ribes} \rfloor$ – число ребер в l -й волне волнового разложения графа X начиная с вершины Y .

Возможно использование и других критериев для оценки возможности совмещения вершин графов. Метод разработки таких критериев основан на волновом разложении графов, начиная с заданной вершины [15]. По мере распространения волны получают подграфы окружения вершин. Сравнивая параметры соответствующих подграфов окружения вершин графов, которые предполагается совмещать, делается вывод о потенциальной возможности или принципиальной невозможности такого совмещения. В приведенных критериях для этого использовались сумма вершин и ребер в подграфах окружения сравниваемых вершин для всех этапов распространения волны.

Сортировка вершин графов производится с целью ускорения нахождения изоморфной подстановки, в случае, если такая подстановка существует. В переборной части алгоритма переставляются только вершины большего графа, в то время, как порядок вершин меньшего графа не меняется. Порядок следования вершин меньшего графа определяется в предварительной части алгоритма.

Пусть $T_{B,t}$ – количество ребер инцидентных вершинам с меньшими номерами и $P_{B,i} = \sum_{j=1}^{|V_i|} M_{j,i}$ – суммарное количество вариантов совмещения вершины i графа G_B с вершинами графа G_A . Тогда порядок сортировки вершин графа G_B следующий: $V_{B,i} = V_{B,k}$, где $T_{B,k} = \min_{j=i+1}^{|V_B|} (T_{B,j})$.

Если $T_{B,i} = T_{B,j}$, то $V_{B,i} = V_{B,k}$, где $P_{B,k} = \min(P_{B,i}, P_{B,j})$.

Т. е. вершины графа G_B сортируются в порядке убывания количества связей с вершинами, имеющими меньшие номера, или в порядке убывания количества вариантов совмещения вершин, если количество связей одинаково. Такой порядок следования вершин обусловлен тем, что чем больше связей с совмещенными вершинами, тем жестче будет ограничивающее условие, включающее новую вершину, и, соответственно, меньше общее количество совмещений, которые необходимо перебрать.

2.2 Основная часть алгоритма

Основная часть алгоритма представляет собой последовательное наложение вершин с возвратом, описывать которое удобно в терминах метода поиска в пространстве состояний.

Вершины графа G_B остаются нетронутыми и каждой из них ставится в соответствие одна из вершин графа G_A . При этом проверяется допустимость такого совмещения. Если удастся найти соответствие всем вершинам графа G_B и при этом, выполнено условие задачи, то найденное состояние возвращается как искомая подстановка.

Пусть $T_{A,i}$ – количество связей вершины i графа G_A с вершинами $V_{A,j} \in \Phi_A(s)$, а $T_{B,i}$ – количество связей вершины i графа G_B с вершинами $V_{B,j} \in \Phi_B(s)$.

Начальному состоянию $\Phi(s)_0 = 0$ соответствует состояние, при котором не совмещено еще ни одной пары вершин.

Для получения i -го состояния для вершины $V_{B,i}$ ищется соответствие среди вершин $V_{A,j}$, таких что:

1. $M_{i,j} = 1$, т. е. вершины совместимы на основании предварительных проверок;
2. $T_{A,i} \geq T_{B,j}$;
3. $N_{A,i} \geq N_{B,i}$;
4. Для $k = 1 \dots i$, если $(v_i, v_k) \in E_A$, то $(\Phi(v_i), \Phi(v_k)) \in E_B$;
5. Для $k = 1 \dots i$, $(r_i, r_k) \in R_A \Rightarrow (\Phi(r_i), \Phi(r_k)) \in R_B$.

Если выполнены все условия, из которых 3, 4 и 5-е являются прямым следствием постановки задачи, то соответствующая пара вершин входит в частичную подстановку и формируется новое состояние $\Phi(s)_i$.

Перебор состояний производится методом поиска в глубину.

3 ИССЛЕДОВАНИЕ ХАРАКТЕРИСТИК АЛГОРИТМА

На сегодняшний день конфигурация большинства вычислительных сетей представляет собой одну из стандартных топологий (звезда, общая шина или кольцо), но быстрыми темпами идет развитие беспроводных компьютерных сетей, при которых становится возможным прямое взаимодействие компьютеров, используя более сложные топологии. Используя общую коммуникационную среду, сетевое взаимодействие может осуществляться между всеми компьютерами, расположенными на достаточно близком расстоянии, при котором уровень сигнала достаточен для поддержания устойчивого соединения. Такой тип взаимодействия даже в рамках одного офисного здания может приводить к формированию достаточно сложных топологий, размещение распределенных задач в которых представляет проблему.

Поэтому для исследования характеристик алгоритма использовались случайно сгенерированные графы с заданной плотностью реберного заполнения. Такие графы могут представлять в грубом приближении описанные

выше коммуникационные среды, формируемые сред- ствами беспроводной передачи данных.

Для исследования формировалась разнородная вы- числительная сеть, состоящая из компьютеров и сете- вых соединений различной производительности.

Вычислительная производительность узлов сети слу- чайным образом выбиралась от 200 до 3000, что про- порционально встречающимся процессорам от 200 МГц до 3000 МГц. Поскольку при качественной оценке про- изводительности алгоритма важны не численные пара- метры сети, а их пропорциональное соответствие ре- альным вычислительным сетям в произвольном мас- штабе, то пересчет производительности в MIPS или MFLOPS не производился. Также опущена разница в производительности, обусловленная различием в по- колениях компьютеров (наборах команд, размерности обрабатываемых данных и использованных методах конвейеризации для ускорения вычислений).

Веса ребер генерировались случайным образом в ди- апазоне от 10 до 100. Плавное изменение пропускной производительности сетевых соединений использован- ное для моделирования сети в экспериментах обуслов- лено тем, что пропускная способность беспроводных

каналов связи убывает с уменьшением интенсивности сигнала в зависимости от расстояния, в отличие от ка- бельных сетей, в которых пропускная способность под- держивается на некоем гарантированном уровне.

Задачи генерировались случайным образом из гра- фа сети, путем удаления части вершин и ребер графа сети и уменьшения приписанных весов. Параметрами являлись величины:

k_{nv} – коэффициент уменьшения числа вершин для графа задачи ($|V_B| = k_{nv}|V_A|$, где $|V_X|$ – число вершин графа X);

k_r – коэффициент числа ребер графа ($|E_B|' = k_r|E_B|$, где $|E_B|$ – число ребер в графа B);

k_w – коэффициент уменьшения весов графа ($N_{B,i}' = k_w N_{B,i}$ и $R_{B,i,j}' = k_w R_{B,i,j}$).

Исследования проводились на компьютере Athlon 1466 MHz, 512 MB RAM.

Некоторые характерные результаты исследований представлены на рис. 1–4.

Как видно из рисунков, алгоритм имеет в общем случае не полиномиальную вычислительную сложность, но при уменьшении числа вершин в графах или плот-

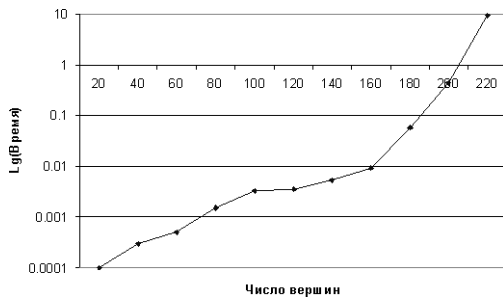


Рисунок 1 – Среднее время нахождения размещения задачи в сети при параметрах: $k_{nv} = 0, 2$, $k_r = 0, 5$, $k_w = 0, 2$, при плотности реберного заполнения характеристических графов сети 50 %

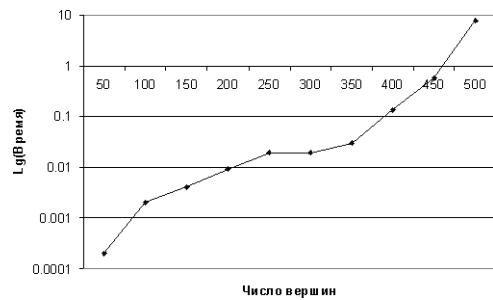


Рисунок 2 – Среднее отклонения времени нахождения размещения задачи в сети от среднего значения при параметрах: $k_{nv} = 0, 2$, $k_r = 0, 5$, $k_w = 0, 2$, при плотности реберного заполнения характеристических графов сети 50 %

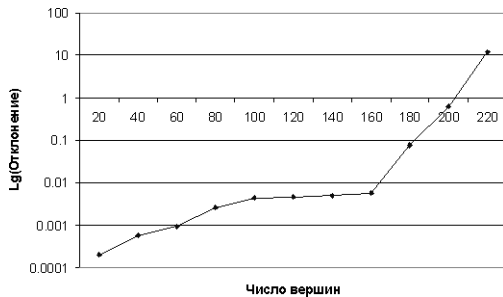


Рисунок 3 – Среднее время нахождения размещения задачи в сети при параметрах: $k_{nv} = 0, 1$, $k_r = 0, 5$, $k_w = 0, 2$, при плотности реберного заполнения характеристических графов сети 30 %

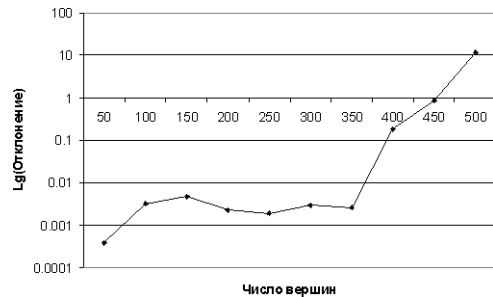


Рисунок 4 – Средние отклонения времени нахождения размещения задачи в сети от среднего значения при параметрах: $k_{nv} = 0, 1$, $k_r = 0, 5$, $k_w = 0, 2$, при плотности реберного заполнения характеристических графов сети 30 %

ности реберного заполнения, большую роль играют предварительные проверки, уменьшающие количество вариантов совмещения вершин в переборной части, поэтому алгоритм имеет значительно большую производительность.

Так для характеристических графов сети, имеющих плотность реберного заполнения 50 % и размер характеристического графа задачи 20 % от размера сети, алгоритм находит размещение менее чем за 1 секунду для графов с числом вершин до 200. А при плотности реберного заполнения 30 % и размере характеристического графа задачи 10 % от размера сети, в течение 1 секунды находится размещение для графов с числом вершин до 450.

В реальных компьютерных сетях плотность реберного заполнения намного меньше, чем 50 % и даже 30 %, таким образом, несмотря на NP-трудный характер задачи, алгоритм может успешно применяться в реальных компьютерных сетях для размещения распределенных задач.

ВЫВОДЫ

Рассмотренный метод статического планирования размещения распределенных задач среди узлов вычислительной сети, построенные на основе алгоритма нахождения граф-подграф изоморфизма для взвешенных графов, может быть основой построения системы управления размещением заданий. Потребности в таких системах существуют на сегодняшний день. Например, в высокопроизводительных кластерах, применяемых для параллельных вычислений при резервировании ресурсов под задачи с заранее известными потребностями в ресурсах; в сетях мобильных операторов, для выделения каналов связи под виртуальные частные сети с гарантированным уровнем сервиса; у hosting-провайдеров, предоставляющих в аренду часть ресурсов распределенного сервера. Рассмотренный подход отличается от методов балансировки нагрузки тем, что не ставится задача постоянного контроля и реконфигурации загрузки сети. Поиск и выделение ресурсов производится однократно при поступлении заявки, а балансировка нагрузки в сети выполняется за счет спланированного выделения ресурсов. Одновременно можно вести учет использования ресурсов сети, накапливая информацию о зарезервированных, выделенных и высвобождаемых ресурсах. Построение такой системы позволит более эффективно использовать имеющиеся вычислительные мощности за счет оптимального размещения задач, мониторинга и эффективного использования свободных мощностей.

В работе представлен метод решения задач размещения распределенных задач среди узлов вычислительных систем обладающих достаточно стабильными характеристиками. Предложены способы сбора статис-

тической информации, на основании которой принимается решение о размещении, и алгоритм, основанный на методе последовательного совмещения с возвратом.

Несмотря на NP-трудную природу задачи, исследования показали, что алгоритм может успешно применяться на современных сетях с числом узлов порядка нескольких сотен.

Дальнейшие усилия будут приложены к оптимизации предложенного алгоритма и построению на его основе системы управления размещением заданий в вычислительной сети.

ПЕРЕЧЕНЬ ССЫЛОК

1. Stone H. S., and Bokha, S. H. Control of Distributed Processes // Computer. – 1978. – July, vol. 11. – P. 97–106.
2. Chow T. C. K., and Abraham J. A. Load Balancing in Distributed Systems, // IEEE Transactions on Software Engineering. – 1982. – July, vol. SE-8(4). – P. 401–412.
3. C.-C. Shen and W.-H. Tsai. A graph matching approach to optimal task assignment in distributed computing system using a minimax criterion // IEEE Transactions on Computers. – 1985. – Mar., vol. C-34, No. 3. – P. 197–203.
4. K. Chow and B. Liu. On mapping signal processing algorithms to a heterogeneous multiprocessor system // 1991 International Conference on Acoustics, Speech, and Signal Processing. – 1991. – May, ICASSP 91, vol. 3. – P. 1585–1588.
5. M. Kafil and I. Ahmad. Optimal task assignment in heterogeneous distributed computing systems // IEEE Concurrency. – 1998. – July-Sept., vol. 6, No.3. – P. 42–51.
6. Chu W. W., Holloway L.J., Lan M.T., and Efe K. Task Allocation in Distributed Data Processing // Computer. – 1980. – Nov., vol. 13. – P. 57–69.
7. Ma P. R., et al, A Task Allocation Model for Distributed Computing Systems // IEEE Transactions on Computers. – 1982. – Jan., vol. C-31. – P. 41–47.
8. Efe K., Heuristic Models of Task Assignment Scheduling in Distributed Systems // Computer. – 1982. – June, vol. 15. – P. 50–56.
9. T. D. Braun, H. J. Siegel, N. Beck, L. Boloni, R. F. Freund, D. Hensgen, M. Maheswaran, A. I. Reuther, J. P. Robertson, M. D. Theys, and Bin Yao, A comparison of eleven static heuristics for mapping a class of independent tasks onto heterogeneous distributed computing systems // Journal of Parallel and Distributed Computing. – 2001. – June, vol. 61, No. 6, – P. 810–837.
10. Gao H., Schmidt A., Gupta A. and Luksch P., Load Balancing for Spatial-Grid-Based Parallel Numeric Simulations on Clusters of SMPs // Proceedings of the PDP2003 Conference. – February 5-7. – Genoa, Italy: IEEE Computer Society Publications. – 2003. – P. 75–82.
11. Ильяшенко М. Б. Разработка и исследование параллельного алгоритма проверки граф-подграф изоморфизма // Радиоэлектроника. Информатика. Управление. – 2006. – № 1. – С. 63–69.
12. Messmer B.T. and Bunke H. A decision tree approach to graph and subgraph isomorphism detection // Pattern Recognition. – 1999. – No. 32. – P. 1979–1998.
13. S. Umeyama. An Eigendecomposition Approach to Weighted Graph Matching Problems // IEEE Transactions on Pattern Analysis and Machine Intelligence. – 1988. – Sept., vol. 10, No. 5. – P. 695–703.
14. Ullman J. R. An algorithm for subgraph isomorphism // Journal of the Association for Computing Machinery. – Vol. 23, No. 1. – P. 31–42.
15. Пинчук В. П. Основанная на волновом разложении система инвариантов для простых графов и алгоритм распознавания изоморфности. – Киев, 1995. – Деп. в ГНТБ Украины 10.05.95, № 1002. – Ук95.

Надійшла 28.09.06
Після доробки 30.01.07

В роботі запропоновано підхід до рішення проблеми розподілення задач в обчислювальних мережах, що мають сталі характеристики використання ресурсів. Наводиться опис та результати дослідження характеристик переборного алгоритму точного вирішення задачі. Запропоновано спосіб побудови систем планування розподілення ресурсів, що забезпечує задану якість обслуговування задач.

This paper presents approach to solving problem of tasks distribution among computing environments with stable resources usage characteristics. Exact enumeration algorithm proposed for solving distribution problem. Algorithm's characteristics investigation also presents. Method for constructing resource distribution planning systems with fixed quality of service invented.

УДК 621.365.5

И. М. Килимник, Д. С. Ярымбаш

ПОВЫШЕНИЕ ЭФФЕКТИВНОСТИ ОБРАБОТКИ ИНФОРМАЦИИ ПРИ РЕГИСТРАЦИИ КОНТРОЛИРУЕМЫХ ПАРАМЕТРОВ В СИСТЕМАХ АВТОМАТИЧЕСКОГО УПРАВЛЕНИЯ

Предложена методика динамической обработки дискретных данных, которая сокращает объем входной информации, обеспечивает вычислительную эффективность и точность описания, позволяет повысить качество автоматического управления. Представлена ее реализация для температурных измерений в мундштуке прессы и измерений токов индукторов в системе автоматического управления режимами электрообогрева при прессовании заготовок подовых и доменных блоков.

ротермические установки, поэтому затраты на электроэнергию сопоставимы с затратами на сырье и оказывают существенное влияние на себестоимость готовой продукции. При производстве подовых и доменных блоков применяются массивные мундштуки для прессов с усилием от 3500 до 6300 кН, оборудованные системами индукционного электрообогрева с автоматическим регулированием температурных режимов прессования. Они, как объекты управления, характеризуются большой потребляемой мощностью, значительной долей времени работы в динамических режимах, частой сменой исходных данных, что необходимо учитывать при расчете управляющих воздействий.

ВВЕДЕНИЕ

На современном этапе развития промышленности Украины основные фонды промышленных предприятий требуют обновления и модернизации. Это обуславливает настоятельную потребность внедрения новых инновационных разработок в области технологии, нового технологического оборудования и средств автоматизации производственных процессов. Эффективность систем автоматизации технологических процессов зависит как от задач, решаемых этими системами, так и от качества информации, используемой для решения. Поэтому, особое внимание должно уделять разработке новых подходов, обеспечивающих повышение качества автоматического управления энергоемкими технологическими процессами, в том числе, на предприятиях металлургической и электродной промышленности.

Существенным фактором, который определяет качественный уровень заготовок подовых и доменных блоков, является соблюдение технологического температурного режима прессования. Нестабильность температурного поля мундштука и заготовки приводит к повышению удельной доли брака, перерасходу сырья и электроэнергии на дополнительные туры прессования заготовок.

Одним из основных требований, предъявляемых к электродной продукции, следует считать высокий уровень качества при минимальной себестоимости, обеспечивающий ее конкурентоспособность и спрос на внутреннем и внешнем рынках.

Для обеспечения стабильности температурных режимов на рабочей поверхности мундштука прессы повышаются требования к качеству управления при синтезе системы оптимального управления его электрообогревом. Анализ и синтез таких систем управления – достаточно сложный, так как технологическая структура «индукторы – мундштук – угольно-графитовая масса» должна рассматриваться в виде объекта с распределенными параметрами, который описывается дифференциальными уравнениями в частных производных. Это обуславливает потребность в новых подходах к обработке сигналов в системах автоматизации электрообогрева мундштука прессы при прессовании

Основным оборудованием на предприятиях графитового производства являются энергоемкие элект-

© Килимник И. М., Ярымбаш Д. С., 2007

заготовок подовых и доменных блоков, так как известные методы и алгоритмы [1, 2] не учитывают ряд особенностей, таких как дискретность данных, наличие скачков и разрывов при регистрации температур в мундштуке и токов в индукторах. Поэтому следует считать актуальной задачу повышения точности обработки динамических массивов исходной информации, адаптированной к автоматическому управлению температурными режимами прессования в режиме реального времени.

ЦЕЛЬ РАБОТЫ

Целью работы является разработка новых подходов, обеспечивающих повышение точности и эффективности при математической обработке динамической исходной информации по температурным и токовым измерениям в процессе прессования заготовок доменных и подовых блоков, улучшение качества автоматического управления электрообогревом мундштука пресса.

МАТЕРИАЛ И РЕЗУЛЬТАТЫ ИССЛЕДОВАНИЙ

В соответствии с работой [3] технологическая система «индукторы – мундштук – угольно-графитовая масса», как объект управления с распределенными параметрами, может быть представлена системой дифференциальных уравнений в частных производных:

$$\theta_{Fo(i)} = L\theta_i + B\theta_i + p_j(\theta_j, J_j), \quad i = 1, 2, 3, 4, \quad j = 1, 2, \quad (1)$$

где θ_i , J_j , Fo – относительные температуры, токовые нагрузки и критерий Фурье, которые определяются соотношениями вида:

$$\left\{ \begin{array}{l} \theta_i = \frac{T_i(\tau) - T_{\min_i}}{T_{\max_i} - T_{\min_i}} \\ J_j = \frac{I_j(\tau) - I_{\min_j}}{I_{\max_j} - I_{\min_j}} \\ Fo = \frac{\lambda(\theta_3) \cdot \tau}{c(\theta_3) \cdot \rho(\theta_3) \cdot S_{\text{п.с.м.}}^2} \end{array} \right. , \quad (2)$$

L – линейный дифференциальный оператор Лапласа; B – нелинейный оператор, учитывающий зависимость теплофизических свойств объекта от температуры; p_j – относительная удельная тепловая мощность; τ – время; $S_{\text{п.с.м.}}$ – площадь поперечного сечения мундштука; λ, c, ρ – коэффициент теплопроводности, удельная теплоемкость и плотность, соответственно; $T_i, T_{\max_i},$

T_{\min_i} – локальная, максимальная и минимальная температура; $I_j, I_{\max_j}, I_{\min_j}$ – действующие значения текущих, максимальных и минимальных токов в индукторах; индексы $i = 1, 2, 3, 4$ – соответствуют геометрическим областям индукторов захода и калибра, мундштука и угольно-графитовой массы и $j = 1, 2$ – относятся к областям индукторов калибра и захода, соответственно.

В работе [3] обоснован высокоэффективный подход, позволяющий преобразовать систему уравнений (1) с граничными условиями III и IV рода к системе обыкновенных дифференциальных уравнений:

$$\{F(\theta_{j,k}, \theta'_{j,k}, \theta_{j+1,k}, \theta_{j-1,k}, Fo) = 0, \quad 1 < k < n-1, \quad j = 1, 2, \quad (3)$$

$$\left\{ \begin{array}{l} \theta_{j,0 \cup n} = \theta_{\text{к.т}}(Fo)_{j,0 \cup n} \\ \theta'_{j,0 \cup n} = \theta'_{\text{к.т}}(Fo)_{j,0 \cup n} \end{array} \right. , \quad (4)$$

где $\theta_{\text{к.т}}(Fo)_{j,0 \cup n}$, $\theta'_{\text{к.т}}(Fo)_{j,0 \cup n}$ – значения относительных температур и скоростей их изменения в контрольных точках характерного поперечного сечения мундштука пресса под индуктором ($k = 0$) и над рабочей поверхностью калибра ($k = n$); n – число зон в характерных поперечных сечениях мундштука пресса.

Таким образом, с учетом (3), (4) для управления электрообогревом по входным дискретным целочисленным значениям температур контрольных точек при регистрации в динамических массивах данных, преобразованных в режиме реального времени в соответствующие значения критерия Фурье $\{Fo_i\}_{i \in I = \text{var}}$, необходимо обеспечить отображение множеств $\{\theta_{\text{к.т.},j,0 \cup n}\}_{i \in I = \text{var}}$ непрерывными функциями с множеством значений – $\{\theta_{\text{к.т}}(Fo_i)_{j,0 \cup n}\}_{i \in I = \text{var}}$ и непрерывными производными с множеством значений – $\{\theta'_{\text{к.т}}(Fo_i)_{j,0 \cup n}\}_{i \in I = \text{var}}$.

Для температурной коррекции токовых нагрузок индукторов калибра и захода в процессе прессования заготовок доменных и подовых блоков достаточно применить метод наименьших квадратов [1] к локальной кубической полиномиальной аппроксимации относительных токовых нагрузок и температур мундштука в характерных сечениях под индукторами по критерию Фурье (2):

$$J(Fo_i)_j = \sum_{l=0}^3 \gamma_{l,j} Fo_i^l \Big|_{i \in I = \text{var}},$$

$$\theta_{\text{н.к}}(Fo_i)_j = \sum_{p=0}^3 \eta_{l,j} Fo_i^l \Big|_{i \in I = \text{var}}, \quad j = 1, 2.$$

Так как функции $J = J(Fo)$ и $\theta_{и.к} = \theta_{и.к}(Fo)$ можно рассматривать, как параметрические, то следует считать определенными функции вида

$$J_j = J(\theta_{и.к})_j, \quad j = 1, 2.$$

При этом среднее значение относительной погрешности для относительной токовой нагрузки не превышает 1,1 % – для индуктора калибра и 2,1 % – для индуктора захода.

По методике, предложенной в работе [4], для отображения исходной информации температурных измерений заменим дискретные множества данных непрерывными функциями сплайн – интерполяции следующим образом:

$$\theta_{к.т}(Fo) = Pf(Fo) = \varphi_0(t)f(Fo_i) + \varphi_1(t)f(Fo_{i+1}) + \varphi_2(t)(Fo_{i+1} - Fo_i)\lambda_i + \varphi_3(t)(Fo_{i+1} - Fo_i)\lambda_{i+1}, \quad (5)$$

с базисными полиномами Эрмита [1]

$$\begin{aligned} \varphi_0(t) &= (2t + 1) \cdot (1 - t^2), \quad \varphi_2(t) = t \cdot (1 - t^2), \\ \varphi_1(t) &= (3 - 2 \cdot t) \cdot t^2, \quad \varphi_3(t) = (t - 1) \cdot t^2, \end{aligned}$$

где $t = \frac{Fo - Fo_i}{Fo_{i+1} - Fo_i}$ на отрезках $Fo_i \leq Fo \leq Fo_{i+1}$.

При определении неизвестных производных функции сплайн-интерполяции $\lambda_i = Pf'(Fo_i)$, $\lambda_{i+1} = Pf'(Fo_{i+1})$ потребуем, чтобы выполнялись следующие условия минимальной кривизны [1] в дифференциальной форме:

$$Pf''(Fo_i) = 0, \quad Pf''(Fo_{i+1}) = 0, \quad (6)$$

или в тождественной интегральной форме

$$\min J(\lambda) = \sum_{i=0}^{I-1} \int_{Fo_i}^{Fo_{i+1}} \left[Pf'' \left(\frac{Fo - Fo_i}{Fo_{i+1} - Fo_i} \right) \right]^2 dFo$$

совместно с условиями интерполяции [1]:

$$\begin{aligned} \delta_i(Pf(Fo))_{i,j,0 \cup n} &= \{\theta_{к.т,i,j,0 \cup n}\}_{i \in I = \text{var}}, \\ Fo_0 &< Fo_1 < \dots < Fo_I. \end{aligned} \quad (7)$$

Таким образом, сплайн-интерполяция базисными полиномами Эрмита (5)–(7) позволяет в режиме реального времени с высокой точностью преобразовывать массивы дискретных целочисленных значений температур в непрерывные функции (5) с непрерывными производными (рис. 1).

Однако, эта методика обладает рядом существенных недостатков, так как объем информации динамически возрастает пропорционально времени и такту ее ре-

гистрации. Следует также обратить внимание на низкую достоверность при определении значений скоростей температур в контрольных точках (рис. 1, б), обусловленную целочисленным представлением информации и необходимостью аппаратного ограничения ее разрядности (при регистрации температур в контрольных точках мундштука пресса – тремя знаками).

Поэтому требуется последующая аппроксимация непрерывной функции (5) на базе тригонометрических функций, что уже было успешно апробировано и реализовано авторами в работе [4] для автоматической системы регистрации токов и напряжений в короткой сети печей графитации.

С учетом того, что процессы электрообогрева и теплопередачи в мундштуке пресса при прессовании заготовок подовых и доменных блоков являются циклическими, то в ортонормированных базисных функциях

$$v(\varpi) = (\varpi^0, \sin(\varpi), \cos(\varpi), \dots, \sin(n\varpi), \cos(n\varpi)) \quad (8)$$

вводится критериальная функция аргумента Fo вида: $\varpi(\beta_1, \beta_2 Fo) = (\beta_1 \cdot Fo)^{\beta_2}$.

Аппроксимация непрерывной функции сплайн-интерполяции (5) на базе (8) имеет вид

$$Af(Fo) = \sum_{r=1}^R \alpha_r v_r(\beta_1, \beta_2, Fo) \quad (9)$$

с непрерывными производными порядка n

$$Af^{(n)}(Fo) = \sum_{r=1}^R \alpha_r v_r^{(n)}(\beta_1, \beta_2, Fo), \quad n \geq 1,$$

где неизвестные коэффициенты аппроксимации $\{\alpha_r\}_{r \in R}$ определяются из достаточного условия

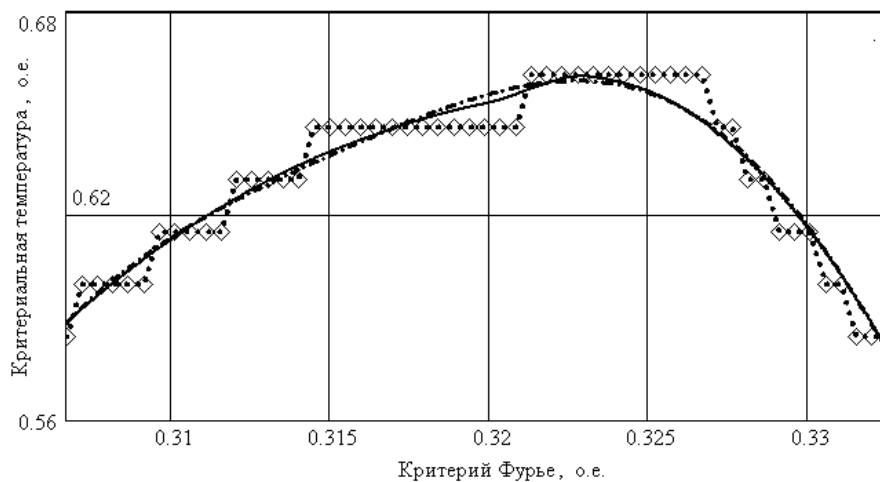
$$\min(\varphi(Af)) = \int_0^{Fo_I} \left[Pf(Fo) - \sum_{r=1}^R \alpha_r v_r(\beta_1, \beta_2, Fo) \right]^2 dFo$$

или соответствующих ему необходимых условий вида

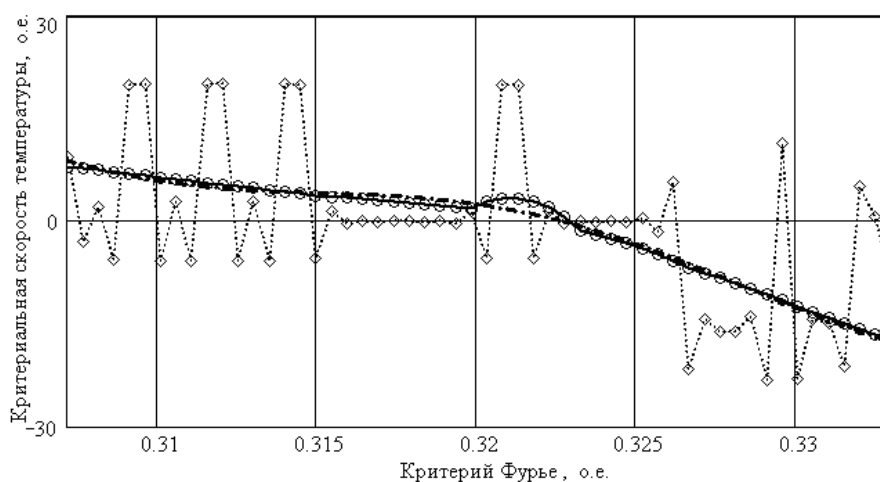
$$\begin{cases} \partial \varphi(Af) / \partial \alpha_r = 0, \quad r \in R \\ \partial \varphi(Af) / \partial \beta_{1,2} = 0 \end{cases}, \quad (10)$$

которые представляют собой систему линейных алгебраических уравнений относительно $\{\alpha_r\}_{r \in R}$ и нелинейных уравнений относительно $\{\beta_{1,2}\}$.

При различных турах прессования в результате обработки динамических данных регистрации температур над рабочей поверхностью мундштука было установ-



а)



б)

Рисунок 1 – Локальная обработка измерений критериальной температуры рабочей поверхности калибра мундштука:

а – критериальная температура, б – критериальная скорость изменения температуры

лено, что при аппроксимации их сплайн-интерполяции базисными полиномами Эрмита (5) достаточно ограничиться только третьими гармоническими составляющими относительно критериальной функции времени

$$\varpi(\beta_1, \beta_2 Fo) = \beta_1 Fo^{\beta_2}$$

в массиве базисных функций (8). Вектор коэффициентов $\{\beta_{1,2}\}$ с достаточной для управления точностью определяется уже на начальных шагах регистрации для каждого технологического цикла, включающего заполнение массного цилиндра, подпрессовку массы и прессование заготовок подовых или доменных блоков. Поэтому размерность вектора $\{\alpha_r\}_{r \in R}$ не

превышает $R = 5$, что более чем на порядок, уменьшает объем исходной информации, необходимой для расчета управляющих воздействий.

По результатам обработки температурных измерений при прессовании заготовок доменных и подовых блоков по методике (8), (10) и сплайн-аппроксимации [1, 5] (рис. 1) следует отметить, что оба подхода обеспечивают практически одинаковую точность, как по значениям контрольных температур (средние значения относительной погрешности – 0,83 % и 0,76 %), так и скоростей их изменения (среднее значение относительной невязки не превышает 2,8 %). При этом методика (8), (10) обеспечивает вычислительную эффективность в 5–7 раз выше по сравнению с методикой [5].

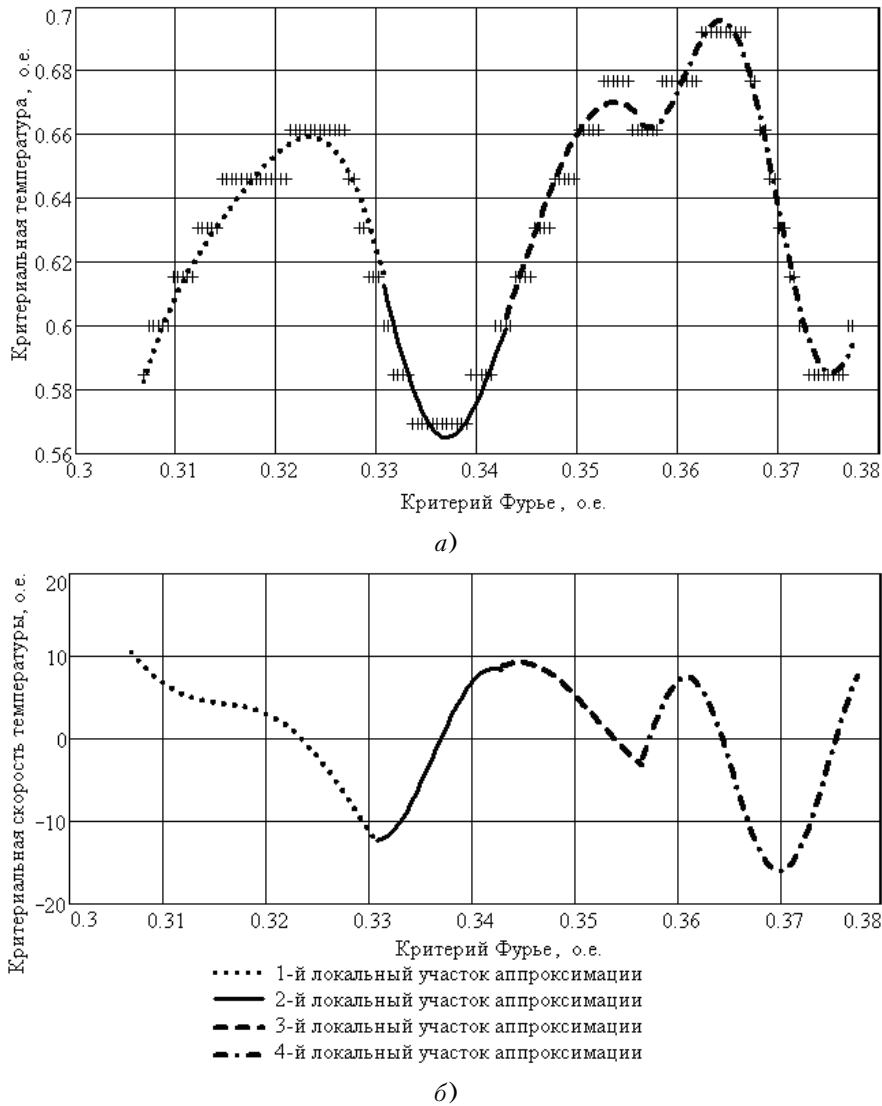


Рисунок 2 – Сшивка локальных участков обработки критериальной температуры рабочей поверхности калибра мундштука:

а – критериальная температура, б – критериальной скорости изменения температуры

В процессе прессования заготовок подовых и доменных блоков данные по измерениям температур и токов поступают последовательно через заданный такт, и их обработка происходит в режиме реального времени для определенного локального временного интервала (рис. 1, а). При этом возникает необходимость «сшивки» результатов по методике (8), (10) на смежных интервалах времени по критерию коллокации вида

$$\begin{aligned} & \left| Af(Fo_{i_1}) - Af(Fo_{i_2}) \right| \cdot \rho_{\square} + \\ & + \left| Af'(Fo_{i_1}) - Af'(Fo_{i_2}) \right| \cdot \rho_{T\square} \leq \varepsilon, \\ & \forall i_1 \in I_1 \cap \forall i_2 \in I_2 \cap \{Fo_{i_1} = Fo_{i_2}\}, \end{aligned} \quad (11)$$

где ρ_{\square} и $\rho_{T\square}$ – весовые коэффициенты, ε – точность коллокации.

Вычислительная эффективность «сшивки» локальной аппроксимации (8), (10) по критерию (11) соответствует методике [5], но при этом обеспечивается непрерывность, как температурной зависимости (рис. 2, а), так и ее скорости с установленной степенью точности ε (рис. 2, б).

ВЫВОДЫ

Предложена методика динамической обработки данных измерений для циклических процессов на основе комбинации методов сплайн-интерполяции полиномами Эрмита и аппроксимации на тригонометрическом базисе (8) с необходимыми условиями (10). Эта методика при одинаковом уровне вычислительной

погрешности с известными подходами позволяет более, чем на порядок, сократить объем входных данных для системы автоматического управления электрообогревом мундштука, повысить вычислительную эффективность обработки входных данных температурных измерений в 5–7 раз. Таким образом, можно существенно снизить требования, предъявляемые к аппаратным средствам автоматизированной системы управления по вычислительным ресурсам и объему памяти, повысить качество управления электрообогревом прессового оборудования, качество заготовок доменных и подовых блоков и сократить производственные затраты.

ПЕРЕЧЕНЬ ССЫЛОК

1. Математика и САПР. Кн. 2. Вычислительные методы. Геометрические методы / Под ред. Волкова Н. Г. – М.: Мир, 1989. – 260 с.
2. Шелевицкий І. В. Методи та засоби сплайн-технології обробки сигналів складної форми – Кривий Ріг: Європейський університет, 2002. – 304 с.
3. Андриенко П. Д., Ярымбаш Д. С. Особенности моделирования температурного состояния технологической системы как объекта управления // Электромашиностроения та електрообладнання. – Одесса, 2006. – № 66. – С. 291–293.
4. Ярымбаш С. Т., Ярымбаш Д. С., Тютюнник А. В. Особенности контроля электрических параметров, мощности и энергопотребления во время кампании графи-

тации в печах переменного тока // Вісник Кременчуцького державного політехнічного університету. – Кременчук, 2006. – Вип. 3 (38), частина 1. – С. 91–95.

5. Ярымбаш Д. С., Тютюнник О. В., Загруный О. Л. Моделирование температурных режимов электротехнологической системы «индукторы – мундштук» на подготовительном этапе тура прессования // Электротехника и электроэнергетика. – Запорожье: ЗНТУ, 2006. – № 1. – С. 56–60.

Надійшла 27.01.07

Запропонована методика динамічної обробки дискретних даних, яка скорочує об'єм вхідної інформації, забезпечує розрахункову ефективність і точність опису, дозволяє підвищити якість автоматичного керування. Представлена її реалізація для температурних вимірювань у мундштуці пресу і вимірювань струмів індукторів у системі автоматичного керування режимами електрообігріву при пресуванні заготовок подових та доменних блоків.

The method of dynamic analysis of the digital data is proposed. The input data compression, the numerical computation efficiency, the mathematical transformation accuracy and the high quality of automatic control. Implementation of this method for the temperature measuring data of pulp press die and current measuring in the heating inductor in the automatic-control system of the pulp pressing electric heating regime of the bottom and the blast-furnace graphite block production are presented.

УДК 004.7

Г. П. Коломеец

ИССЛЕДОВАНИЕ ВОЗМОЖНОСТЕЙ ПРОГРАММЫ «АНАЛИЗАТОР СЕТЕЙ ETHERNET» ДЛЯ МОДЕЛИРОВАНИЯ ПРОЦЕССА ВОЗНИКНОВЕНИЯ КОЛЛИЗИЙ В КОМПЬЮТЕРНЫХ СЕТЯХ ETHERNET

Выполнено исследование построенной с помощью программы «Анализатор сетей Ethernet» модели сети и ее сравнение с разработанной моделью возникновения коллизий в сегменте с двумя рабочими станциями. Выявлено хорошее соответствие для зависимости времени возникновения первой коллизии от длины сегмента, однако обнаружены некоторые ограничения программы.

ВВЕДЕНИЕ

Технология Ethernet используется в компьютерных сетях уже более 30 лет и сегодня является самой распространенной технологией локальных сетей. За это время скорость передачи данных посредством этой технологии возросла в 1000 раз. При этом для большин-

ства современных реализаций Ethernet остаются единственными формат пересылаемых кадров и используемый метод доступа к среде передачи – метод множественного доступа с обнаружением несущей и детектированием коллизий (CSMA/CD) [1]. Характерной чертой этого метода является возникновение коллизий (столкновение кадров разных рабочих станций) в совместно используемой несколькими рабочими станциями передающей среде и обработка этих коллизий. Поскольку от интенсивности коллизий зависит производительность сети, исследованию метода доступа CSMA/CD и процесса возникновения коллизий посвящено достаточно много работ [2, 3].

ПОСТАНОВКА ЗАДАЧИ

Благодаря значительному росту производительности персональных компьютеров в последнее время стали достаточно распространенными специализированные программы, позволяющие выполнять исследование процессов в компьютерных сетях путем моделирования. В то же время алгоритмы, на которых базируется такое программное обеспечение, как правило, неизвестны, поэтому нам представляется достаточно интересным тестирование подобных средств моделирования с помощью разработанной нами для простейшей конфигурации сети аналитической модели.

МОДЕЛЬ ВОЗНИКНОВЕНИЯ КОЛЛИЗИЙ В СЕГМЕНТЕ ETHERNET С ДВУМЯ РАБОЧИМИ СТАНЦИЯМИ

Рассмотрим сегмент сети Ethernet на коаксиальном кабеле (спецификация 10Base2) длиной L , состоящий из двух рабочих станций WS1 и WS2 (рис. 1). Здесь L – длина сегмента сети; $L_{1К}$, $L_{2К}$ – длины переданных в сеть станциями WS1 и WS2 частей кадров в момент столкновения.

Пусть в момент времени $t_1^0 = 0$ (начало отсчета времени) станция WS1 начинает передачу кадра, высылая первый бит информации в передающую среду. Через некоторое время t_2^0 (относительно начала отсчета), за которое высланный первой станцией кадр еще не достигает второй (поэтому вторая станция считает, что среда свободна), станция WS2 высылает свой кадр. Очевидно, что в данном случае произойдет коллизия, искажающая битовые сигналы обоих кадров. С учетом скорости распространения сигналов по кабелю $V = c \cdot k$, где c – скорость света в вакууме ($c = 3 \cdot 10^8$ м/с), k – зависящий от используемой среды передачи коэффициент (для коаксиального кабеля $k = 0,66$) [4], время возникновения коллизии t_K (относительно начала отсчета) находится в интервале:

$$0 < t_K < L/V. \quad (1)$$

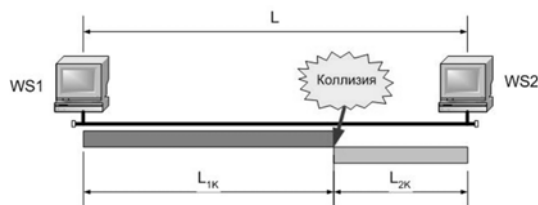


Рисунок 1 – Схема возникновения коллизии в совместно используемом сегменте сети Ethernet

В момент возникновения коллизии длина находящейся в кабеле части кадра станции WS1 равна:

$$L_{1К} = V \cdot t_K, \quad (2)$$

а длина находящейся в кабеле части кадра станции WS2 равна:

$$L_{2К} = V \cdot (t_K - t_2^0). \quad (3)$$

Сумма этих длин кадров равна длине кабеля, соединяющего рабочие станции $L = L_{1К} + L_{2К}$, откуда:

$$t_K = \frac{L}{2V} + \frac{t_2^0}{2}. \quad (4)$$

Из выражения (4) следует, что в случае, если обе станции начинают передачу своих кадров одновременно ($t_1^0 = t_2^0 = 0$), время возникновения коллизии равно:

$$t_K = \frac{L}{2V}, \quad (5)$$

т. е. коллизия происходит посередине сегмента. Если же станция WS2 начинает передачу непосредственно перед получением первого бита кадра, посланного WS1 ($t_2^0 \approx \frac{L}{V}$), то время возникновения коллизии:

$$t_K \approx \frac{L}{V}, \quad (6)$$

т. е. примерно равно времени распространения сигнала по всей длине сегмента.

Ситуация, когда коллизия возникает ближе к станции WS1, также может иметь место, однако она не соответствует сделанному предположению о том, что станция WS1 первой начинает передачу кадра.

На практике важными процессами являются регистрация и обработка коллизий. Первой коллизию регистрирует станция, расположенная ближе к месту столкновения кадров, в нашем случае это станция WS2. Момент времени регистрации t_2^{P0} представляет собой время, необходимое первому биту кадра WS1 для достижения места коллизии, плюс время, за которое первый бит кадра WS1 (уже искаженный) пройдет от места возникновения коллизии до WS2:

$$t_2^{P0} = t_K + (t_K - t_2^0) = 2t_K - t_2^0. \quad (7)$$

Очевидно, что станция WS1 узнает о коллизии через время:

$$t_1^{P0} = 2t_K. \quad (8)$$

Станция WS2, узнав о коллизии, в соответствии с алгоритмом CSMA/CD [1] высылает JAM-кадр длиной 32 бита, первый бит которого достигнет станции WS1 в момент времени t_1^P :

$$t_1^P = t_2^{P0} + \frac{L}{V} = 2t_K - t_2^0 + \frac{L}{V} = \frac{2L}{2V} + \frac{2t_2^0}{2} - t_2^0 + \frac{L}{V} = \frac{2L}{V}. \quad (9)$$

Аналогично, станция WS1, получив сигнал коллизии в момент времени t_1^{P0} , также вышлет JAM-кадр, первый бит которого достигнет станции WS2 в момент времени t_2^P :

$$t_2^P = t_1^{P0} + \frac{L}{V} = \frac{2L}{2V} + t_2^0 + \frac{L}{V} = \frac{2L}{V} + t_2^0. \quad (10)$$

После получения сигнала коллизии обе станции начинают отсчет времени задержки t_3 , которое согласно алгоритму CSMA/CD выбирается случайным образом [1]:

$$t_3 = \text{RAND}(0, 2^{\min(N, 10)}) \cdot 512\text{BT}, \quad (11)$$

где RAND – генератор случайных целых чисел в интервале от 0 до 2^N , N – номер попытки повторной высылки претерпевшего коллизии кадра (максимальное значение $N = 10$), BT – Bit Time – время передачи одного бита (для Ethernet $1 \text{ BT} = 1/10^7 \text{ бит/с} = 0,1 \text{ мкс}$).

Вероятность возникновения коллизий для повторно отправляемых кадров снижается с ростом номера попытки повторной передачи N (за счет расширения диапазона целых чисел, из которого производится случайный выбор).

Общая задержка времени $t_3^{\text{кол}}$ в результате возникновения коллизии складывается из времени регистрации коллизии, времени случайной задержки и равна для станции WS1:

$$t_{31}^{\text{кол}} = t_2^P + t_3 = \frac{2L}{V} + t_3, \quad (12)$$

а для станции WS2:

$$t_{32}^{\text{кол}} = t_2^P + t_3 = \frac{2L}{V} + t_2^0 + t_3. \quad (13)$$

Очевидно, что при повторной высылке кадров станциями также может возникнуть их коллизия. При этом вероятность возникновения коллизии зависит от выбранной задержки (для $N = 1$ (0, 51,2, 102,4) мкс)

и длины кадра. Поскольку для сегмента сети 10Base2 длиной 100 м время двойного оборота равно:

$$\frac{2L}{V} = \frac{2 \cdot 100 \text{ м}}{0,66 \cdot 3 \cdot 10^8 \text{ м/с}} = 1,01 \text{ мкс}, \quad (14)$$

то максимальная задержка для второй передачи кадра согласно выражению (13) определяется временем задержки повторной высылки и примерно равна 103 мкс. За это время по сети без коллизии может быть передан кадр длиной $103 \text{ мкс} / 1 \text{ BT} = 103 / 0,1 = 1030$ байт. В случае же передачи кадра большей длины (или/и уменьшения длины сегмента), должна произойти повторная коллизия.

ИССЛЕДОВАНИЕ ПРОЦЕССА МОДЕЛИРОВАНИЯ КОЛЛИЗИЙ В ПРОГРАММЕ «АНАЛИЗАТОР СЕТЕЙ ETHERNET»

Мы использовали приведенную выше аналитическую модель для исследования времен возникновения коллизий в модели простейшей сети, аналогичной показанной на рис. 1, построенной с помощью программы «Анализатор сетей Ethernet» (версия 1.1.0 Демо) [5].

Эта достаточно простая в использовании программа позволяет выполнять моделирование сетей Ethernet (а также Fast и Gigabit Ethernet) с заданием таких параметров, как длина сегмента, количество рабочих станций в сегменте, задаваемое для каждой станции начальное время высылки кадров, размер и количество высылаемых кадров, интервал времени между высылаемыми кадрами. В качестве регистрируемых параметров программа позволяет выбрать времена возникновения коллизий в сегменте, число переданных и потерянных кадров, среднюю длину буфера передачи, среднее время ожидания кадров в буфере и др.

Мы построили модель сегмента Ethernet с двумя рабочими станциями на его концах и установили для обеих станций параметры, приведенные на рис. 2. После этого запускался процесс моделирования и фиксировалось время возникновения в сегменте первой коллизии. Такая процедура повторялась для различных длин сегмента и в результате была получена зависимость модельного времени возникновения первой коллизии от длины сегмента (рис. 3).

Видно, что эта зависимость линейна, что согласуется с выражением (4). Тангенс угла наклона прямой равен $1/2V = 5,0067 \cdot 10^{-9} \text{ с/м}$, откуда можно получить $V = 1 \cdot 10^8 \text{ м/с}$ и коэффициент $k = 0,33$ для $V = c \cdot k$, где c – скорость света в вакууме ($c = 3 \cdot 10^8 \text{ м/с}$), k – зависящий от используемой среды передачи коэффициент (для коаксиального кабеля $k = 0,66$). И хотя по-

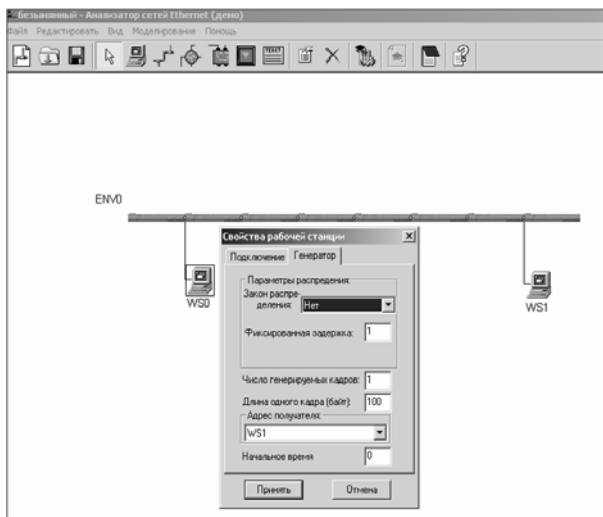


Рисунок 2 – Модель сегмента Ethernet с двумя рабочими станциями и используемые параметры для трафика рабочих станций

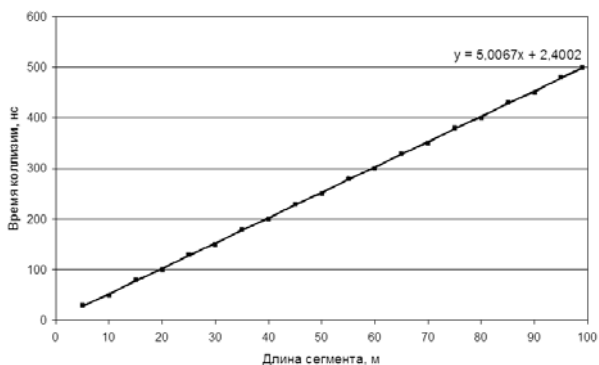


Рисунок 3 – Зависимость времени возникновения первой коллизии от длины сегмента

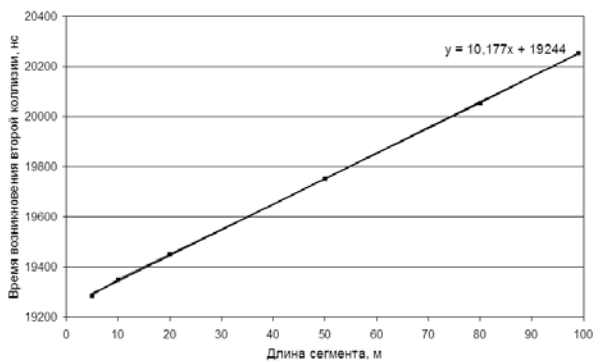


Рисунок 4 – Зависимость времени возникновения второй коллизии от длины сегмента

лученное значение коэффициента k в два раза ниже упоминаемого в [4] значения для коаксиального кабеля, линейный характер зависимости вполне соответствует выражению (4) и мы можем сделать вывод о соответствии результатов моделирования разработанной нами аналитической модели. Кроме того, очевидно, что используемые в программе для длин сегмента безразмерные величины (в диапазоне от 0 до 99) соответствуют метрам.

Следующим шагом была попытка исследования зависимости времени возникновения коллизий от времени задержки начала передачи кадра второй станцией t_2^0 и проверка полученных данных на соответствие выражению (4). Однако выполнить такое исследование не удалось, т. к. при установлении времени задержки t_2^0 в минимально возможное в программе значение (0,001 с) коллизия не происходила даже для максимально возможной в программе длины сегмента ($L = 99$ м) и максимально возможной длины передаваемых кадров (9999 байт). В то же время, как известно [1], для кадра Ethernet максимальной длины в 1522 байта время его передачи при скорости передачи данных 10 Мбит/с составляет $1522 \cdot 8 \cdot 100$ нс = 1,218 мс и при t_2^0 , меньшем или равном этому значению, должна возникать коллизия. Гипотетический кадр размером 9999 байт должен передаваться за $9999 \cdot 8 \cdot 100$ нс = 7,999 мс и, тем более, должна происходить коллизия при задержке начала передачи второй станцией в 1 мс.

После этого мы попытались определить зависимость времени возникновения второй коллизии от размера передаваемого кадра при одновременной высылке обеими станциями первых единичных кадров для фиксированных длин сегмента и зависимость времени возникновения второй коллизии от длины сегмента. Для длины сегмента $L = 5$ м были получены времена возникновения первой и второй коллизий при одновременной передаче по одному кадру размером 100 байт станциями WS1 и WS2. Фиксировались результаты (20 отсчетов) в случае возникновения более одной коллизии (были случаи возникновения только одной коллизии). Полученное среднее время возникновения второй коллизии не зависело от длины передаваемых кадров и составило $19,2884 \pm 0,0082$ мкс. Зависимость времени возникновения второй коллизии от длины сегмента была обнаружена (рис. 4).

Однако абсолютные значения разницы между временем возникновения первой и второй коллизии не превышали 20,5 мкс, в то же время, если повторная передача кадра начинается обеими станциями не одновременно, разница времен возникновения второй коллизии должна составлять не менее 51,2 мкс.

Полученные результаты свидетельствуют в пользу одновременной повторной высылки кадров обеими станциями после первой коллизии. При увеличении

длины передаваемых кадров картина наблюдалась аналогичная. Наносекундный разброс времен возникновения второй коллизии в этом случае, скорее всего, связан с разбросом задержек кадров в буферах рабочих станций.

ВЫВОДЫ

Основываясь на стандартных положениях метода доступа к разделяемой среде CSMA/CD, используемого в сетях Ethernet, разработана простая аналитическая модель, позволяющая рассчитать время возникновения коллизий и временные задержки при передаче кадров в сегменте сети Ethernet с двумя рабочими станциями. С использованием этой модели выполнено исследование программы моделирования «Анализатор сетей Ethernet». Полученная в программе зависимость времени возникновения первой коллизии от длины сегмента демонстрирует хорошее соответствие разработанной аналитической модели. Однако зависимость времени возникновения коллизий от задержки начала передачи кадра второй станцией в программе установить не удалось, что, очевидно, следует отнести к ограничениям используемых в ней алгоритмов. Также получена зависимость времени возникновения второй коллизии от длины сегмента для случая одновременной повторной отправки кадров.

УДК 025.4.03

М. В. Михайлова

ПОРІВНЯННЯ АЛГОРИТМІВ НЕЧІТКОГО ПОШУКУ В ТЕКСТАХ УКРАЇНСЬКОЮ МОВОЮ

У роботі проведено теоретичне та експериментальне порівняння ефективності роботи ряду методів нечіткого пошуку у текстових документах українською мовою. Показано, що найшвидшим та найефективнішим методом для пошуку слів з помилками є пошук за метрикою Левенштейна, найефективнішими для пошуку спільнокореневих слів – пошук за регулярним виразом та метод пошуку максимального спільного рядка з літерами, що розташовані послідовно, розроблений автором.

ВСТУП

Розвиток теорії і практики створення та використання баз даних приводить до поняття сховища даних – централізованої бази різномірних даних, отриманих з різних джерел, які часто зберігаються у текстовому

© Михайлова М. В., 2007

ПЕРЕЧЕНЬ ССЫЛОК

1. Таненбаум Э. Компьютерные сети. 4-е изд. – СПб.: Питер, 2003. – 992 с.
2. Абросимов Л. И. Анализ и проектирование вычислительных сетей: Учебное пособие. – М.: Издательство МЭИ, 2000. – 52 с.
3. Wang J. and Keshav S. Efficient and accurate Ethernet simulation // Technical Report TR99-749, Department of Computer Science, Cornell University (USA). – 1999. – Pp. 182–191.
4. Новиков Ю. В., Кондратенко С. В. Локальные сети: архитектура, алгоритмы, проектирование. – М.: ЭКОМ, 2000. – 312 с.
5. Анализатор сетей Ethernet – официальный сайт [Электронный ресурс] / Авт. Яшкин Д. – Электрон. дан. (7 файлов). – [Рязань?], [200–?]. – Режим доступа: <http://yd.ryazan.ru/index.html>, свободный. – Загл. с экрана. – Яз. рус.

Надійшла 22.09.06
Після доробки 1.12.06

Виконано дослідження побудованої за допомогою програми «Аналізатор мереж Ethernet» моделі мережі та її порівняння з розробленою моделлю виникнення колізій у сегменті із двома робочими станціями. Отримано добру відповідність для залежності часу виникнення першої колізії від довжини сегмента, однак виявлені деякі обмеження програми.

Research of the model constructed with the “Analyzer of Ethernet network” program and its comparison with the developed model of collisions occurrence in a segment with two workstations is carried out. Good correspondence for the dependence of the first collision occurrence time on segment length is detected, some restrictions of the program however are found out.

вигляді [1]. Основною мовою даних сховища звичайно є національна мова.

Одним з основних засобів створення запитів у сховищах даних є алгоритми пошуку в текстових документах необхідної інформації за рядом ознак: власне за змістом документа або за вторинними ознаками (дата, обсяг, розташування, природа). Найефективнішим засобом перевірки релевантності (відношення до теми) документу є пошук у ньому ключових слів.

За рядом причин точний пошук за ключовими словами (пошук у тексті безпосередньо заданих слів) не дає бажаних результатів, тому провідне місце серед алгоритмів пошуку за ключовими словами займають алгоритми нечіткого пошуку. Головна ідея нечіткого пошуку полягає у тому, що у тексті зустрічаються слова,

певною мірою схожі на шукані, вживання яких дозволяє віднести документ до шуканої теми.

Розробка і дослідження методів нечіткого пошуку ведеться багато десятиріч, але результати часто незастосовні до української мови через її істотні граматичні відмінності від мов, для яких проводилися дослідження. Інформація щодо застосування методів нечіткого пошуку до текстів українською мовою практично відсутня.

Предметом даної роботи є дослідження і порівняння ефективності ряду існуючих та розроблених автором методів нечіткого пошуку, застосованих до текстів українською мовою.

1 ПРЕДМЕТНА ОБЛАСТЬ. ПОСТАНОВКА ЗАДАЧІ

Перші згадки про задачу нечіткого пошуку відносяться до 60-х – 70-х років, коли вона з'явилася у ряді різних областей, головними з яких були тоді і залишаються сьогодні обчислювальна біологія (порівняння ланцюжків ДНК з можливими мутаціями), обробка сигналів (задачі перетворення сигналів із датчиків у зручну для зберігання та обробки форму, наприклад, розпізнавання мови, оптичне розпізнавання символів, виправлення помилок тощо) та текстовий інформаційний пошук. Сучасні застосування включають виявлення вірусів та злочинного проникнення в комп'ютерні системи, стиснення графічних файлів, розпізнавання образів, порівняння файлів та оновлення зображень на екрані.

Дана робота присвячена порівнянню методів розв'язання задачі текстового інформаційного пошуку – пошуку у великих підбірках текстових документів інформації, яка стосується заданого питання. Точного пошуку ключових слів недостатньо, оскільки стрімке зростання обсягів доступних текстів, яке спостерігається у теперішній час, супроводжується ростом кількості помилок, а точний пошук не знайде слово, яке було неправильно написано при внесенні до бази даних. Схожим застосуванням розглянутих методів може бути перевірка орфографії слів в ході їх набору.

З технічної точки зору задача нечіткого пошуку полягає у тому, щоб знайти серед запропонованого набору документів текст, в якому з'являється слово, яке схоже на задане користувачем слово або регулярний вираз (далі «зразок»). Текст розглядається як лінійна послідовність символів.

У даній роботі розглядається лише пошук в режимі прямого доступу, тобто алгоритми, які не вимагають попередньої обробки тексту для побудови індексу (допоміжної структури, яка будується заздалегідь і спрощує та пришвидшує пошук). Цей вибір пояснюється тим, що використання індексів має ряд недоліків, а всі

алгоритми з індексацією базуються на алгоритмах прямого доступу.

Якість нечіткого пошуку залежить від того, наскільки повно його алгоритм ураховує причини, з яких шукане слово може відрізнитися від свого аналогу в тексті. Коротко розглянемо ці причини.

1. Помилки у написанні слів, які з'явилися при внесенні тексту в базу даних:

- орфографічні помилки – 1,5–2,5 % тексту;
- помилки, зроблені при друкуванні – 1–3,2 % тексту;
- помилки, зроблені при оптичному розпізнаванні тексту – 7–16 % тексту;
- помилки, пов'язані з диктовкою тексту – досить специфічні, але зустрічаються у ряді застосувань, наприклад, у базах іншомовних прізвищ.

80 % друкарських помилок описуються моделлю Дамерау і виправляються лише одною вставкою, видаленням, заміною чи перестановкою. Орфографічні помилки та помилки розпізнавання звичайно також підходять під цю модель з більшими витратами [2].

2. Шукане слово присутнє у тексті у іншому відмінку, числі, часі або роді, ніж заданий варіант. Іноді з технічної точки зору це можна розглядати як помилку («шукана» – «шукане»), інші випадки потребують іншого підходу («малий» – «малесенький»).

Можливо два підходи до розв'язання цієї задачі: для кожного слова тексту визначити його основу і порівнювати з основою шуканого слова або за певними правилами будувати словоформи шуканого слова і шукати вже їх.

3. У тексті присутнє слово, спільнокореневе з шуканим, яке несе таке саме смислове навантаження («залежний» – «співузалежність»). Стеммери – це сімейство методів, призначених спеціально для пошуку спільнокореневих слів. Їх сутність полягає у тому, щоб за певним списком правил видалити зі слова префікси, суфікси та закінчення, отримати його корінь і порівняти його з коренем іншого слова.

4. У тексті присутній синонім шуканого слова, який не має з ним нічого спільного, крім смислу. Ця задача розв'язується використанням систем типу *WordNet*. *WordNet* – це лексична база даних, слова в якій організовані у синсети – набори синонімів, які визначають смисл поняття. Синсети, у свою чергу, організовані у деревоподібні структури, які описують інші семантичні відношення між поняттями – антонімію, меронімію (відношення «частина-ціле»), гіпонімію (відношення приналежності). Таким чином, при пошуку можна знайти не тільки синоніми заданого слова, а й інші слова, які несуть на собі схоже смислове навантаження.

Методи, розглянуті у даній роботі, враховують лише перші три з вказаних причин.

У ході роботи не було знайдено безкоштовних програмних продуктів, які б використовували алгоритми нечіткого пошуку в текстах українською мовою.

Пошук аналогічних продуктів для російської мови, яка є граматично найближчою до української, також не дав результатів. Тому всі досліджувані методи було реалізовано власноруч.

Робота орієнтована на максимальне можливе алгоритмічне пришвидшення методів, тобто зменшення часу роботи алгоритмів за рахунок зменшення кількості операцій. При цьому не використовуються засоби пришвидшення, які передбачають запам'ятовування вже оброблених слів, оскільки це вимагає обсягу пам'яті порядку розміру вихідних текстів. Не проводилася оптимізація пам'яті, якої вимагають методи, оскільки більшість методів використовує незначний обсяг пам'яті (порядку добутку довжини тексту на довжину шуканого слова).

2 ДОСЛІДЖУВАНІ МЕТОДИ

2.1 Метрика Левенштейна

Модель помилок Дамерау припускає існування 4 типів помилок, які можуть спотворити слово: вставка зайвого символу, видалення потрібного символу, заміна одного символу на інший і перестановка сусідніх символів. Кожен з типів помилок має свою вартість (яка може залежати від літер та від позиції помилки), отже, будь-який спосіб перетворення одного слова на інше має свою вартість. Операції заміни і перестановки можуть бути змодельовані як пара операцій вставка-видалення за більшу ціну, але у випадку друкованих текстів це недоцільно, оскільки вони є типовими помилками при друку.

На цій моделі помилок базується сімейство методів під назвою узагальнені відстані редагування. Відстанню між двома рядками $d(x, y)$ називається мінімальна вартість послідовності операцій, яка перетворює рядок x на рядок y (якщо така існує). Неважко переконатися, що якщо прямим і оберненим операціям призначені однакові вартості, відстань є метрикою у просторі слів. Залежно від обраних типів помилок та їх вартостей можна отримати ряд методів з істотно різними властивостями, наприклад, відстань Хемінга, «відстань подій» та узагальнена відстань Левенштейна [2].

Для реалізації обрано відстань Левенштейна з усіма чотирма операціями вартості як один із найтипівіших представників методів, призначених для пошуку слів з помилками. Алгоритм динамічного програмування складності $O(|A| \cdot |B|)$ [3]:

1. $D(0, a) = D(a, 0) = a$.
2. Для всіх $1 \leq a \leq |A|, 1 \leq b \leq |B|$

$$D(a, b) = \min \begin{cases} D(a, b-1) + 1, \\ D(a-1, b) + 1, \\ D(a-1, b-1) + \text{індикатор}(A[a] \neq B[b]), \\ (D(a-2, b-2) + 1) \cdot \text{індикатор}(A[a-1] = B[b], A[a] = B[b-1]). \end{cases}$$

3. Повернути $D(|A|, |B|)$.

Тут A і B – порівнювані слова, $A_{1\dots a}$ – префікс слова A довжини a , $|A|$ – довжина слова A , $D(a, b)$ – відстань Левенштейна між $A_{1\dots a}$ і $B_{1\dots b}$.

Всі ідеї, які покращують час роботи алгоритму у найгіршому випадку, полягають у тому, щоб обчислювати матрицю динамічного програмування не по рядках, а по діагоналях [2, 4, 5]. Можливо спланувати порядок обчислення елементів матриці так, щоб вирахувати лише «корисні» клітинки, тобто не дуже віддалені від «найкращого шляху» (набору клітинок, які описують послідовність дій, яка перетворить перше слово у друге за мінімальною кількістю операцій). Міркування, викладені в [5], узагальнені на випадок $|A| \geq |B|$. Їх наслідок – змінений порядок обчислення елементів матриці.

Розрахунок матриці починається з її осі і ведеться вздовж діагоналей. Коли з'являється елемент осі зі значенням $N+3$, слід повернутися і розрахувати діагоналі номер 1 вгору і вниз від осі, а потім перерахувати цей елемент. Коли з'явиться елемент осі зі значенням $N+2n+1$, слід повернутися назад, розрахувати діагоналі номер n вгору і вниз від осі і перерахувати цей елемент.

При цьому при розрахунку клітинок крайньої діагоналі (номер n від осі) слід використовувати значення лише двох клітинок матриці – попередньої клітинки діагоналі n і сусідньої клітинки діагоналі номер $n-1$. Сусідню клітинку діагоналі номер $n+1$ можна не використовувати з тих самих міркувань.

Шукана відстань $D(|A|, |B|) = D$ є максимальним елементом на осі матриці. Тоді $m = \left\lceil \frac{D-N}{2} \right\rceil$ – це номер останньої діагоналі від осі, яка була розрахована. Загальна кількість розрахованих діагоналей $N+2m+1 \cong D+1$; на кожній діагоналі не більше $|A|$ клітинок; отже, складність алгоритму $O(|A| \cdot (D(|A|, |B|) + 1))$.

Запропоновано два засоби зменшення часу роботи для задачі пошуку слів, які відрізняються від шуканого не більше ніж на k операцій (тут і далі k – параметр пошуку):

- 1) слова, для яких $||A| - |B|| > k$, можна не розглядати взагалі, оскільки для перетворення потрібно щонайменше $||A| - |B||$ операцій (засіб « $||A| - |B|| \leq k$ »);
- 2) якщо у ході обчислень знайдено елемент осі, значення якого перевищує k , то шукана відстань буде більше k , і обчислення можна припинити (засіб «недорозрахування»).

2.2 Пошук максимального спільного підрядка (МСП)

Метод пошуку МСП приймає мірою подібності двох слів довжину їх найдовшого спільного підрядка. Підрядок рядка отримуємо методом виключення з самого рядка нуля чи більше символів, не обов'язково сусідніх. Спільний підрядок двох рядків – це рядок, який є підрядком кожного з них. Найдовший з таких рядків називається максимальним спільним підрядком [6]. МСП використовується, коли істотним є порядок літер, а не відстані між ними. Метод МСП є другим з обраних типових методів пошуку слів з помилками.

Метод МСП є частковим випадком узагальненої відстані редагування з операціями вставки і видалення одиничної вартості [2]. Якщо d – це довжина МСП, а D – відстань редагування, то неважко переконатися, що $d = \frac{|A| + |B| - D}{2}$.

Для реалізації обрано алгоритм Вагнера і Фішера – алгоритм динамічного програмування складності $O(|A| \cdot |B|)$ [7, 8]:

1. $L(0, b) = L(a, 0) = 0$.
2. Для всіх $1 \leq a \leq |A|, 1 \leq b \leq |B|$

$$L(a, b) = \begin{cases} \max(L(a-1, b), L(a, b-1)), \\ \text{якщо } A[a] \neq B[b], \\ 1 + L(a-1, b-1), \text{ інакше.} \end{cases}$$

3. Повернути $L(|A|, |B|)$.

Запропоновано два засоби зменшення часу роботи для задачі пошуку слів, які мають із шуканим МСП довжини не менше за k :

- 1) слова, для яких $|A| < k$, можна не розглядати взагалі (засіб « $|A| \geq k$ »);
- 2) якщо на якомусь кроці обчислень отримано МСП довжини k , це слово вже можна вважати схожим на шукане, оскільки довжина МСП не спадає (засіб «недоразування»).

2.3 Пошук МСП з літерами, які розташовані посліпль

Класичний метод МСП є недосконалим для використання для пошуку спільнокореневих слів. Так, слова «*підпослідовність*» і «*піролюзіт*» мають довжину МСП 6, але не є спільнокореневими. Для задачі пошуку спільнокореневих слів істотно, що спільні літери таких слів розташовані поруч, утворюючи основу.

Пропонується модифікація методу МСП – метод МСП, який складається з літер, що розташовані посліпль, наприклад, «*підпослідовність*» і «*непослідовний*» (9). На жаль, цей метод буде реагувати і на слова, які мають довгу спільну кінцівку і короткі різні основи, як «*малесенький*» і «*гарнесенький*» (8). Алгоритм:

1. $L'(0, b) = L'(a, 0) = 0$.
2. Для всіх $1 \leq a \leq |A|, 1 \leq b \leq |B|$

$$L(a, b) = \begin{cases} 0, \text{ якщо } A[a] \neq B[b], \\ 1 + L'(a-1, b-1), \text{ інакше.} \end{cases}$$

3. Повернути $\max_{\substack{1 \leq a \leq |A|, \\ 1 \leq b \leq |B|}} L'(a, b)$.

Засоби пришвидшення роботи – аналогічно класичному методу МСП.

2.4 Метод n -грам

Метод n -грам узятो як представник методів, які не використовують модель заміни підрядків. Інші методи цього класу (відстань Хемінга, яка визначається як кількість позицій, в яких символи слів не співпадають; відстань інверсій, яка припускає інверсії підрядків; блочна відстань, яка дозволяє перестановки підрядків; відстань замін та інші) ще менше відповідають меті даної роботи, оскільки не моделюють типові причини спотворення слів.

Метод n -грам використовує в якості міри подібності двох слів кількість n -ок літер (n -грам), які зустрічаються в обох словах. Так, слова «агорафобія» і «графофобія» мають 5 спільних триграм («раф», «афо», «фоб», «обі», «бія»).

Одиницями порівняння можуть бути послідовності літер будь-якої довжини, від 1 (порівняння наборів літер, з яких складаються слова) до довжини коротшого слова (точне порівняння слів). Чим більша довжина n -грам, тим жорсткіші умови відбору подібних слів, і тим менша «нечіткість» метода.

Цьому методу притаманний один істотний недолік: одна помилка в середині слова спотворює n n -грам, що згубно впливає на результат порівняння.

Метод n -грам є єдиним з розглянутих методів, який може визнати як подібні пару слів, одне з яких складається з двох основ, а інше – з них же, але в іншому порядку. Прикладом таких пар є «графолог» і «логограф», «логотип» і «типологія». У [9] ця властивість методу n -грам вважається важливою перевагою. Але цей метод розроблявся і тестувався на текстах німецькою мовою. Її специфіка полягає в тому, що слова часто складаються з багатьох простих основ, як в українській – речення зі слів. У німецькій слова, які містять одні й ті самі основи в різному порядку, напевно мають близьке значення. В українській таких пар слів небагато, і слова пари мають різні значення, тому ця відмінність не має вирішального значення.

В цілому цей метод можна використовувати для пошуку спільнокореневих слів та слів форм заданого слова. Нижче буде показано, що на практиці цей метод еквівалентний методу МСП з літерами, які розташовані посліпль, але працює повільніше.

Було реалізовано метод, який порівнює окремі слова (а не слово з текстом, як у [9]) за кількістю спільних триграм, причому кожна триграма має вагу 1 (у загальному випадку вага триграми може залежати від її позиції у словах, від довжини слів та від самої триграми).

1. Для слів, які порівнюються, складається список триграм, які до них входять.

2. Отримані списки триграм порівнюються, і визначаються спільні триграми.

3. Вага спільних триграм сумується і порівнюється із заданим пороговим значенням. Якщо сумарна вага більша за порогове значення, то слова вважаються схожими.

Засоби пришвидшення роботи для задачі пошуку слів, які мають щонайменше k спільних триграм: слова, для яких $|A| < k + 2$, можна не розглядати взагалі, оскільки вони не містять k триграм (засіб « $|A| \geq k + 2$ »).

2.5 SoundEx

SoundEx є представником фонетичних методів, не дуже ефективних для розв'язання поставленої задачі, але цікавих для дослідження. *SoundEx* – це загальна назва сімейства алгоритмів, які перетворюють запис слова на певний код, який описує вимову слова. При пошуку порівнюються саме ці фонетичні коди. Запис слова може бути спотворений рядом факторів:

- слово пишеться не так, як вимовляється;
- людина, яка диктує слово, має дефекти мовлення, регіональні або особистісні особливості вимови чи іноземний акцент;
- людина, яка записує слово, недочула або неграмотно записала почуте.

Алгоритм застосовний лише в системах, які допускають високі рівні помилок першого і другого роду. Областю застосування цих алгоритмів є пошук власних назв (прізвищ або географічних назв) в базах даних значного обсягу. Перший алгоритм *SoundEx* використовувався Бюро перепису населення США для індексації прізвищ у базах даних. До сімейства *SoundEx* належать також коди *LEAA* (*Law Enforcement Alliance of America*), які використовуються в базах даних з запобігання злочинам, і коди *Cutter Tables*, якими користуються бібліотеки для кодування імен авторів книг [10].

Запропоновано оригінальний алгоритм перетворення запису слова на його фонетичну характеристику для української мови, побудований за тими ж принципами, що і англійський прототип. Алгоритм базується на фонетичній класифікації звуків української мови з урахуванням специфіки вимови йотованих букв. Букви 'ю', 'я', 'ї', 'є' є йотованими, 'у', 'е', 'і', 'и', 'а', 'о' – голосними, а решта букв – приголосними.

1. Заміняємо букви слова, крім першої, на спеціальні символи за правилом, яке схематично наведено у табл. 1.

Таблиця 1 – Правило заміни символів для методу *SoundEx*

'б', 'п', 'в', 'ф'	'0'
'г', 'г', 'х', 'к'	'1'
'ч', 'ц', 'щ'	'2'
'ж', 'ш'	'3'
'д', 'т'	'4'
'с', 'з'	'5'
'м', 'н'	'6'
'р'	'7'
'л'	'8'
'й'	'9'
голосна	'_'
'ь' або '' + йотована буква	'9_'
'ь' без йотованої букви	' '
йотована буква після голосної	'9_'
йотована буква після приголосної	'_'

2. З отриманого рядку видаляємо всі пари сусідніх однакових символів.

3. З отриманого рядку видаляємо всі символи '_'.

4. Для порівняння двох слів порівнюються їх фонетичні характеристики; якщо вони співпадають, то слова вважаються схожими.

У якості засобу пришвидшення роботи можна використати те, що слова з різними першими символами можна не порівнювати (засіб « $A[1] = B[1]$ »).

2.6 Використання регулярних виразів

Метод використання регулярних виразів є алгоритмом ідентифікації. У типовій задачі ідентифікації послідовностей задаються текст і множина послідовностей-образів (задана регулярним виразом); необхідно знайти входження образів у текст. Такі задачі часто виникають при роботі з текстами, тому методи їх розв'язання добре вивчені.

Існує ряд теорем, які використовуються при аналізі регулярних виразів та реалізації розпізнавачів образів. Одна з них стверджує, що для кожного регулярного виразу можна побудувати недетермінований скінченний автомат (НСА). НСА M – це п'ятірка (S, I, δ, s_0, F) , де:

- 1) S – скінченна множина станів пристрою керування (ПК);
- 2) I – скінченний алфавіт вхідних символів;
- 3) $\delta: S \times I \rightarrow 2^S$ – функція переходів;
- 4) $s_0 \in S$ – початковий стан ПК;
- 5) $F \subseteq S$ – множина кінцевих станів ПК.

Стан НСА задається парою (s, w) , де $s \in S$ – поточний стан ПК, а $w \in I^*$ – невикористана частина вхідного слова. Початковий стан НСА – (s_0, w_0) , де w_0 – слово, яке порівнюється з регулярним виразом.

Введемо бінарне відношення \mapsto на множині станів НСА. Якщо $s' \in \delta(s, a)$, то $\forall w \in I^* (s, aw) \mapsto (s', w)$. Транзитивне замикання цього відношення визначає регулярну мову, яку розпізнає цей автомат: якщо $\exists s \in F : (s_0, w) \mapsto^* (s, "")$, то w розпізнається автоматом.

За 1 крок автомат обробляє перший символ поточного слова і видаляє його; для кожного стану ПК робляться всі можливі переходи під впливом цього символу. Якщо множина станів, у якій може опинитися ПК автомату після завершення обробки слова, містить хоча б один кінцевий стан, то вхідне слово розпізнається автоматом. Якщо в ході обробки слова автомат потрапляє у ситуацію, коли з жодного стану ПК неможливо здійснити перехід, робота автомата припиняється, і слово вважається нерозпізнаним. [11]

Метод регулярних виразів обрано як найпростіший метод, який дозволяє використати знання користувача про відмінювання шуканого слова та його спільнокореневі слова, але, на відміну від стеммерів, не вимагає представлення в реалізації алгоритму загальних лінгвістичних знань з української мови. Метод регулярних виразів доцільно використовувати у випадках, коли користувач знає, як може видозмінюватися потрібне йому слово, і може описати це за допомогою регулярного виразу.

У нашому випадку для опису множини шуканих слів використовуються регулярні вирази спеціального вигляду: в їх запис можуть входити всі символи алфавіту I та спеціальні символи '*' та '?', які означають будь-яку послідовність літер алфавіту I (можливо, нульової довжини) та будь-яку одну літеру алфавіту I відповідно. Наведемо процедуру побудови НСА за таким регулярним виразом.

1. Побудова починається з порожнього автомату $M_0 = (\{s_0\}, I, \emptyset, s_0, \emptyset)$, який описується порожнім регулярним виразом «».

2. За кожний крок до регулярного виразу, яким описується поточний автомат $M_t = (\{s_0, \dots, s_m\}, I, \delta_t, s_0, \emptyset)$, дописується один символ заданого виразу a , і в залежності від того, який це символ, автомат змінюється за такими правилами (рис. 1):

1) $a \in I$: до станів ПК додається новий стан $S'_t = \{s_0, \dots, s_{m+1}\}$, функція переходів розширюється на перехід $(s_m, a) \rightarrow \{s_{m+1}\}$.

2) $a = '?'$: до станів ПК додається новий стан $S'_t = \{s_0, \dots, s_{m+1}\}$, функція переходів розширюється на переходи $\forall b \in I (s_m, b) \rightarrow \{s_{m+1}\}$.

3) $a = '*'$: функція переходів розширюється на переходи $\forall b \in I (s_m, b) \rightarrow \{s_m\}$.

3. Останній стан ПК приймається за кінцевий: $M_{fin} = (\{s_0, \dots, s_n\}, I, \delta, s_0, \{s_n\})$.

Для реалізації було створено алгоритм динамічного програмування складності $O(|A| \cdot |B|)$, який моделює роботу НСА, але не будує його:

1. $P(0, b) = \text{false}$ для $1 \leq b \leq |B|$.

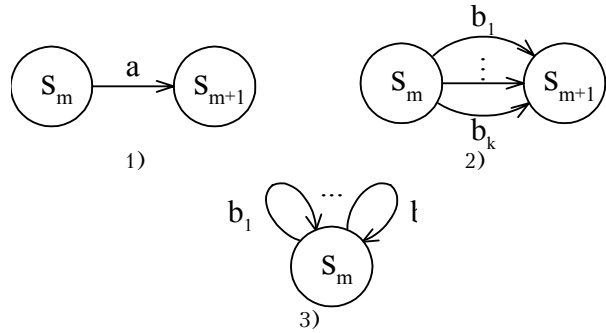


Рисунок 1 – Графічне представлення правил формування НСА

$P(0, 0) = \text{true}$, $P(1, 0) = \text{індикатор}(A[1]=*)$, $P(a, 0) = \text{false}$ для $1 < a \leq |A|$.

2. $P(a, b) = (P(a, b-1) \vee P(a-1, b)) \& (A[a] = *) \vee P(a-1, b-1) \& (A[1] \text{ in } \{*, ?, B|b\})$.

3. Повернути $P(|A|, |B|)$.

Засіб пришивдення часу роботи: слова, довжина яких менша за довжину регулярного виразу без символів '*', можна не розглядати (засіб $\langle |A| \geq l_{\min} \rangle$).

3 РЕЗУЛЬТАТИ ДОСЛІДЖЕННЯ

3.1 Критерії порівняння методів

Методи порівнювалися за двома критеріями: час роботи і якість результатів.

Час роботи включає в себе час, витрачений на пошук документів, що задовольняють умовам пошуку, та на розбиття кожного документу на слова, але цей «технічний» час, спільний для всіх методів, оцінено окремо, і він не впливає на порівняння. На заміри часу можуть впливати випадкові фактори (наприклад, переключення між процесами, характерне для Windows). За рахунок цього та з-за квантування часу похибка вимірювань становить близько 1 сек. Тому про перевагу одного метода над іншим можна говорити лише коли різниця часів їх роботи перевищує кілька секунд.

Якість результатів пошуку характеризується рівнями помилок першого і другого роду. Результатом пошуку *RES* назвемо множину слів, які алгоритм визначив як подібні до шуканого. Бажаними словами *WANT* назвемо множину слів, які користувач визначив би як подібні до шуканого і хотів би бачити в результаті пошуку. Помилка I роду має місце, коли в результаті пошуку відсутнє бажане слово, II роду – коли в результаті пошуку присутнє небажане слово.

Помилки I і II роду мають різну вагу. Прийняття рішення про те, чи відноситься документ до теми пошуку, залишається користувачу. Для виправлення помилки II роду йому достатньо продивитися результат пошуку і виключити зайві слова. Єдиний спосіб знайти

помилку I роду – це прочитати всі документи, що є неможливим у практично важливих випадках.

Множина бажаних слів залежить не лише від шуканого слова, а і від мети користувача. Так, якщо користувач задав слово «десять», невідомо, чи цікавить його чисельник «шістдесят». В загальному випадку користувача цікавлять словоформи слова, саме слово з помилками, і, можливо, спільнокореневі слова. В кожному окремому випадку якість результатів є окремим питанням. Для роботи визначимо бажані слова як спільнокореневі (включаючи словоформи). Випадки, у яких шукане слово написано з помилкою, враховуються окремо, оскільки експериментальні дані містять мало таких випадків.

Визначимо рівень помилки I роду як відношення кількості бажаних слів, які відсутні у результаті, до загальної кількості бажаних слів: $\epsilon_1 = \frac{|WANT \setminus RES|}{|WANT|} = 1 - \frac{|WANT \cap RES|}{|WANT|}$, рівень помилки II роду – як відношення кількості небажаних слів, які присутні у результаті, до загальної кількості слів у результаті: $\epsilon_2 = \frac{|RES \setminus WANT|}{|RES|} = 1 - \frac{|RES \cap WANT|}{|RES|}$. Рівні помилок аналізуються окремо, без утворення інтегрального показника, оскільки питання про визначення вагових коефіцієнтів потребує окремого дослідження.

3.2 Дані, на яких проводилися експерименти

Для експериментів було використано три типи даних.

Час роботи методів порівнювався на документах з предметної області «Психологія», поданих у текстовому форматі або підготованих редактором Word, обсягом 63 МБ. Шукалися слова, характерні для предметної області. Для проведення експериментів використовувався ПК з частотою процесора 1.8 ГГц.

Штучні документи створювалися для слів, обраних для ілюстрації певної переваги чи недоліку одного з методів. Для кожного слова виписувалися спільнокореневі слова та словоформи, а також слова, які не є спорідненими, але схожі на це слово і можуть бути обрані одним з методів. Результати експериментів було використано для якісного аналізу переваг і недоліків методів, але ці результати не є статистично достовірними і не можуть бути використані для обчислення помилок I і II роду.

У якості зразків природних документів використовувалися статті з предметних областей «Захист інформації» та «Психологія», щоб перевірити роботу методів на релевантних і нерелевантних документах. Слова, на яких проводилися експерименти (і для яких було створено штучні документи), було обрано як характерні для предметної області «Захист інформації».

Параметри методів, при яких проводився пошук, обиралися залежно від довжини шуканого слова та довжини його кореня.

3.3 Порівняння методів

Час роботи. Для кожного методу наведено графіки його часових характеристик, проведено порівняння ефективності засобів прискорення роботи та аналіз залежності часу роботи від параметрів пошуку.

Для всіх досліджуваних методів, які можуть мати різний ступінь нечіткості (залежний від параметра пошуку k , як описано вище), при використанні засобів прискорення, які його враховують, час роботи зростає зі зростанням ступеня нечіткості; для решти час роботи постійний.

Для метрики Левенштейна (рис. 2) найефективнішим засобом прискорення виявилася комбінація усіх запропонованих, але кожен з них окремо дає істотний ефект.

Час роботи методу без пришвидшення та без «діагонального порядку обчислення» не залежить від пара-

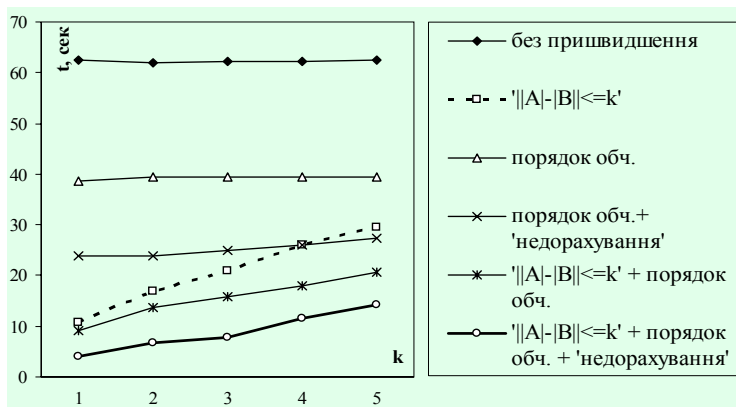


Рисунок 2 – Часові характеристики методу Левенштейна

метру пошуку. У цих методах для кожного слова тексту і шуканого слова спочатку вираховується відстань між ними, а потім ця відстань порівнюється з параметром пошуку.

Засіб $\|A\| - \|B\| \geq k$ покращує час роботи методу від 15 % до 50 %, час роботи зростає зі збільшенням k . Метод порівнює шукане слово та слова тексту, довжини яких задовольняє умові $|B| - k \leq |A| \leq |B| + k$; зі збільшенням k зростає кількість таких слів.

Діагональний порядок обчислення зменшує час роботи на 30 %, тобто в середньому третина клітинок матриці динамічного програмування виявляється «зайвою». «Недорахування» зменшує час роботи ще на 15 % (від вихідного часу) за рахунок того, що для слів, які не схожі на шукане, обчислення осі матриці не доводиться до кінця. Час роботи методу «з недорахуванням» незначно збільшується при збільшенні параметра k : збільшується кількість клітинок, які слід обчислити, щоб переконатися у несхожості слів.

Для класичного методу МСП (рис. 3) найефективнішим засобом пришвидшення є $|A| \geq k$. Застосований окремо, він зменшує час роботи на 50–60 %, причому час роботи зменшується із ростом k , оскільки кількість слів, для яких $|A| \geq k$, зменшується.

Засіб «недорахування» виявився неефективним. Застосований окремо, він зменшує час роботи методу всього на 8 %, а у комбінації з $|A| \geq k$ не пришвидшує роботу взагалі, оскільки він пришвидшує роботу для слів, схожих на шукане, яких значно менше, ніж несхожих. Для $k = 5$ застосування цього засобу навіть погіршило час роботи методу через те, що на кожному кроці методу виконується додаткова перевірка «чи це слово вже не схоже на шукане».

Через доведену неефективність «недорахування» цей засіб не використовувався для МСП з літерами, розташованими посліпль. МСП з літерами, розташованими посліпль, працює майже вдвічі швидше, ніж класичний МСП, оскільки його реалізація простіша (рис. 3). Оптимізація $|A| \geq k$ для нього так само ефективна і зменшує час роботи на 40–60 %.

При порівнянні МСП з літерами, розташованими посліпль, і метода триграм (рис. 4) бачимо, що без додаткового прискорення метод триграм працює швидше, оскільки метод триграм використовує розбиття шуканого слова на триграми, виконане один раз, і його складність ближча до $O(|A|)$. При використанні еквівалентних засобів прискорення $|A| \geq k$ для МСП та $|A| \geq k + 2$ для триграм МСП працює швидше.

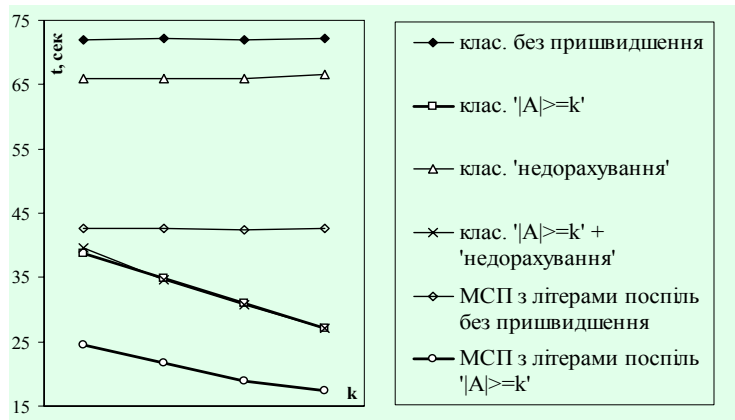


Рисунок 3 – Часові характеристики МСП

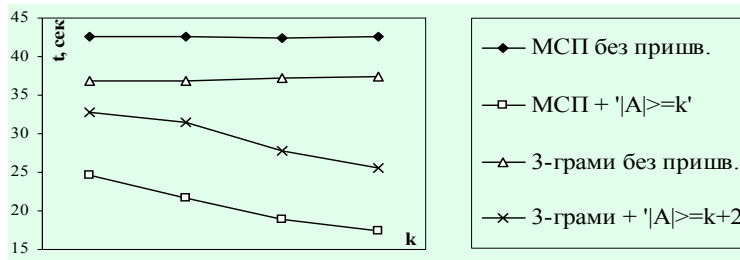


Рисунок 4 – Часові характеристики методу МСП з літерами, які розташовані посліпль, та методу триграм

Для методу регулярних виразів єдиний запропонований засіб прискорення $|A| \geq l_{\min}$ зменшує час роботи на 30 %. Для методу *SoundEx* засіб $A[1] = B[1]$ пришвидшує роботу на 90 %, оскільки його використання одразу відкидає близько $\frac{29}{30}$ усіх слів тексту.

Найповільнішим з методів є метод регулярних виразів (навіть із пришвидшенням $|A| \geq l_{\min}$), хоча шуканий регулярний вираз коротший за шукане слово. Найшвидшим з досліджуваних методів є *SoundEx* із пришвидшенням $A[1] = B[1]$, який через це найзручніший для побудови індексів. З методів власне для пошуку, найшвидшим є метрика Левенштейна із максимальним пришвидшенням.

Результати пошуку у штучних документах. Проаналізуємо переваги і недоліки методів, які проявляються на словах, спільнокореневих до шуканого.

Метрика Левенштейна знаходить досить багато словоформ шуканого слова, але лише дуже незначну частину спільнокореневих слів. Це пояснюється тим, що спільнокореневе слово зазвичай досить сильно відрізняється від шуканого, і метрика цієї відмінності перевищує рівень нечіткості пошуку, в той же час словоформа відрізняється від слова лише закінченням. Закінчення може бути довгим і не підпадати під це правило, але взагалі тенденція зберігається.

Якщо порогове значення методу МСП не менша за довжину незмінюваної основи слова, то обидва методи МСП знаходять всі спільнокореневі слова. При пошуку слова з чергуванням в корені класичний МСП знаходить більше слів, ніж МСП з літерами, які розташовані посліпль. Класичний МСП з параметром порядку довжини шуканого слова дозволяє знаходити слова з помилками.

Метод триграм дає такі самі результати, як і метод МСП з літерами, які розташовані посліпль, оскільки зазвичай спільна частина спільнокореневих слів складається із літер, які ідуть посліпль, і на них реагують обидва методи.

Метод *SoundEx* може знайти щонайбільше декілька словоформ шуканого слова за певних обмежень на нього. Цей досить специфічний метод справді підхо-

дить лише для індексації баз даних, які заповнювалися «на слух».

Метод регулярних виразів при правильному підборі шуканого регулярного виразу дозволяє знайти всі спільнокореневі слова. При пошуку слів без чергування в основі цей метод еквівалентний методу МСП з літерами, які розташовані посліпль, з певним числовим параметром, але в загальному випадку його можливості значно ширші. До того ж користувач має можливість обмежити множину бажаних слів. На жаль, цей метод не пристосований для пошуку слів з помилками: якщо помилка виникла у частині слова, яка жорстко задається регулярним виразом, то слово не буде знайдено.

Результати пошуку у природних документах. У таблиці 2 наведено приклад рівнів помилок I і II роду для результатів пошуку слова «керувати» у природних документах з області «Захист інформації».

Наведемо тут висновки, для яких недостатньо аналізу результатів пошуку у штучних документах.

Як і очікувалось, для всіх методів при збільшенні рівня нечіткості зростає обсяг результату пошуку; помилка другого роду зростає швидше, ніж зменшується помилка першого роду, що робить невиправданим збільшення рівня нечіткості за певною межею.

Метрика Левенштейна не дає очікуваних результатів при пошуку спільнокореневих слів. Рівень помилки першого роду становить 60–95 %, що робить цей метод практично незастосовним.

Класичний метод пошуку МСП для порівняно довгих слів дає дуже велику помилку другого роду. Крім того, для слів, які мають довге закінчення при короткій основі, він вважає схожими на них всі слова, які мають таке саме закінчення і перші літери основи. Прикладом такого випадку є пошук слова «керувати», закінчення якого є типовим для дієслів. Отже, цей метод не пристосований до задачі пошуку спільнокореневих слів. Кращі результати він дає при пошуку слів з помилками, але на відміну від метрики Левенштейна не враховує типову друкарську помилку перестановки сусідніх символів.

Таблиця 2 – Приклад результатів обробки експериментальних даних

$ W = 6$	Левенштг 2	Левенштг 3	Левенштг 4	МСП клас 6	МСП клас 7	МСП послп 5	МСП послп 6	Зграми 3	Зграми 4	SoundEx	Рег. вирази
$ R $	0	3	21	26	4	23	6	23	6	0	5
$ W \cap R $	0	1	3	1	0	1	1	1	1	0	5
ε_1	1	0.83	0.5	0.83	1	0.83	0.83	0.83	0.83	1	0.17
ε_2	0	0.67	0.86	0.96	1	0.96	0.83	0.96	0.83	0	0

При однаковому ступені нечіткості класичний метод МСП дає результат, обсяг якого на порядок більший за обсяг результату методу МСП з літерами, що розташовані поспіль. Метод пошуку МСП з літерами, які розташовані поспіль, краще відповідає задачі пошуку спільнокореневих слів, але залишається занадто чутливим до слів з таким самим закінченням, як у шуканого.

При детальнішому аналізі експериментів бачимо, що результати методу триграм відрізняються від результатів методу МСП з літерами, які розташовані поспіль, на 10–11 слів з усіх. Практично метод триграм еквівалентний методу МСП; відмінності не істотні, оскільки результати все одно значно гірші за результати методу регулярних виразів.

SoundEx має найвищі показники рівня помилки першого роду.

Єдиний метод, який стабільно дає низькі (не більше 17 %) рівні обох помилок, – це метод регулярних виразів. Це пояснюється тим, що цей метод дозволяє врахувати не лише написання шуканого слова, а і правила його відмінювання та утворення спільнокореневих слів. Чудовим прикладом цього є пошук за виразом «*кер?в*» для слова «керувати», який враховує зміну на «керований» та «керівний». Рівень помилки I роду 17 %, показаний на слові «керувати», насправді дещо завищений, оскільки в текстах зустрілася єдина нерозпізнана словоформа «керує» і не зустрілося ряду розпізнаних словоформ.

Взагалі високі рівні помилок другого роду пояснюються тим, що слова, для яких проводилися експерименти, є загальноживаними і типовими для орфографії української мови, тому для них існує багато слів зі схожим написанням. Методи нечіткого пошуку призначені для пошуку специфічних слів, які визначають тему документу (технічних та наукових термінів, прізвищ), тому при їх практичному застосуванні рівні помилок будуть менші за отримані у роботі.

ВИСНОВОК

У роботі досліджено ряд класичних та авторських методів нечіткого пошуку. Показано, що найшвидшим та найефективнішим методом пошуку слів з помилками є пошук за метрикою Левенштейна з рядом оптимізацій; найефективнішими для пошуку слів, спільнокореневих із шуканим, є пошук за регулярним виразом та (у спеціальних випадках) пошук МСП з літерами, що розташовані поспіль. Метод *SoundEx* взагалі не розв'язує поставлену задачу. Деякі із запропонованих алгоритмічних засобів пришвидшення пошуку виявилися достатньо ефективними. Питання про те, яким чином пришвидшити пошук та покращити його результати без включення до алгоритмів пошуку спеціальних знань щодо граматики використовуваної мови, потребує подальшого вивчення.

ПЕРЕЛІК ПОСИЛАНЬ

1. Спирли, Эрик. Корпоративные хранилища данных. Планирование, разработка, реализация. Т.1.: Пер. с англ. – М.: Издательский дом «Вильямс», 2001. – 400 с.
2. A Guided Tour to Approximate String Matching [Електронний ресурс] / Gonzalo Navarro. – ACM Press, 2001. – Режим доступу: <http://portal.acm.org/citation.cfm?doid=375360.375365>, вільний. – Загол. з екрану. – Англ.
3. Levenshtein Distance, in Three Flavors [Електронний ресурс] / Michael Gilleland. – Merriam Park Software, [199–?] – Режим доступу: <http://www.merriampark.com/ld.htm>, вільний. – Загол. з екрану. – Англ.
4. An Extension of Ukkonen's Enhanced Dynamic Programming ASM Algorithm [Електронний ресурс] / Hal Berghel, David Roach. – Електрон. стаття (5 ілюстрацій). – Publications of Hal Berghel, 1996. – Режим доступу: <http://berghel.net/publications/asm/asm.php>, вільний. – Загол. з екрану. – Англ.
5. Lazy Dynamic Programming can be Eager [Електронний ресурс] / L. Allison. – Monash University, 1992. – Режим доступу: <http://www.csse.monash.edu.au/~lloyd/tilde-Strings/Alignment/92.IPL.html>, вільний. – Загол. з екрану. – Англ.
6. Общие подпоследовательности. Дистанция [Електронний ресурс] / Илья Кантор, при поддержке компании Relmax. – [М.], [200–?] – Режим доступу: <http://algotist.manual.ru/search/lcs/index.php>, вільний. – Загол. з екрану. – Рос.
7. Метод динамического программирования Вагнера и Фишера [Електронний ресурс] / Илья Кантор, при поддержке компании Relmax. – [М.], [200–?] – Режим доступу: <http://algotist.manual.ru/search/lcs/vagner.php>, вільний. – Загол. з екрану. – Рос.
8. Нахождение наибольшей общей подпоследовательности [Електронний ресурс] / David Eppstein, пер. с англ. Илья Кантор. – [М.], [200–?] – Режим доступу: http://algotist.manual.ru/search/lcs/simple_lcs.php, вільний. – Загол. з екрану. – Рос.
9. Text-Detector. Fault-tolerant Retrieval Made Simple [Електронний ресурс] / Reinhard Rapp. – [Hildesheim], [199–?] – Режим доступу: www.cs.um.edu.mt/~mros/doc/Text-Detector.doc, вільний. – Загол. з екрану. – Англ.
10. 10. Understanding Classic SoundEx Algorithms [Електронний ресурс]: Search Names & Phrases Based on Phonetic Similarity / D. J. Repici. – [Creativyst], [200–?] – Режим доступу: <http://www.creativyst.com/Doc/Articles/SoundEx1/SoundEx1.htm>, вільний. – Загол. з екрану. – Англ.
11. Ахо А., Хопкрофт Дж., Ульман Дж. Построение и анализ вычислительных алгоритмов. – М.: Радио и связь, 1979. – 536 с.

Надійшла 19.09.2006

В работе проведено теоретическое и экспериментальное сравнение эффективности работы ряда методов нечеткого поиска в текстовых документах на украинском языке. Показано, что самым быстрым и самым эффективным методом поиска слов с ошибками является поиск по метрике Левенштейна, а самым эффективным для поиска однокоренных слов – авторский поиск по регулярному выражению и метод поиска максимальной общей подстроки с идущими подряд буквами.

This paper is devoted to theoretical and experimental comparison of efficiency of several main fuzzy search methods applied to the task of search in textual documents in Ukrainian. It is shown that the fastest and the most effective method of misspelled words search is a method that uses Levenshtein distance, while the most effective method of cognate words search is a regular expression matching search and a method searching the longest common substring with adjacent letters designed by the author.

М. Ю. Сенашова, А. Г. Рубцов, М. Г. Садовский

КИНЕТИЧЕСКАЯ МАШИНА КИРДИНА И ЗАДАЧА ВОССТАНОВЛЕНИЯ УТЕРЯННЫХ ДАННЫХ

Изложены предварительные результаты решения проблемы восстановления пропусков в символьных последовательностях при помощи кинетической машины Кирдина (далее КМК). КМК это мелкозернистый высокопараллельный распределенный вычислительный формализм. Предложен модифицированный имитатор КМК, специализированный для данной задачи. Имитатор работает с частью частотного словаря, содержащей лишь реагентоспособные слова. Решение задачи проиллюстрировано рядом примеров. Обсуждены дальнейшие направления применения КМК для задач восстановления утраченных данных.

ВВЕДЕНИЕ

Восстановление утраченных данных является актуальной проблемой как для фундаментальных, так и для прикладных областей науки. Результаты восстановления существенно зависят от способа, которым производилось восстановление, от характера данных и от самих утраченных данных. Очевидно, что некоторые утраченные данные не могут быть восстановлены на основе только той информации, которая содержится в известных частях последовательности. Любые данные, по крайней мере, теоретически, всегда можно рассматривать как символьную последовательность из некоторого конечного алфавита. В рамках настоящей статьи мы будем рассматривать лишь такие данные, что не уменьшает общности изложенных результатов.

Методы восстановления утраченных данных, как правило, опираются на знание различных свойств этих данных. Ключевым вопросом здесь является знание этих свойств. Если ограничиться собственно теми данными (фрагментами данных), которые доступны для исследования, то единственное их свойство, известное исследователю, заключается в знании частот отдельных малых фрагментов данных, имеющих в распоряжении исследователя. Под частотами понимается отношение количества одинаковых фрагментов заданной длины к общему числу фрагментов такой длины в тексте. Как правило, восстановление данных носит комплексный характер, и для этого привлекаются самые разные сведения о них, в том числе и информация, не содержащаяся непосредственно в этих данных (например, способы интерпретации или выявления значения для случая языковых данных). Рассмотрим проблему восстановления утраченных данных в максимально общей и строгой формулировке.

© Сенашова М. Ю., Рубцов А. Г., Садовский М. Г., 2007

1 ПОСТАНОВКА ЗАДАЧИ

В качестве данных будем рассматривать конечные символьные последовательности. Будем полагать, что алфавит, в котором записаны изучаемые последовательности, заранее известен. Отсутствие части такой последовательности будем рассматривать как потерю данных; будем также считать, что длина утраченной части последовательности известна, а сама утраченная часть (далее будем называть ее лакуной) представляет собой связный диапазон. В такой постановке проблема весьма актуальна для самых различных областей знания – от теории передачи данных до молекулярной биологии.

Решение указанной задачи требует развития алгоритмов заполнения лакун в символьных последовательностях. В нестрогом изложении принцип заполнения заключается в следующем: заполнять лакуны надо таким образом, чтобы последовательность, получающаяся после заполнения (восстановленная) была наиболее похожа на те части последовательности, которые имеются в распоряжении исследователя, однако требуется сделать это так, чтобы восстановленная последовательность несла в себе минимум дополнительной информации. Такой принцип имеет две формулировки: принцип максимума энтропии восстановленного частотного словаря (для случая восстановления по опорному частотному словарю) и минимума условной энтропии (для случая восстановления по полному частотному словарю). Перейдем теперь к строгим формулировкам и точным утверждениям.

2 КРИТЕРИИ ПРАВИЛЬНОСТИ ЗАПОЛНЕНИЯ ЛАКУН В ПОСЛЕДОВАТЕЛЬНОСТИ

Пусть N – длина всей (восстанавливаемой) последовательности:

$$N = N_1 + L + N_2,$$

где N_1 и N_2 – длины известных частей последовательности, а L – длина участка, который необходимо восстановить. Словом длины q будем называть любую связную последовательность этой длины, составленную

из символов алфавита Ω . Опорным частотным словарем W толщины q будем называть список всех слов этой длины, встречающихся в доступных исследователю частях, с указанием частот этих слов f_ω . Пополненным частотным словарем \bar{W} будем называть частотный словарь (толщины q), составленный по той последовательности, которая возникает в результате заполнения лакуны. Очевидно соотношение

$$1 \leq q \leq \min\{N_1, N_2\}.$$

Левой (соответственно, правой) опорой длины t , $0 \leq t \leq q - 1$ будем называть слово этой длины, расположенное сразу слева (соответственно, сразу справа) от лакуны. Тем самым, в зависимости от величины t , восстанавливаемая часть имеет длину $L + 2t$, при том условии, что первые и последние t символов являются фиксированными. Отметим, что для построения заполнения длина лакуны L должна быть больше длины слова q .

Построить заполнение лакуны означает построить из слов ω_i длины q цепочку

$$\omega_1, \omega_2, \omega_3, \dots, \omega_{L+2t-q}, \omega_{L+2t-q+1} \quad (1)$$

длиной $L + 2t$, у которого первые t и последние t символов заданы, а для каждой пары соседних слов выполнялось условие

$$\omega_j = i_1 \bar{\omega}; \quad \bar{\omega} i_q = \omega_{j+1},$$

т. е. два соседних слова пересекаются по общему подслову $\bar{\omega}$ длины $q - 1$, а первое слово в этой цепочке (соответственно, последнее) начиналось (соответственно, заканчивалось) левой опорой α_l (соответственно, правой опорой α_r). Если существует цепочка вида (1), составленная из слов опорного словаря и она единственна, то задача построения заполнения решена. Если существует несколько цепочек вида (1), составленных из слов опорного словаря, то среди всех возможных следует выбрать ту цепочку, которая обеспечивает максимум энтропии

$$\tilde{S} = -\sum_{\omega} \{\tilde{f}_\omega \cdot \ln \tilde{f}_\omega\} \quad (2)$$

пополненного частотного словаря $\bar{W}(q)$, где \tilde{f}_ω – частота слов, вычисленная по тексту, полученному в результате заполнения лакуны. Вне зависимости от того, существует или нет заполнение лакуны словами из опорного словаря, можно строить заполнение лакуны всеми возможными в данном алфавите словами. Очевидно, что такое заполнение существует всегда и не единственно.

Возможен иной подход к выбору наилучшего заполнения (1). Он реализует принцип максимального подобия заполнения имеющимися частям последовательности. В этой ситуации выбрать следует такое заполнение (1), для которого условная энтропия

$$\tilde{S} = \sum_{\omega} \tilde{f}_\omega \cdot \ln \left(\frac{f_\omega}{\tilde{f}_\omega} \right) \quad (3)$$

опорного частотного словаря $\bar{W}(q)$ относительно пополненного $\bar{W}(q)$ достигнет минимума. Здесь f_ω – частота слов в опорном частотном словаре, а \tilde{f}_ω – частота слов в словаре, построенном по всей последовательности, полученной в результате заполнения лакуны (пополненного); понятно, что $f_\omega = 0$ для некоторых ω , в то время, как $\tilde{f}_\omega > 0$.

Эти два принципа выбора наилучшего заполнения не являются взаимоисключающими либо конкурирующими. Каждый из них может быть применен независимо в одной и той же ситуации, каждый из них позволит выбрать то или иное заполнение.

**3 КИНЕТИЧЕСКАЯ МАШИНА КИРДИНА
В ЗАДАЧЕ ВОССТАНОВЛЕНИЯ
УТЕРЯННЫХ ДАННЫХ И ЕЕ
ПОСЛЕДОВАТЕЛЬНЫЙ ИМИТАТОР**

Прежде, чем излагать процедуру заполнения утраченных данных, кратко рассмотрим понятие кинетической машины Кирдина (далее КМК). КМК [1–3] является идеальным мелкозернистым параллельным вычислительным формализмом, по степени абстракции подобным машине Тьюринга. Известно, что КМК алгоритмически полна [3], то есть любой мыслимый алгоритм представим в ее терминах. Отличительной чертой КМК является мелкозернистый параллелизм. Обрабатываемой единицей является ансамбль слов M из алфавита Ω , который отождествляется с функцией F_M с конечным носителем на Ω^* , принимающей неотрицательные целые значения $F_M: \Omega^* \rightarrow N \cup \{0\}$. Значение $F_M(\omega)$ интерпретируется как число экземпляров слова ω в ансамбле M .

Обработка ансамблей в КМК состоит в совокупности элементарных событий, происходящих недетерминировано и параллельно. Элементарное событие $S: M \rightarrow M'$ состоит в том, что из ансамбля M изымается ансамбль K^- и добавляется ансамбль K^+ , т. е. $F_{M'} = F_M - F_{K^-} + F_{K^+}$. Ансамбли K^- и K^+ однозначно задаются правилами или командами, которые объединяются в программу. Команды могут быть только трех видов: распад $uvw \rightarrow uf + gw$, синтез $ik + dk \rightarrow usw$ и прямая замена $uvw \rightarrow usw$. Здесь u и w – произволь-

ные слова, а слова v, f, g, k, d и s суть некоторые фиксированные слова из Ω^* (подробности см. в [1–6]).

Неформально КМК можно представить себе как аналог химического реактора, в котором происходят реакции [7]. Имеется химический реактор идеального смешения, в котором плавают слова. В реактор добавляются правила-катализаторы; одни из них, взаимодействуя со словами, способствуют их распаду, другие, встречая пару подходящих слов, способствуют их синтезу, и, наконец, третьи заменяют в словах некоторые подцепочки.

Перейдем теперь к изложению способа заполнения лакуны в последовательности в терминах КМК. Для этого приведем программы для КМК, описывающие алгоритмы построения частотного словаря и заполнения лакун. Пусть у нас имеется текст T , по которому требуется составить частотный словарь W_q . Программа для КМК, реализующая этот процесс, состоит из одной команды и выглядит следующим образом:

$$uf^1v^{q-1}g^1w \rightarrow uf^1v^{q-1} + v^{q-1}g^1w,$$

где в качестве M нужно взять ансамбль, состоящий из одного слова T . Верхним индексом обозначено количество букв в слове. После того, как машина остановится, ансамбль M будет содержать все слова длины q , встречающиеся в исходном тексте, с учетом их кратности.

Программа, реализующая процесс заполнения лакуны в терминах КМК выглядит следующим образом:

$$\begin{aligned} \alpha_l + \alpha_l v^{q-t} &\rightarrow \alpha_l v^{q-t*}; \\ v^{q-t} \alpha_r + \alpha_r &\rightarrow *v^{q-t} \alpha_r, \end{aligned} \quad (4a)$$

$$\begin{aligned} uv^{q-1}* + v^{q-1}v^1 &\rightarrow uv^{q-1}v^1*; \\ v^1v^{q-1} + *v^{q-1}w &\rightarrow *v^1v^{q-1}w, \end{aligned} \quad (4b)$$

$$u^* + *v \rightarrow uv. \quad (5)$$

Первые две строчки программы осуществляют инициализацию затравок, т. е. обеспечивают взаимодействие правой (левой, соответственно) опоры длины с подходящим словом длины q . Третья и четвертая строчки осуществляют рост затравок, в ходе которого, собственно, и строится заполнение лакуны. И, наконец, последняя строка склеивает левые и правые части. Символ $< * >$ не принадлежит алфавиту Ω и используется в программе, чтобы помечать те слова, которые успешно прошли стадию инициации.

Исходным ансамблем для программы, реализующей заполнение лакуны, является некоторое количество копий «затравок» (левая (α_l) и правая (α_r) опоры) и не-

которое количество копий словарей, полученных применением к опоре предыдущей программы.

В КМК элементарные события происходят недетерминировано и параллельно. Тем не менее, программа построена так, что вначале в ансамбле просто нет таких слов, к которым бы могли примениться три последние команды. Таким образом, всю программу работы КМК по заполнению лакуны в последовательности можно представить состоящей из трех последовательных этапов.

Первый этап – инициализации затравок – заключается в присоединении к левой (α_l) и правой (α_r) опорам слов из W .

Второй этап – рост заполнений – начинается, как только в ансамбле будет достаточное количество слов, помеченных символом $< * >$. Следует отметить, что число слов, помеченных символом $< * >$, зависит от структуры фрагментов, имеющихся в распоряжении исследователя; может случиться так, что опорный словарь W не будет содержать ни одного слова, которое обеспечивало бы инициализацию. Последовательное и многократное применение команд этого этапа позволяет получить слова вида $\alpha_l u^*$ и $*v \alpha_r$, длина которых составляет приблизительно $L/2$.

На **третьем этапе** остается «склеить» слова вида $\alpha_l u^*$ и $*v \alpha_r$. Полученные таким образом слова будут составлять финальный словарь для данной программы и вышеописанного исходного ансамбля, так как ни одна из команд программы будет к ним уже не применима. Поскольку КМК функционирует не детерминированно и параллельно, в этом финальном ансамбле будут слова разной длины. Самое короткое из них может иметь длину $q+2$. Теперь нам нужно выбрать из ансамбля слова длины $L+2t$ и исследовать полученные заполнения лакуны в соответствии с предложенными критериями.

Работа КМК аналогична кинетике химической реакции. Справедливым здесь будет утверждение о том, что время надежного вычисления нужной (в нашем случае) цепочки, являющейся заполнением лакуны, существенно зависит от «концентрации» тех слов, которые могут породить подходящие продолжения опоры (точнее, время построения подходящей цепочки определяется произведением таких «концентраций» с точностью до некоторого коэффициента, не зависящего от них). Высокий параллелизм вычислений для заполнения лакуны означает, что мы можем параллельно вычислять любое наперед заданное число продолжений одной и той же опоры.

КМК является идеальным вычислительным устройством, обеспечивающим высокий уровень распараллеливания вычислений. Тем не менее, задача физического построения такого устройства далека от разрешения; поэтому мы будем строить такой имитатор КМК

на обычной последовательной машине фон Неймановского типа, который необходим для решения нашей конкретной задачи, а не всех алгоритмов, которые могут быть представимы в КМК. Мы делаем это для исключения нерезультативных шагов КМК и создания более оптимальной программы, решающей нашу задачу заполнения лакун. Эффективность работы такого имитатора очень сильно падает с уменьшением «концентрации» реагентоспособных слов. В первом приближении эффективность работы последовательного имитатора КМК может быть оценена по аналогии с оценками скорости химической реакции [7]. Для повышения эффективности заполнения лакуны с помощью последовательного имитатора КМК он был модифицирован.

Имитатор КМК должен до определенной степени отражать работу параллельного вычислительного устройства. Для этого число копий затравок (левых опор, для определенности) бралось достаточно большим; в наших вычислительных экспериментах это число составляло 10^2 – 10^5 копий. Так как имитатор перебирал все затравки последовательно, лишь имитируя параллельную работу, эффективность работы такого имитатора существенно определялась «концентрацией» тех слов, которые могли провзаимодействовать с затравкой (опорой, либо словом, прошедшим инициализацию). Для повышения эффективности построения заполнений лакун в символьной последовательности последовательный имитатор КМК был модифицирован; всего было внесено три модификации.

Первая. Все затравки росли только в одном направлении – слева направо, для определенности. Как только цепочка достигала нужной длины $L + 2t$, она проверялась на включение в нее правой опоры. Если правая опора в нее входила, то эта цепочка считалась одним из возможных заполнений, в противном случае она отбрасывалась.

Вторая. Модификации подвергся словарь, по которому строилось заполнение. Для реализации этапов, соответствующих работе команд (4), исходный словарь, из которого брались слова, заменялся на модифицированный. Модифицированный словарь содержал только те слова, которые имели начала, соответствующие затравке. Модификация частотного словаря означает построение на ансамбле M^* новой функции $\bar{F}_M: \Omega^* \rightarrow N \cup \{0\}$, такой, что $\bar{F}_M = F_M(uv^1v^{q-1}*) + F_M(*v^{q-1}v^1w)$ для всех v^1, v^{q-1} , для которых выполняются команды подпрограмм (4) и (5) и равна нулю во всех остальных случаях. Поскольку в общем случае у одной опоры существует несколько продолжений, из всех возможных продолжений случайным образом выбиралось одно (для данной затравки) с вероятностью, пропорциональной доле этого продолжения. Для реализации этапов, соответствующих работе

команд (4) КМК, исходный словарь, из которого брались слова, заменялся на тот, который содержал только те слова, которые имели начала, соответствующие слову, прошедшему инициализацию.

Третья. Периодически проводилась селекция всех слов, являющихся продолжениями опор, построенных в силу команд (4) КМК. Среди продолжений слова $uv^1v^{q-1}*$ могут быть такие, которые сами уже не имеют никаких продолжений среди слов из используемого в текущем вычислительном эксперименте частотного словаря. Тем самым, возможна ситуация, в которой для некоторых слов команда (4) не выполняется никогда. С точки зрения повышения эффективности работы последовательного имитатора КМК, такие «тупиковые» слова следует удалить. С другой стороны, удаление таких «тупиковых» слов на каждом шаге времени существенно понижает эффективность работы имитатора: приходится сравнивать большое количество слов. Соответственно, селекция (удаление «тупиковых» слов) проводилось не постоянно, а дважды за время роста продолжений. Для этого вся лакуна, которую необходимо было заполнить, разбивалась на три интервала равной длины. Понятно, что некоторые затравки дали такие продолжения, которые обрывались на длине, меньшей длины ее первого фрагмента – см. описание работы КМК выше. По достижении остальными словами этой пороговой длины (равной трети длины лакуны) «тупиковые» слова из всего множества слов, с которыми работает имитатор КМК, удалялись. Затем те слова, которые достигли этой критической длины, удваивались (либо, в общем случае, их число увеличивалось в k раз) и процедура построения заполнения в силу команд (4) продолжалась до тех пор, пока эти слова не достигали следующей длины, на которой проводилась селекция. По достижении этой длины (составляющей две трети от длины лакуны) оставшиеся слова опять «размножались».

4 РЕЗУЛЬТАТЫ

Целью настоящей работы является изучение качества восстановления лакуны в символьной последовательности с помощью КМК. Изучение качества восстановления проводилось с помощью вычислительных экспериментов по заполнению лакун в символьной последовательности. В качестве тест-объектов использовались следующие символьные последовательности: двоичные последовательности, троичные последовательности; генетические последовательности, тексты из русского языка.

В работе рассмотрены предварительные результаты восстановления утерянных данных, восстановленных в силу двух различных принципов выбора наилучшего заполнения (2, 3).

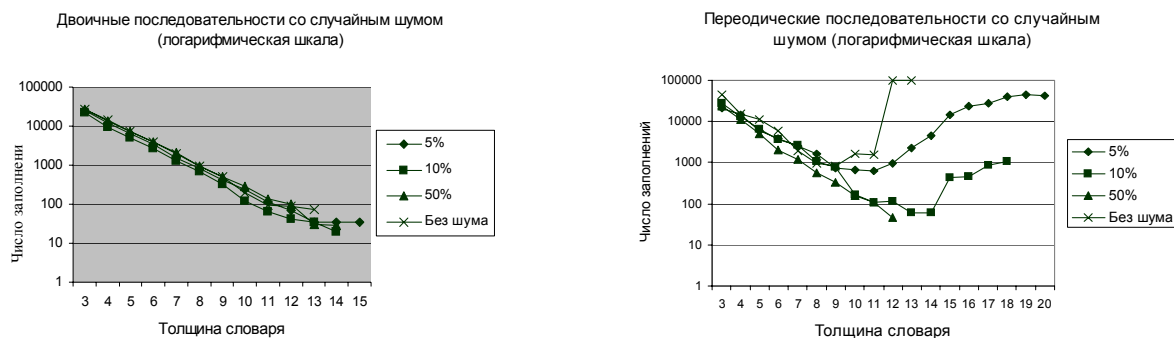


Рисунок 1 – Число затравок, доросших до полного заполнения для последовательностей с различной долей шума: по горизонтали – толщина словаря, по которому строилось заполнение; по вертикали – число успешно доросших затравок

Двоичные числа. Во всех вычислительных экспериментах использовались одинаковые параметры: длина последовательности составляла не менее 6000 символов; длина лакуны – от 50 до 200 символов. Расположение лакун внутри исходной последовательности выбиралось произвольным образом. Последовательность из алфавита $\{0, 1\}$ получалась выписыванием подряд без пробелов натуральных чисел, записанных в двоичной системе счисления: 110111001011101111000100110101011... и так далее. Полученная таким образом $\{0, 1\}$ – последовательность затем подвергалась случайным искажениям. Из нее порождались четыре другие последовательности, в которых 5 %, 10 % и 50 % случайно выбранных символов заменялись на противоположные. На рис. 1 показаны результаты восстановления последовательностей с внесенным шумом.

Троичные числа. Последовательность (по-прежнему, из алфавита $\{0, 1\}$) получалась выписыванием подряд без пробелов натуральных чисел, записанных в троичной системе счисления, из которой затем вычеркивался символ 2: 110111011001011011011111011100010 и так далее. Полученная таким образом последовательность затем также подвергалась случайным искажениям с той же интенсивностью мутаций.

Периодические последовательности. Порождалась короткая (длиной 100 символов) последовательность, которая затем периодически повторялась до тех пор, пока не возникала последовательность общей длины 6000 символов.

Четырехбуквенные последовательности. В качестве рассматриваемой последовательности был взят текст генетической последовательности (AB012132 – геном вируса парагриппа человека). Длина текста составляла 14573 символа, длина лакуны – 900 символов. Рассматривались заполнения с толщиной словаря от 3 до 8 символов. Число затравок для каждой толщины словаря бралось равным 100000. Число заполнений, у которых концы совпали с правой опорой, уменьшалось с ростом толщины словаря от 1 % (для словаря толщи-

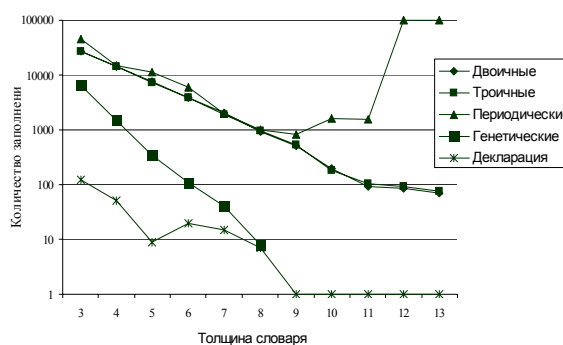


Рисунок 2 – Число затравок, доросших до полного заполнения для различных видов последовательностей: по горизонтали – толщина словаря, по которому строилось заполнение, по вертикали – число успешно доросших затравок

ны 3) до 0,001 % (для словаря толщины 8). Число несовпавших позиций составляло, в среднем, около 70 %.

Тридцатидвухбуквенные последовательности. В качестве последовательности был взят текст Всемирной декларации прав человека. Из оригинального текста декларации были удалены пробелы и знаки препинания. Получившаяся символьная последовательность составила около 7500 символов, длина лакуны – 100 символов. Использовались словари толщиной от 3 до 11 символов. Число затравок для каждой толщины словаря бралось равным 100000. Число заполнений, совпавших с правой опорой составляло для всех словарей около 0,0001 %. Число несовпавших позиций составляло, в среднем, около 90 %. Результаты построения заполнений для различных видов текста показаны на рис. 2.

В таблицах 1 и 2 представлены значения абсолютной и условной энтропии для наилучшего заполнения лакуны, полученных с помощью КМК; 2 – двоичные последовательности, 3 – троичные, p – периодические, 4 – из четырехбуквенного алфавита (генетический текст), и 32 – последовательность из русского языка.

Таблиця 1 – Значення абсолютної ентропії для найкращого заповнення, побудованого по опорному словарю, в силу критерія максимуму абсолютної ентропії

q	Види текстових послідовностей				
	2	3	p	4	32
3	2.07377	2.07389	2.68792	4.12773	7.02394
4	2.76145	2.76465	2.68761	5.49251	7.74578
5	3.45149	3.45531	3.3222	6.84594	8.10222
6	4.1406	4.14536	3.86802	8.16154	8.31051
7	4.83561	4.83543	4.22472	9.30995	8.44191
8	5.52357	5.52325	4.41634	10.0243	8.53062
9	6.20128	6.19984	4.50374		8.60683
10	6.83363	6.8333	4.56362		8.65466
11	7.36101	7.36099	4.60516		8.69932
12	7.79148	7.79196	4.60516		
13	8.1064	8.10517	4.60516		

Таблиця 2 – Значення умовної ентропії для найкращого заповнення, побудованого по повному словарю, в силу критерія мінімуму умовної ентропії

q	Види текстових послідовностей				
	2	3	p	4	32
3	$2.11001 \cdot 10^{-7}$	$2.11731 \cdot 10^{-7}$	$1.31713 \cdot 10^{-6}$	$1.08111 \cdot 10^{-6}$	0.002501334
4	$8.63964 \cdot 10^{-7}$	$1.15992 \cdot 10^{-6}$	$5.68907 \cdot 10^{-7}$	$7.28976 \cdot 10^{-6}$	0.004114464
5	$8.80001 \cdot 10^{-6}$	$9.15474 \cdot 10^{-6}$	$4.37778 \cdot 10^{-6}$	$3.55110 \cdot 10^{-5}$	0.005274570
6	$3.25038 \cdot 10^{-5}$	$2.83748 \cdot 10^{-5}$	$2.17838 \cdot 10^{-5}$	0.000127584	0.006469274
7	$8.54041 \cdot 10^{-5}$	$9.11896 \cdot 10^{-5}$	$2.43775 \cdot 10^{-5}$	0.000383415	0.007102911
8	0.000250881	0.000225592	$1.81110 \cdot 10^{-5}$	0.000727322	0.006994414
9	0.000533898	0.000531811	$9.28852 \cdot 10^{-6}$		0.007878056
10	0.00102765	0.00104568	$1.12619 \cdot 10^{-5}$		0.00755565
11	0.0018486	0.00176425	$1.32075 \cdot 10^{-5}$		0.00835729
12	0.0025176	0.00244363	$1.43811 \cdot 10^{-5}$		
13	0.0034176	0.00342195	$1.55263 \cdot 10^{-5}$		

ВИВОДИ

Роботи, посвященні відновленню пропущених даних, в основній своїй масі посвящені багатомірним даним [8–14]. В цих роботах об'єкти (дані) представляються точкою в багатомірному просторі, а параметри об'єкта являються координатами цієї точки. При цьому для відновлення пропущених координат зазвичай вимагається деяка апріорна інформація. Наукова новизна викладеного тут методу полягає в тому, що він працює з символічними послідовностями, при цьому втрачена частина символічної послідовності відновлюється за допомогою тільки тієї інформації, яка міститься в самій символічній послідовності (частотні словари). Крім того, завдання побудови

всіх можливих заповнень є переборною; вона вирішується за допомогою імітатора високопаралельного мелкозернистого процесора.

На практиці метод може застосовуватися в усіх завданнях, які вимагають відновлення пропусків в символічних послідовностях – від теорії передачі даних до молекулярної біології.

ПЕРЕЧЕНЬ ССЫЛОК

1. Горбунова Е. О. Формально-кінетична модель безструктурного мелкозернистого паралелізму // Сиб. журн. числ. математики. – 1999. – Т. 2, № 3. – С. 239–256.
2. Gorban A. N., Gorbunova K. O., Wunsch D. C. Liquid Brain: Kinetic Model of Structureless Parallelism // Advances in Modelling & Analysis. – ASME. – 2000. – V. 5, No. 5. – Pp. 33–56.

3. Gorban A. N., Gorbunova K. O., Wunsch D. C. Liquid Brain: The Proof of Algorithmic Universality of Quasichemical Model of Fine-Grained Parallelism // Neural Network World. – 2001. – № 4. – P. 391–412.
4. Неменчинская Е. О., Кондратенко Ю. В., Садовский М. Г. Предварительные результаты в проблеме восстановления утерянных данных с помощью кинетической машины Кирдина // Вычислительные технологии. – 2004. – Т. 9, № 1. – С. 42–57.
5. Gorbunova E. O., Kondratenko Yu.V., Sadovsky M. G. Data loss reparation due to indeterminate fine-grained parallel computation // ICCS, LNCS 2658, Springer-Verlag, Berlin Heidelberg, 2003. – P. 794–801.
6. Nemenchinskaya E. O., Kondratenko Yu.V., Sadovsky M. G. Entropy based approach to data loss reparation through the indeterminate fine-grained parallel computation // Open Systems & Information Dynamics. – 2004. – V. 11, № 2. – P. 161–175.
7. Яблонский Г. С., Быков В. И., Горбань А. Н. Кинетические модели каталитических реакций. – Новосибирск: Наука, 1983. – 253 с.
8. Загоруйко Н. Г., Елкина В. Н., Тимеркаев В. С. Алгоритм заполнения пропусков в эмпирических таблицах (алгоритм «ZET») // Вычислительные системы. – Новосибирск, 1975. – Вып. 61. Эмпирическое предсказание и распознавание образов. – С. 3–27.
9. Злоба Е., Яцкив И. Статистические методы восстановления пропущенных данных // Компьютерные модели и новые технологии. – 2002. – Т. 6, № 1. – С. 51–61.
10. Голяндина Н., Осипов Е. Метод «Гусеница»-SSA для анализа временных рядов с пропусками // Сб. научн. тр. НИИХ. – 2005. – С. 23–37.
11. Jaap P. L. Brand. Development, implementation and evaluation of multiple imputation strategies for the statistical analysis of incomplete data // Print partners ispkamp, Enschede. – 1999. – 333 с.
12. Литтл Р. Дж.А., Рубин Д. Б. Статистический анализ данных с пропусками. – М.: Финансы и статистика, 1991.
13. Горбань А. Н., Макаров С. В., Россиев А. А. Нейронный конвейер для восстановления пробелов в таблицах и построения регрессии по малым выборкам с неполными данными // V Международная конференция «Математика, компьютер, образование»: Тез. докл. – Дубна, 1998. – С. 53.
14. Горбань А. Н. Проблема скрытых параметров и задачи транспонированной регрессии // V Всеросс. семинар «Нейроинформатика и ее приложения»: Тез. докл. – Красноярск, 1997. – С. 15–16.

Надійшла 21.07.06
Після доробки 29.01.07

Викладені попередні результати розв'язку проблеми відновлення послідовності символів за допомогою кінетичної машини Кірдіна (КМК). КМК це мілкозернистий високопаралельний розподілений обчислювальний формалізм. Запропонований модифікований імітатор КМК, спеціалізований для даної задачі. Імітатор працює з частиною частотного словника, що включає лише реагентоздатні слова. Розв'язок задачі проілюстровано рядом прикладів. Розглянуті наступні напрямки використання КМК для задач відновлення загублених даних.

Preliminary results of the solving of the lost data recovery problem by means of the Kirdin kinetic machine (the ideal fine-grained and highly parallel distributed computing formalism) are presented. The modified simulator of the Kirdin kinetic machine (KKM) specialized for effective construction of the fillings is proposed. The simulator works with the part of the frequency dictionary containing only reactionable words, with periodic removal of those variants of fillings which had no suitable continuations. The decision of a problem is illustrated by a number of examples. The further directions in application KKM for problems of restoration of the lost data are discussed.

УДК 621.744:004:62-92

А. Л. Становский, Т. В. Лысенко, Н. П. Худенко

УПРАВЛЕНИЕ СОСТОЯНИЯМИ В СЛОЖНЫХ ТЕХНИЧЕСКИХ ОБЪЕКТАХ

Предложена система управления состояниями в технических объектах, положительный эффект от деятельности которой основан на достижении с помощью управления совпадения во времени событий в подсистемах тепло и массопереноса, дающем значительный качественный эффект.

ВВЕДЕНИЕ

При управлении процессами, распадающимися на отдельные подпроцессы, важной задачей является обеспечение оптимальной параллельности течения этих подпроцессов.

Существующие методы сетевого планирования [1], а также различные синхронизирующие алгоритмы [2, 3], позволяют в значительной степени снизить время про-

текания процесса в целом, а также расходы материальных, энергетических, трудовых и прочих ресурсов. Задачей синхронизации в этом случае является также сохранение качества продукции. Распараллеливание обработки на подпроцессы не должно приводить к его ухудшению.

В то же время, многие процессы, протекающие в объектах управления, не нуждаются в принудительном распараллеливании, т. к. их течение определяется физическими законами сохранения и переноса. Примером такого процесса является теплообмен, который можно рассматривать как параллельно протекающие подпроцессы тепло- и массообмена, которые нельзя реализовать последовательно.

ПОСТАНОВКА ЗАДАЧИ

В настоящее время одним из основных и наиболее эффективных подходов к решению задач управления большой размерности является декомпозиция [4, 5]. Этот подход объединяет группу методов, основанных на разложении исходной задачи большой размерности на подзадачи, каждая из которых существенно проще исходной и решается независимо от других. Связь между отдельными подзадачами осуществляется с помощью «координирующей» задачи, которая тоже проще исходной. Для этого задачу управления приводят к виду, удовлетворяющему требованиям декомпозируемости, основными из которых [6] являются: аддитивность целевой функции; блочный характер ограничений; наличие блочных связей. Однако при решении практических задач синтеза оптимального управления большой размерности зачастую сложно удовлетворить перечисленным требованиям. Например, качество отливки может оцениваться критерием весьма общего типа, который может быть несепарабельным по отношению к задачам управления отдельными подсистемами тепло и массообмена. Поэтому при приведении исходной задачи управления к виду, удовлетворяющему требованиям декомпозируемости, неизбежны как различные упрощения, аппроксимации, так и различные варианты разбиения задачи на локальные подзадачи, т. е. блоков ограничений и межблочных связей. Все эти факторы влияют как на качество решения, так и на сложность расчетов при поиске оптимального решения [7].

РЕШЕНИЕ ЗАДАЧИ

Пусть некоторый процесс, например, в песчаной литейной форме описывается зависимостью:

$$\mathbf{Y} = \mathbf{F}(\mathbf{X}, \mathbf{u}, \mathbf{d}), \quad (1)$$

где \mathbf{Y} – вектор переменных выхода; \mathbf{X} – вектор переменных состояния; \mathbf{u} – вектор управлений и \mathbf{d} – вектор возмущений.

Задача управления процессом в общем виде формулируется следующим образом [8]: найти такое \mathbf{u} , чтобы при заданных \mathbf{X} и допустимых \mathbf{d} гарантированно обеспечить требуемые \mathbf{Y} .

Применительно к литейной форме эта задача может звучать так: найти такое значение управления \mathbf{u} (например, толщины стенки и плотности формовочного материала), которое для данной технологии \mathbf{X} (например, литья в оболочковые формы) в пределах допусков на свойства исходных материалов \mathbf{d} (например, огнеупорного наполнителя и пульвербакелита) обеспечивает требуемое качество поверхности отливки \mathbf{Y}



Рисунок 1 – Схема к определению области существования вектора качества \mathbf{Y} и промежуточного вектора \mathbf{u}

(например, шероховатости поверхности и наличия макродефектов).

Естественно предположить, что переменные, входящие в \mathbf{u} , влияют на \mathbf{Y} , то есть, объект управляем.

В песчаной литейной форме протекают одновременно несколько подпроцессов, влияющих на конечный результат. Поэтому качество отливки \mathbf{Y} является интегральной характеристикой, зависящей от отдельных подпроцессов, протекающих в форме (тепло-, массоперенос, химические реакции, фазовые превращения и т. п.):

$$\mathbf{Y} = \mathbf{Y}(y_1, y_2, \dots, y_n), \quad (2)$$

где индексы 1, 2, ..., n относятся к соответствующим подпроцессам.

Важно подчеркнуть, что в выражении (2) вектор параметров качества отливки \mathbf{Y} «появляется» только после окончания техпроцесса (рис. 1) и далее не является функцией времени, а y_1, y_2, \dots, y_n существуют только во время техпроцесса и могут быть функциями времени.

Тогда целью управления может быть обеспечение минимальной разницы между фактическим изменением векторов $y_{1\text{факт}}(\tau), y_{2\text{факт}}(\tau), \dots, y_{n\text{факт}}(\tau)$ во времени и заданным изменением $y_{1\text{зад}}(\tau), y_{2\text{зад}}(\tau), \dots, y_{n\text{зад}}(\tau)$, т. е. минимизация функционала:

$$I = \sum_{i=1}^n \int_{\tau_0}^{\tau_i} (y_{i\text{факт}}(\tau) - y_{i\text{зад}}(\tau))^2, \quad (3)$$

которая обеспечивает заданное качество \mathbf{Y} .

С точки зрения реальных технологических возможностей управления здесь возможны следующие варианты:

- каждый из процессов управляется различными, не зависящими друг от друга параметрами (в литейной форме это практически неосуществимо);
- все процессы управляются одним параметром, поразному влияющим на их протекание (например, если единое u – это плотность набивки форм, то оно влияет и на процесс теплопереноса, и на процесс массообмена в песчаной форме).

Пусть выход оценивается одной переменной y , управление – одной переменной u , а возмущения отсутствуют. Тогда (1) с учетом того, что управление также является функцией времени, сведется к выражению:

$$y = f(\mathbf{X}, u, \tau), \quad (4)$$

что при $\mathbf{X} = \text{const}$ может быть представлено, например, в виде семейства кривых, приведенных на рис. 2.

Пусть некоторое значение выхода y_s интерпретируется как особое состояние – событие S (например, если y – это концентрация растворимой примеси, то $y_s = 0$ означает окончание растворения; если y – температура, то $y_s = T_{\text{лик}}$ означает событие «начало кристаллизации» и т. п.).

Как видно из рис. 2, время наступления события S зависит от управления u :

$$\tau_s = \varphi(u), \quad (5)$$

что может быть представлено, например графиком, приведенным на рис. 3, и интерпретировано как *управление событием S*.

При изменении управления в пределах технологической возможности (например, от u_1 до u_4) время наступления события S также изменяется, не обязательно монотонно.

Если рассмотреть n подпроцессов, то уравнение (1) можно представить в виде системы:

$$y_1 = f_1(\mathbf{X}_1, \mathbf{u}_1, \mathbf{d}_1); \quad (6)$$

$$y_2 = f_2(\mathbf{X}_2, \mathbf{u}_2, \mathbf{d}_2); \quad (7)$$

...

$$y_n = f_n(\mathbf{X}_n, \mathbf{u}_n, \mathbf{d}_n), \quad (8)$$

а если $\mathbf{u}_1, \mathbf{u}_2, \dots, \mathbf{u}_n$ сводится к единому управляющему воздействию u , то уравнение (5) – к системе:

$$\tau_{s1} = \varphi_1(u); \quad (9)$$

$$\tau_{s2} = \varphi_2(u); \quad (10)$$

...

$$\tau_{si} = \varphi_i(u). \quad (11)$$

При $i = 2$ график, приведенный на рис. 3, приобретает вид (рис. 4).

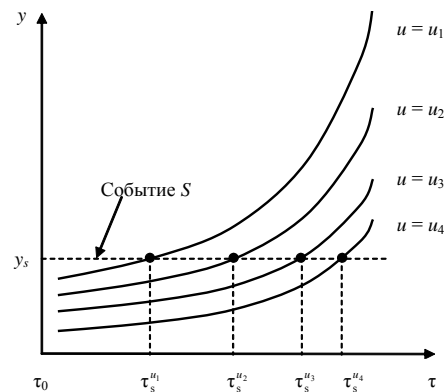


Рисунок 2 – Зависимость выхода подпроцесса от времени при различных управлениях

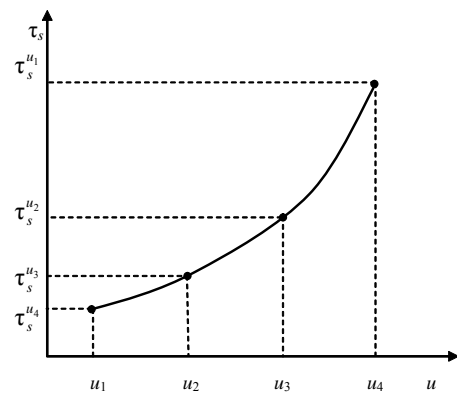


Рисунок 3 – Схема к понятию «управление событием»

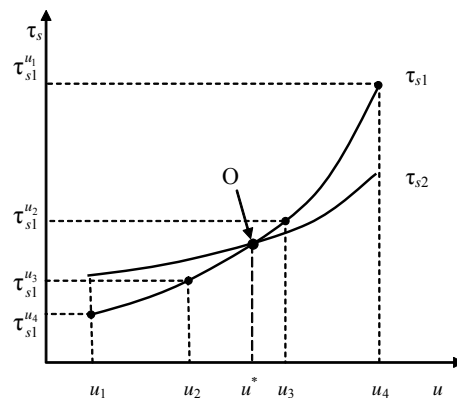


Рисунок 4 – Схема к определению понятия «оптимизация событий»

Точка O на этом графике – особая, она соответствует ситуации, когда, благодаря управлению u^* , события S_1 и S_2 наступают одновременно:

$$\tau_{s1}(u^*) = \tau_{s2}(u^*). \quad (12)$$

Поиск (расчет) такого u^* является главной задачей управления событием.

ВЫВОДЫ

Предложенная в работе система управления событиями в литейном производстве позволила разработать новые и существенным образом расширить возможности существующих технологических решений по повышению качества отливок. Система прошла многостороннюю производственную проверку с положительным технико-экономическим эффектом при проектировании и управлении технологическими процессами и оборудованием для литья в песчано-глинистые, песчано-смоляные и керамические формы, облицованные кокили, а также при литье по выплавляемым моделям.

ПЕРЕЧЕНЬ ССЫЛОК

1. Хэмди А. Таха. Введение в исследование операций. – М.: Вильямс, 2001. – 235 с.
2. Дерябина Р. М. Моделирование и оптимизация синхронизирующих расписаний. – Дисс. ... доктор техн. наук. – Кемерово, 2000. – 119 с.

3. Вермей Е. И. Введение в современные методы оптимизации систем управления // Экспонента. про. Математика в приложениях. – 2004. – № 1(5). – С. 32–35.
4. Месарович М., Мако Д, Такахара И. Теория иерархических многоуровневых систем. – М.: Мир, 1973. – 276 с.
5. Аэсдон Л. С. Оптимизация больших систем. – М.: Мир, 1973. – 237 с.
6. Балтовский А. А. Синтез оптимального закона управления большой системой на основе композиции локальных оптимальных решений // ААЭКС. – 2005. – № 2. – С. 107–110.
7. Овезгельдыев А. О., Петров Э. Т., Петров К. Э. Синтез и идентификация моделей многофакторного оценивания и оптимизации. – К.: Наукова думка, 2002. – 148 с.
8. Рей У. Методы управления технологическими процессами. – М.: Мир, 1983. – 368 с.

Надійшла 20.11.06
Після доробки 6.02.07

Запропоновано систему управління станами в технічних об'єктах, позитивний ефект від діяльності якої заснований на досягненні збігу в часі станів у різних підсистемах технічного об'єкта, що дає значний якісний ефект.

The control system of conditions in technical objects is offered, the positive effect from which activity is based on achievement of concurrence in time of conditions in various subsystems of the technical object, giving significant qualitative effect.

ПРОГРЕСИВНІ ІНФОРМАЦІЙНІ ТЕХНОЛОГІЇ

ПРОГРЕССИВНЫЕ ИНФОРМАЦИОННЫЕ ТЕХНОЛОГИИ

PROGRESSIVE INFORMATION TECHNOLOGIES

УДК 681.3.+681.5.007

Н. В. Алипов, Н. М. Кораблев, М. И. Хиль, М. В. Гусятин

ПРИМЕРЫ ПОСТРОЕНИЯ ОРИЕНТИРОВАННЫХ ГРАФОВ ПЕРЕХОДОВ ЦИФРОВОГО АВТОМАТА С ПСЕВДОСЛУЧАЙНЫМИ ПЕРЕХОДАМИ

Методом индукции строятся для конкретных параметров виртуальной последовательности ориентированные графы переходов цифровых автоматов с псевдослучайными переходами (ЦАПП).

ВВЕДЕНИЕ

Новым направлением в развитии методов защиты информации при ее передаче и хранении является направление, основанное на использовании цифровых автоматов с псевдослучайными переходами [1]. К настоящему моменту разработаны структуры таких цифровых автоматов и алгоритмы формирования виртуальных последовательностей [2]. Однако к настоящему моменту еще не рассмотрены вопросы, связанные с построением ориентированных графов переходов ЦАПП.

Цель исследования является разработка примеров построения ориентированных графов переходов ЦАПП.

ПРИМЕРЫ ПОСТРОЕНИЯ ОРИЕНТИРОВАННЫХ ГРАФОВ ПЕРЕХОДОВ ЦАПП

Следует заметить, что функционирование ЦАПП для систем защиты информации осуществляется на основе помехоустойчивых к тем или иным виртуальным последовательностям алгоритмов поиска точки x на отрезке единичной длины ($x \in [0, 1]$). Подавление виртуальных последовательностей осуществляется алгоритмически. Алгоритмическое подавление виртуальных последовательностей также как и параметрическое эффективно только для особых (хороших) последовательностей, под действием которых автомат, переходя псевдослучайным образом из одного состояния в другое, в конце концов, достиг бы конечного состояния, поставленного в соответствии кодируемому символу входного алфавита. Такие виртуальные последовательности описаны в работе [2].

Алгоритмическое подавление виртуальных последовательностей основано на теории помехоустойчивого

одномерного поиска точки с характерным признаком на отрезке единичной длины [3]. Заметим, что поиск осуществляется путем уменьшения исходного интервала неопределенности относительно исходной точки x . Исходным интервалом неопределенности является отрезок $[0, 1]$. Алгоритм поиска формирует некоторые эталонные значения на j -м шаге поиска x_p^j , где $\rho = \overline{1, k}$, k – количество эталонных величин, формирующих алгоритмом поиска, и выполняет эксперимент (сравнение координат эталонных точек и координат искомой точки x):

$$P\{x \geq x_p^j\}.$$

При этом если

$$P\{x \geq x_p^j\} = 0,$$

то искомая точка как будто находится в интервале $[x_{l-1}^j, x_n^{j-1}]$; где $[x_{l-1}^j, x_n^{j-1}]$ – полуоткрытый интервал неопределенности, выделенный на предыдущем шаге алгоритма.

Если же $P\{x \geq x_p^j\} = 1$, то формируется полуоткрытый интервал $[x_l^j, x_n^{j-1}]$.

Поиск осуществляется в условиях воздействия на x либо на x_p^j виртуальной последовательности (ВП), поэтому выделенные относительно точки x полуоткрытые неопределенности будут недостоверными. В теории поиска [4] разрабатываются специальные правила формирования нового полуоткрытого интервала неопределенности относительно x . Такие правила называются *решающими правилами*. Если алгоритм формирует на каждом шаге только одну эталонную точку, то его называют *последовательным*, если же формируется более одной эталонной точки ($k > 1$) на каждом шаге алгоритма, такой алгоритм является *параллельно-последовательным*. В данной работе решающие правила строятся на основании принципа «пересечения». Для пояснения принципа «пересечения» ниже рассмотрен пример построения решающих правил.

Пусть $k = 1$ и известна максимальная амплитуда ВП, равная « a » [2]. Тогда в результате сравнения координат на первом шаге алгоритма может один из исходов:

$$a) x \leq x_1^1; \quad б) x > x_1^1.$$

Для несимметричной нерегулярной виртуальной последовательности типа $0 \rightarrow 1$ будем иметь [5]:

– для исхода а) (ВП вследствие малой амплитуды не смогла сместить точку x за точку x_1^1)

$$x \in [0, x_1^1];$$

– для исхода б) однозначно нельзя утверждать, что в момент совершения первого шага алгоритма не проявлялась виртуальная последовательность.

Поэтому точка x могла принадлежать полуоткрытому интервалу $[0, x_1^1)$. Максимальное ее отдаление от точки x_1^1 равно амплитуде виртуальной последовательности. Для исхода б) выделяется следующий полуоткрытый интервал:

$$x \in [x_1^{1,1}, 1),$$

где

$$x_1^{1,2} = \begin{cases} x_1^1 - a\delta, & x_1^1 - a\delta \geq 0; \\ 0 & \text{в противном случае,} \end{cases}$$

где δ – дискретность преобразования по уровню.

При воздействии симметричной ВП точка x может смещаться как в направлении $0 \rightarrow 1$, так и в направлении $1 \rightarrow 0$. Поэтому при возникновении исхода типа а) его расширяют вправо, при возникновении исхода типа б) его расширяют влево и получают следующие интервалы неопределенности:

– для исхода а) будем иметь

$$x \in [0, x_1^{1,2});$$

– для исхода б) будет справедливым соотношение

$$x \in [x_1^{1,1}, 1),$$

где

$$x_1^{1,2} = \begin{cases} x_1^1 + a\delta, & x_1^1 + a\delta \geq 1; \\ 1 & \text{в противном случае,} \end{cases}$$

$$x_1^{1,1} = \begin{cases} x_1^1 - a\delta, & x_1^1 - a\delta \geq 0; \\ 0 & \text{в противном случае.} \end{cases}$$

Заметим, что под шагом понимают формирование эталонных точек, выполнение эксперимента и выделение нового интервала неопределенности.

В теории поиска [4] имеется еще одно важное понятие – *стратегия поиска*. Стратегия поиска – правило выбора точек нового эксперимента во вновь выделенном полуоткрытом интервале неопределенности.

Пусть на первом шаге алгоритма некоторым образом выбраны k точек первого эксперимента $x_1^1, x_2^1, \dots, x_k^1$. В результате выполнения первого шага алгоритма получено:

$$x \geq x_1^1, \quad x > x_2^1 \dots \geq x \geq x_2^1, \quad x < x_{2+1}^1 \dots \geq x < x_k^1.$$

В этом случае выделяется интервал неопределенности

$$[x_z^{1,1}, x_{z+1}^{1,2}), \quad Z = \overline{1, k},$$

где

$$x_z^{1,1} = \begin{cases} x_z^1 - a\delta, & x_z^1 - a\delta \leq 0; \\ 0 & \text{в противном случае;} \end{cases}$$

$$x_{z+1}^{1,2} = \begin{cases} x_{z+1}^1 + a\delta, & x_{z+1}^1 + a\delta \leq 1; \\ 1 & \text{в противном случае.} \end{cases}$$

Стратегии поиска бывают оптимистическими, пессимистическими и смешанными. Для оптимистической стратегии справедливо соотношение

$$x_{z_1}^2 \in (x_z^1, x_{z+1}^1), \quad z_1 = \overline{1, k}.$$

Для пессимистической стратегии имеют место соотношения

$$x_{z_1}^2 = x_z^1, \quad \text{где } z_1 = \overline{1, k}.$$

Эту стратегию еще называют принципом «повторных сравнений». Для смешанной стратегии и симметричной последовательности справедливы соотношения

$$x_1^2 = x_z^1; \quad x_q^2 \in (x_z^1, x_{z+1}^1); \quad q = \overline{2, k-1}; \quad x_k^2 = x_{z+1}^1.$$

Для несимметричной ВП вида $0 \rightarrow 1$ справедливы соотношения

$$x_1^2 = x_z^1; \quad x_{q_1}^2 \in (x_z^1, x_{z+1}^1); \quad q_1 = \overline{2, k}.$$

Для последовательных алгоритмов существуют две стратегии [5]: оптимистическая и пессимистическая. Для оптимистической стратегии при симметричной ВП характерно размещение точки эксперимента во вновь выделенном интервале неопределенности относительно x . Для пессимистической стратегии характерно то, что $x_1^2 = x_1^1$. При этом могут возникать следующие совокупности исходов:

$$\text{а) } x + \xi(t_1) \leq x_1^1; \quad \text{б) } x + \xi(t_1) > x_1^1,$$

$$\text{а}_1) x + \xi(t_2) \leq x_1^2; \quad \text{а}_2) x + \xi(t_2) > x_1^2,$$

$$\text{б}_1) x + \xi(t_2) > x_1^2; \quad \text{б}_2) x + \xi(t_2) < x_1^2.$$

Пусть $l = 1$ [5]. Тогда совокупность исходов а), а₁), б), б₁) подтверждает истинность соотношений

$$x \in [0, x_1^1]; \quad x \in [x_1^1, 1).$$

Исходы а) и а₂), б) и б₂) противоречат друг другу. Поэтому третий шаг выполняется таким образом:

$$x_1^3 = x_1^1.$$

При этом могут возникнуть такие исходы:

$$\text{а}_{21}) x + \xi(t_3) \leq x_1^3;$$

$$\text{а}_{22}) x + \xi(t_3) > x_1^3;$$

$$\text{б}_{21}) x + \xi(t_3) \geq x_1^3;$$

$$\text{б}_{22}) x + \xi(t_3) < x_1^3.$$

Совокупность исходов а), а₂₁) подтверждает истинность выражения $x \in [0, x_1^1]$. Совокупность исходов а₂), а₂₂) подтверждает истинность соотношения $x \in [x_1^1, 1)$. Совокупность исходов б), б₂₁) подтверждает истинность выражения $x \in [x_1^1, 1)$. Совокупность исходов б₂), б₂₂) подтверждает истинность утверждения $x \in [0, x_1^1]$.

Для несимметричных ВП пессимистическая стратегия применяется в случае, когда возникает исход

$$\text{б) } x + \xi(t_1) \geq x_1^1, \quad x \in [x_1^{1,1}, 1).$$

Будем предполагать, что оптимальный последовательный алгоритм поиска разбивает исходный полуоткрытый интервал неопределенности $[0, 1)$ за i шагов на $\varphi(i, 1)$ равных частей, за $(i-1)$ – на $\varphi(i-1, 1)$ равных частей и т. д.

При выполнении второго шага алгоритма ($x_1^2 = x_1^1$) могут возникнуть такие исходы

$$\text{б}_1) x + \xi(t_2) \leq x_1^2; \quad \text{б}_2) x + \xi(t_2) > x_1^2.$$

Для исхода б₁) характерно то, что виртуальная последовательность действовала на первом шаге алгоритма и по условию еще $(h-1)$ шаг она не будет проявляться. На этих шагах на отрезке $[x_1^{1,1}, x_1^1]$ действует классический алгоритм поиска (применяется дихотомия). По предположению выделений на $(h+1)$ -м шаге алгоритма интервал неопределенности будет разбит $\varphi(i-h-1, 1)$ равные части. Поскольку первоначальный полуоткрытый интервал неопределенности $[x_1^{1,1}, x_1^1]$ посредством применения дихотомии был разбит на 2^{h-1} равные части, а затем каждая из них помехоустойчи-

вым алгоритмом были разбиты на $\varphi(i-h-1, 1)$ равных частей, то полуоткрытый интервал неопределенности $[x_1^{1,1}, x_1^1)$ будет разбит $2^{h-1}\varphi(i-h-1, 1)$ равных частей.

Поскольку $l = 1$, то исход $b_2)$ подтверждает исход б) и $x \in [x_1^1, 1)$ виртуальная последовательность при такой длительности выброса не может действовать на двух соседних шагах алгоритма.

Истрчено два шага алгоритма и выделенный на втором шаге алгоритма полуоткрытый интервал будет разбит на $\varphi(i-2, 1)$ равных частей.

По итогам выполнения второго шага алгоритма для исхода $b_1)$ могут быть истинными такие соотношения:

$$l([x_1^{1,1}, x_1^1)) = \delta\{2^{h-1}\varphi(i-h-1, 1)\};$$

$$l([x_1^{1,1}, x_1^1)) < \delta\{2^{h-1}\varphi(i-h-1, 1)\};$$

$$l([x_1^{1,1}, x_1^1)) > \delta\{2^{h-1}\varphi(i-h-1, 1)\}.$$

Если истинным будет первое соотношение, то на втором шаге алгоритма следует применять пессимистическую стратегию; если второе неравенство справедливо, то можно попробовать применить на втором шаге оптимистическую стратегию; если справедливым будет третье выражение, то для заданного параметра h помехоустойчивого алгоритма не существует (необходимо увеличить h либо уменьшить амплитуду выброса).

Пусть будет справедливым второе неравенство. Тогда для исхода б) на втором шаге попробуем применить оптимистическую стратегию $x_1^2 \in [x_1^1, 1)$.

При этом в результате выполнения второго шага алгоритма может сформироваться один из исходов:

$$c_1) x + \xi(t_2) \leq x_1^2; c_2) x + \xi(t_2) > x_1^2.$$

Для исхода $c_1)$ на основании принципа «пересечения» устанавливаем $x \in [x_1^{1,1}, x_1^2)$.

Если на третьем шаге использовать пессимистическую стратегию $x_1^3 = x_1^1$, то в результате его выполнения может появиться один из исходов:

$$c_{11}) x + \xi(t_3) \leq x_1^3; c_{12}) x + \xi(t_3) > x_1^3.$$

Поскольку исход $c_{11})$ противоречит исходу б), то для несимметричной виртуальной последовательности устанавливаем: виртуальная последовательность действовала на первом шаге алгоритма, на втором и третьем шагах она отсутствовала, $x \in [x_1^{1,1}, x_1^1)$; она также не будет проявляться на последующих $(h-2)$ шагах алгоритма. Применяем на выделенном полуоткрытом интервале $[x_1^{1,1}, x_1^1)$ первоначально на $(h-2)$ шагах дихотомию: затем на последующих $(i-h-1)$ шагах к вы-

деленному дихотомией интервалу неопределенности помехоустойчивый алгоритм, разобьем этот полуоткрытый интервал неопределенности на $2^{h-2}\varphi(i-h-1, 1)$ равных частей.

Если при этом имеет место соотношение:

$$l([x_1^{1,1}, x_1^1)) \leq \delta(2^{h-2}\varphi(i-h-1, 1)), \quad (1)$$

то на втором шаге применяем оптимистическую стратегию; если это соотношение нарушается, то на втором шаге применяем пессимистическую стратегию, и продолжаем процесс построения алгоритма.

Если по итогам третьего шага формируется исход $c_{12})$, то с учетом исхода б) устанавливаем $x \in [x_1^1, x_1^2)$.

Поскольку в распоряжении алгоритма осталось $(i-3)$ шага, то по предположению его за оставшиеся шаги алгоритма помехоустойчивым алгоритмом разобьем на $\varphi(i-3, 1)$ равных частей.

Если на втором шаге применена оптимистическая стратегия и возник исход $c_2)$, то на основании принципа «пересечения» устанавливаем:

$$x \in [x_1^{2,1}, 1), x_1^{2,1} = \begin{cases} x_1^2 - a\delta, x_1^2 - a\delta \geq x_1^1; \\ x_1^1 \text{ в противном случае.} \end{cases}$$

В дальнейшем этот исход разрешаем таким же образом как исход б).

Следует заметить, что $\varphi(i, 1)$ определяется как сумма количества равных частей отдельных отрезков, на которые разбивается исходный полуоткрытый интервал неопределенности $[0, 1)$.

Соотношения, определяющие применение оптимистической стратегии для симметричной виртуальной последовательности, строятся описанным способом.

Эффективность алгоритмов поиска оценивается длиной интервала неопределенности, полученного на последнем шаге алгоритма. В общем случае длина интервала неопределенности определяется как местоположением точки x , так и алгоритмом. Поэтому в качестве оценки эффективности S -го алгоритма возьмем величину

$$L_S = \max_{x \in [0, 1]} \{l_i(x, S)\},$$

где i – длина поиска (число шагов алгоритма); $l_i(x, S)$ – длина интервала неопределенности, полученного на последнем i -м шаге S -го алгоритма.

Оптимальным алгоритмом назовем такой алгоритм, оценка которого удовлетворяет соотношению

$$L_{S_1} = \min_{S \in M_1} \max_{x \in [0, 1]} \{l_i(x, S)\}, \quad (2)$$

где M_1 – множество возможных алгоритмов.

Обратную величину от L_S обозначим $\varphi(i, k)$. Она показывает, на сколько равных частей разбивает исходный интервал неопределенности S -й алгоритм. Критерий (2) в таком случае можно записать следующим образом:

$$\varphi_{S_1} = \max_{z \in M_1} \varphi_z(i, k). \quad (3)$$

Итак, алгоритм поиска будет характеризоваться длиной поиска (количеством шагов алгоритма i), количеством точек эксперимента k и функцией $\varphi_z(i, k)$.

Заданными считаются: параметры виртуальной последовательности ($h_1, h_2, l_1, l_2, a_1, a_2$), параметры алгоритма i, k . Требуется построить алгоритм, удовлетворяющий критерию оптимальности (3). Заметим, что для ряда задач (защита информации, АЦП и др.) строят и неоптимальные алгоритмы.

Построение алгоритмов функционирования ЦАПП осуществляют методом индукции: сначала строят одношаговый, затем двухшаговый алгоритм и т. д., рассматривая при этом всевозможные исходы.

Примеры построения ориентированных графов переходов ЦАПП.

Рассмотрим случай, для которой виртуальная последовательность – симметричная; параметры алгоритма $k = 1, i = 6$; параметры виртуальной последовательности

$$l = 1, a = 4, h_1 = 5.$$

Показано, что для $i = 0, 1, 2$ значение функции соответственно равно:

$$\varphi(0, 1) = 1; \varphi(1, 1) = 1; \varphi(2, 1) = 1. \quad (4)$$

Пусть $i = 3$ и в некоторой точке $x_1^1 \in [0, 1]$ выполняется первый эксперимент. При этом возможен исход:

$$a) \ x + \xi(t_1) \leq x_1^1.$$

Тогда, поскольку виртуальная последовательность симметричная, то достоверно нельзя утверждать, что $x \in [0, x_1^1]$. По этой причине применяем на втором шаге алгоритма пессимистическую стратегию:

$$x_1^2 = x_1^1.$$

В результате выполнения второго шага алгоритма может появиться один из исходов:

$$a_1) \ x + \xi(t_2) \leq x_1^1;$$

$$a_2) \ x + \xi(t_2) > x_1^1.$$

Поскольку исход a_1) подтверждает исход а), то для виртуальной последовательности, для которой $l = 1$, устанавливаем:

$$x \in [0, x_1^1].$$

Этот полуоткрытый интервал неопределенности за оставшийся последние шаги алгоритма будет разбит на $\varphi(3-2, 1)$ равные части. Отсюда устанавливаем

$$l([0, x_1^1]) = \delta\varphi(3-2, 1) = \delta\varphi(1, 1) = \delta, \quad (5)$$

где δ – дискретность преобразование по уровню, $l(a, b)$ – длина отрезка (a, b) .

Для исхода a_2) возникшее противоречие свидетельствует о том, что виртуальная последовательность (ВП) имела место либо на первом шаге, либо на втором шаге алгоритма. По этой причине следует на третьем шаге алгоритма применить снова пессимистическую стратегию:

$$x_1^3 = x_1^1.$$

В результате совершения третьего шага возможен один из исходов:

$$a_{21}) \ x + \xi(t_3) < x_1^1;$$

$$a_{22}) \ x + \xi(t_3) > x_1^1.$$

Поскольку исход $a_{21})$ для $l = 1$ подтверждает исход а), то устанавливаем истинность соотношения:

$$x \in [0, x_1^1].$$

Поскольку все эксперименты выполнены, то с учетом соотношений (1) будем иметь

$$l([0, x_1^1]) = \delta\varphi(3-3, 1) = \delta. \quad (6)$$

Соотношение (6) не противоречит соотношению (5). Для исхода $a_{22})$ характерно то, что ВП имела место на первом шаге алгоритма, ее проявление на втором и третьем шагах алгоритма не имела места. В этой ситуации, когда все шаги алгоритма использована, то с учетом соотношений (1) будем иметь:

$$l([x_1^1, 1]) = \delta\varphi(3-3, 1) = \delta. \quad (7)$$

Если при совершении первого шага алгоритма возник исход

$$б) \ x + \xi(t_1) > x_1^1,$$

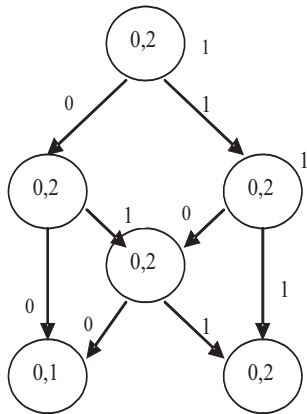


Рисунок 1 – Ориентированный граф переходов ЦАПП

то его разрешают аналогично исходу а). В этом случае на основании соотношений (5), (6), (7) устанавливаем:

$$\varphi(3, 1) = 2. \tag{8}$$

С учетом принятых соглашений и основных соотношений для рассматриваемых исходов ориентированный граф переходов ЦАПП приведен на рис. 1.

На рисунке вершина графа соответствует состоянию автомата; дуга (стрелка) – возможному переходу автомата из конкретного состояния; дуги графа помечены символами «0» (возникают исходы $x + \xi(t_j) \leq x_1^j$); либо «1» (возникают исходы $x + \xi(t_j) > x_1^j$); справа от вершин графа записаны координаты эталонных точек; внутри вершин графа координаты начала и конца исходных интервалов неопределенности относительно точки с характерным признаком.

Пусть $i = 4$. Тогда в результате совершения эксперимента в точке x_1^1 может быть сформирован один из исходов

$$\text{а) } x + \xi(t_1) \leq x_1^1; \text{ б) } x + \xi(t_1) > x_1^1. \tag{9}$$

Для исхода а) применяем пессимистическую стратегию

$$x_1^2 = x_1^1,$$

в результате которой может быть сформирован один из исходов типа а₁) либо а₂).

Для исхода а₁) выделяется такой полуоткрытый интервал неопределенности $[0, x_1^1)$ (см. построение алгоритма для $i = 3$). Поскольку в рассмотренном алгоритме осталось два шага, то на основании соотношений (3) устанавливаем

$$\text{а}_1) I([0, x_1^1)) = \delta\varphi(4 - 2, 1) = \delta. \tag{10}$$

Поскольку ВП симметричная, то для исхода б) имеет место соотношение

$$I([x_1^1, 1)) = \delta. \tag{11}$$

С учетом соотношения (10), (11) устанавливаем

$$\varphi(4, 1) = 2. \tag{12}$$

Пусть $i = 5$. Тогда в результате выполнения первого эксперимента может появиться один из исходов а) либо б). Пусть таковым исходом был сформирован исход а). Для исхода а) на втором шаге алгоритма применяем оптимистическую стратегию:

$$x_1^2 < x_1^1.$$

В результате выполнения второго шага может появиться один из исходов

$$\text{а}_1) x + \xi(t_2) \leq x_1^2; \text{ а}_2) x + \xi(t_2) > x_1^2.$$

Для исхода а₁) характерно то, что он подтверждает исход а). Поскольку $l = 1$, то исход а) является достоверным (ВП не могла действовать на двух соседних шагах алгоритма). На этом основании устанавливаем

$$x \in [0, x_1^1). \tag{13}$$

Считая второй шаг первым шагом четырех шагового алгоритма, с учетом соотношения (12), интервал неопределенности $[0, x_1^1)$ будет разбит на $\varphi(4, 1)$ равных частей (один эксперимент был совершен в точке x_1^1). По этой причине имеет место соотношение

$$I([0, x_1^1)) = 2\delta. \tag{14}$$

Для исхода а₂) в точке x_1^1 на третьем шаге применяем пессимистическую стратегию:

$$x_1^3 = x_1^1.$$

При этом может появиться один из исходов

$$\text{а}_{21}) x + \xi(t_3) \leq x_1^1; \text{ а}_{22}) x + \xi(t_3) > x_1^1.$$

Если возможен исход а₂₁), то точке, x_1^2 применяем пессимистическую стратегию:

$$x_1^4 = x_1^2.$$

В этом случае может появиться один из исходов

$$a_{211}) x + \xi(t_4) \leq x_1^2; a_{212}) x + \xi(t_4) > x_1^2.$$

Для исхода a_{211}) на пятом шаге применяем пессимистическую стратегию $x_1^5 = x_1^2$. При выполнении пятого шага может появиться один из исходов

$$a_{2111}) x + \xi(t_5) \leq x_1^2; a_{2112}) x + \xi(t_5) > x_1^2.$$

Для исхода a_{2111}) возникшее противоречие при выполнении экспериментов в точке x_1^2 , разнесенных во времени, свидетельствует о действии ВП на втором шаге алгоритма. В этом случае устанавливаем

$$x \in [0, x_1^2).$$

Поскольку все эксперименты выполнены, то на основании равенств (4) устанавливаем

$$l([0, x_1^2)) = \delta\varphi(0, 1) = \delta. \quad (15)$$

Для исхода a_{2112}), возникшее противоречие свидетельствует о действии ВП на четвертом шаге алгоритма. На этом основании утверждаем

$$x \in [x_1^2, x_1^1).$$

Поскольку в этом случае все эксперименты выполнены, то на основании соотношений (1) устанавливаем

$$l([x_1^2, x_1^1)) = \delta\varphi(0, 1) = \delta. \quad (16)$$

С учетом соотношений (12, 13) устанавливаем

$$l([0, x_1^1)) = l([0, x_1^2)) + l([x_1^2, x_1^1)) = 2\delta. \quad (17)$$

Для исхода a_{212}) характерно то, что $x \in [x_1^2, x_1^1)$.

Поскольку в рассмотрении алгоритма остался один шаг, то на основании равенства (13) устанавливаем

$$l([x_1^2, x_1^1)) = \delta\varphi(5-4, 1) = \delta. \quad (18)$$

Поскольку соотношение (18) не противоречит соотношению (16), то этим обосновывается справедливость равенства (17).

Для исхода a_{22}) характерно то, что ВП могла действовать на первом либо на третьем шагах алгоритма.

В этом случае на четвертом шагах алгоритма применяем пессимистическую стратегию

$$x_1^4 = x_1^1.$$

При этом может возникнуть один из исходов

$$a_{221}) x + \xi(t_4) \leq x_1^1; a_{222}) x + \xi(t_4) > x_1^1.$$

Поскольку исход a_{22}) подтверждает исход a_2), то с учетом исхода a_{221}) устанавливаем

$$x \in [x_1^2, x_1^1).$$

Возникшее противоречие (см. исходы a), a_{22}), a_{221})) свидетельствует о проявлении ВП на третьем шаге алгоритма. Поскольку в рассмотренном алгоритме поиска осталось совершить один шаг алгоритма в условиях отсутствия ВП, то на пятом шаге алгоритма применяют дихотомию, посредством которой отрезок $[x_1^2, x_1^1)$ будет разбит на две равные части

$$l([x_1^2, x_1^1)) = 2\delta. \quad (19)$$

Исходя из минимального критерия, устанавливаем

$$l([x_1^2, x_1^1)) = \min\{\delta, 2\delta\} = \delta. \quad (20)$$

Для исхода a_{222}) характерно то, что ВП действовала на первом шаге алгоритма. Она не проявлялась на втором, третьем и четвертом шагах алгоритма и будет не проявляться на пятом шаге алгоритма (см. параметры ВП). Действуя на пятом шаге алгоритма, дихотомией разобьем полуоткрытый интервал неопределенности на две равные части.

На этом основании будет справедливым соотношение:

$$l([x_1^1, 1)) = 2\delta. \quad (21)$$

С учетом соотношений (17), (21) и симметрии ВП, устанавливаем

$$l([0, 1)) = 4\delta. \quad (22)$$

Поскольку исход b) симметричен исходу a), то и в этом случае отрезок $[0, 1]$ за пять шагов будет разбит на четыре равные части, а функция $\varphi(5, 1)$ при таком аргументе будет иметь значение

$$\varphi(5, 1) = 4. \quad (23)$$

Рассматривая процесс анализа для других значений i , к примеру $i = 6$, можно получить ориентированный граф переходов ЦАПП. Такой граф приведен на рис. 2, 3 ($i = 6$, $k = 1$; $0 \rightarrow 1$, $1 \rightarrow 0$; $a = 4$, $l = 1$, $h = 5$.)

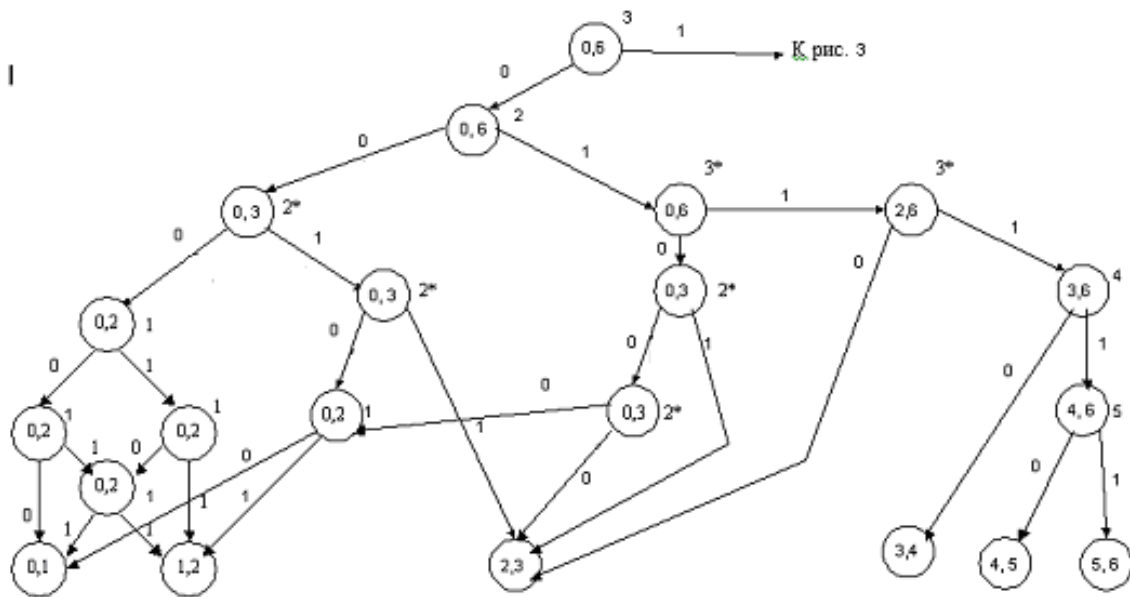


Рисунок 2 – Ориентированный граф переходов ЦАПП

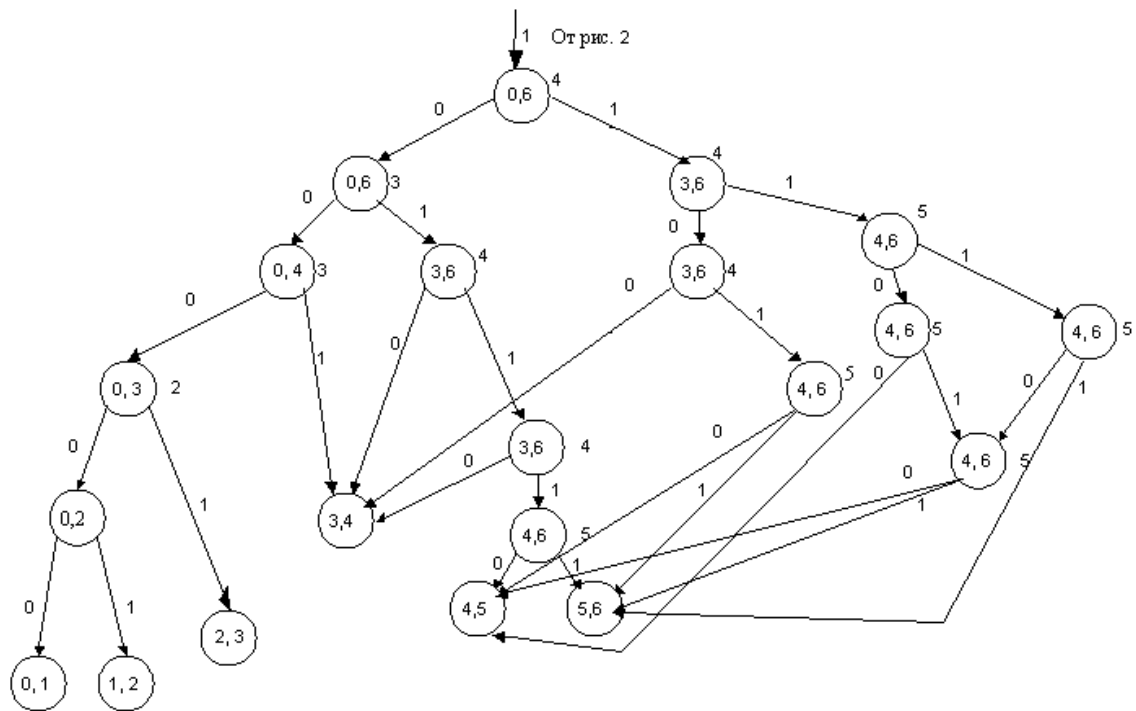


Рисунок 3 – Ориентированный граф переходов ЦАПП

ЗАКЛЮЧЕНИЕ

Показана возможность использования для синтеза алгоритмов функционирования цифровых автоматов с псевдослучайными переходами принципов «пересечения» и «повторных сравнений»; разработаны пра-

вила выбора стратегий функционирования таких автоматов, используя которые можно для конкретных параметров виртуальной последовательности разработать ориентированный граф переходов цифрового автомата с псевдослучайными переходами. Такие автоматы используются в системах защиты информации.

ПЕРЕЧЕНЬ ССЫЛОК

Надійшла 23.03.06

1. Алипов Н. В. Дискретные автоматы с псевдослучайными переходами и подстановочные методы защиты информации на их основе // Радиоэлектроника и информатика. – 2001. – № 4. – С. 95–98.
2. Алипов Н. В., Кораблев Н. М., Хиль М. И., Гусятин М. В. Структура цифрового автомата с псевдослучайными переходами из начального состояния в одно и тоже конечное состояние // Радиоэлектроника. Информатика. Управління. – 2006. – № 2. – С. 102–109.
3. Алипов Н. В. Помехоустойчивый поиск с характерным признаком и кодирование информации // Радиоэлектроника и информатика. – 2000. – № 4. – С. 82–86.
4. Альсведе Р., Вегенер И. Задачи поиска. – М.: Мир, 1982. – 368 с.
5. Алипов Н. В., Алипов И. Н., Хиль М. И., Ребезюк Л. Н. Последовательные алгоритмы поиска точки с характерными, помехоустойчивые к симметричному нерегулярными виртуальными последовательности // Радиотехника. – 2003. – Вып. 133. – С. 36–45.

Описано особливості застосування принципів «перетинання» і «повторних порівнянь» для рішення задачі синтезу алгоритмів функціонування цифрових автоматів із псевдоімовірними переходами; приведений приклад синтезу алгоритмів функціонування таких автоматів, що є генераторами псевдоімовірних підстановок у системах захисту інформації.

The specifics of using principles of «intersection» and «repeating comparisons» for solving the problem of synthesis of algorithms with pseudorandom transitions were described; the example of synthesis of the operation algorithm of such automates, which are the generators of pseudorandom replacements in information protection systems, was given.

УДК 004.057.2

А. О. Говоров, Г. И. Никулищев, Г. Л. Козина

АНАЛИЗ СТАНДАРТОВ ЭЛЕКТРОННОЙ ЦИФРОВОЙ ПОДПИСИ

В статье рассмотрены современные стандарты электронной цифровой подписи на эллиптических кривых ГОСТ Р 34.10-2001 (Россия) и ECDSA (США). Предложена реализация этих стандартов в специализированном математическом пакете, написаны процедуры базовых операций над точками эллиптической кривой. Показаны отличия стандартов во времени и процедурах формирования и проверки подписи.

ВВЕДЕНИЕ

Актуальность задач исследования современных асимметричных криптографических систем, в частности, систем электронной цифровой подписи (ЭЦП), обусловлена существующими проблемами анализа и разработки отечественных средств комплексной защиты информации. Решение этих задач позволяет обеспечить целостность и аутентичность информации, обрабатываемой и передаваемой в автоматизированных системах. Принятие национального стандарта ЭЦП ДСТУ 4145-2002, Закона Украины «Об электронной цифровой подписи» от 28 декабря 2003 г. [1] и ряда других законодательных и нормативных актов, которые регулируют деятельность в сфере электронного документооборота, способствует развитию инфраструктуры открытых ключей в Украине и ставит задачу разработки программного обеспечения для Центров Сертификации [1].

Так как электронный документооборот не ограничивается масштабами Украины, то для эффективной работы Центров Сертификации необходимо учитывать

и стандарты ЭЦП, принятые в других странах, в частности России и США.

ПОСТАНОВКА ЗАДАЧИ

В стандартах электронной цифровой подписи США (ECDSA) и России (ГОСТ Р 34.10-2001) на эллиптических кривых [2] в качестве математического аппарата используются эллиптические кривые над простым полем Галуа.

Перед авторами стояла задача проанализировать работу алгоритмов в каждом из стандартов, проверить их на совместимость в рамках единой программной реализации и разработать такую реализацию. Каждый из алгоритмов содержит три основных этапа:

1. Генерация ключей.
2. Формирование подписи.
3. Проверка подписи.

Поскольку в основе обоих стандартов лежит общий математический аппарат, их алгоритмы схожи, но на некоторых этапах наблюдаются различия.

АНАЛИЗ АЛГОРИТМОВ

Алгоритмы электронной цифровой подписи ECDSA и ГОСТ Р 34.10-2001 в целом схожи, но имеют небольшие отличия на некоторых шагах [2]. Сравнение этапов работы алгоритмов представлено в табл. 1–3.

Таблиця 1 – Сравнение этапов генерации ключей ECDSA и ГОСТ Р 34.10-2001

Генерация ключей	
ECDSA	ГОСТ Р 34.10-2001
<ol style="list-style-type: none"> 1. Выбираем эллиптическую кривую E, определенную на $GF(p)$. 2. Выбираем точку $P \in E(Zp)$ простого порядка n. 3. Выбираем случайное число $d \in [1, n - 1]$. 4. Вычисляем $Q = dP$. 5. Секретным ключом является d, открытым – (E, P, n, Q) 	<ol style="list-style-type: none"> 1. Выбираем эллиптическую кривую E, определенную на $GF(p)$. 2. Выбираем точку $P \in E(Zp)$ простого порядка n. 3. Выбираем случайное число $d \in [1, n - 1]$. 4. Вычисляем $Q = dP$. 5. Секретным ключом является d, открытым – (E, P, n, Q)

Таблиця 2 – Сравнение этапов формирования подписи ECDSA и ГОСТ Р 34.10-2001

Формирование подписи	
ECDSA	ГОСТ Р 34.10-2001
<ol style="list-style-type: none"> 1. Выбираем случайное число $k \in [1, n - 1]$. 2. Вычисляем $kP = (x_1, y_1)$ и $r = x_1 \bmod n$. Если $r \neq 0$, переходим к 3, если нет – возвращаемся к 1. 3. Вычисляем $k^{-1} \bmod n$. 4. Вычисляем $s = k^{-1}(h + dr) \bmod n$. Если $s \neq 0$, переходим к 5, если нет – к 1. 5. Подпись – пара целых чисел (r, s) 	<ol style="list-style-type: none"> 1. Выбираем случайное число $k \in [1, n - 1]$. 2. Вычисляем $kP = (x_1, y_1)$ и $r = x_1 \bmod n$. Если $r \neq 0$, переходим к 3, если нет – возвращаемся к 1. 3. Вычисляем $s = (kh + dr) \bmod n$. Если $s \neq 0$, переходим к 4, если нет – к 1. 4. Подпись – пара целых чисел (r, s)

Таблиця 3 – Сравнение этапов проверки подписи ECDSA и ГОСТ Р 34.10-2001

Проверка подписи	
ECDSA	ГОСТ Р 34.10-2001
<ol style="list-style-type: none"> 1. Если r и s не целые и r и $s \in [1, n - 1]$, подпись не верна. 2. Вычисляем $w = s^{-1} \bmod n$. 3. Вычисляем $u_1 = hw \bmod n$ и $u_2 = rws \bmod n$. 4. Вычисляем $u_1P + u_2Q = (x_0, y_0)$ и $v = x_0 \bmod n$. 5. Подпись верна, если $v = r$ 	<ol style="list-style-type: none"> 1. Если r и s не целые и r и $s \in [1, n - 1]$, подпись не верна. 2. Вычисляем $w = h^{-1} \bmod n$. 3. Вычисляем $u_1 = sw \bmod n$ и $u_2 = -rw \bmod n$. 4. Вычисляем $u_1P + u_2Q = (x_0, y_0)$ и $v = x_0 \bmod n$. 5. Подпись верна, если $v = r$

Как видно из табл. 1, процедуры генерации секретного и открытого ключей в обоих алгоритмах совпадают.

На втором этапе (постановки подписи) видно различие на 3-м и 4-м шаге. В стандарте ECDSA необходимо выполнять дополнительное вычисление (шаг 3) – обращать элемент группы по модулю. Несмотря на то, что алгоритм вычисления обратного элемента полиномиальный, для большого значения модуля ($n > 2^{256}$) будет заметно замедление алгоритма формирования подписи по сравнению с ГОСТом, что в дальнейшем подтверждается при тестировании алгоритмов

Процессы проверки подписи вычислительно совпадают за исключением промежуточных входных параметров.

РЕАЛИЗАЦИЯ СТАНДАРТОВ

Основной проблемой при реализации стандартов электронной цифровой подписи является генерация параметров [5]. Ни в одном из рассматриваемых стандартов не приводится рекомендаций по выбору параметров алгоритмов. К сожалению в работе [5] приводится только контрольный пример из ГОСТа.

Для рассмотренных стандартов электронной цифровой подписи на эллиптических кривых авторами создана программная реализация, которая включает в себя базовые процедуры алгоритмов на эллиптических кривых, процедуру генерации ключей, процедуру формирования подписи, процедуру проверки подписи.

Проведено тестирование данных процедур для входных параметров различной длины. В том числе рассмотрен контрольный пример из стандарта ГОСТ Р 34.10-2001 [2], который представлен на рис. 1.

В виде отдельных функций реализованы базовые алгоритмы операций на эллиптических кривых:

- вычисление координат удвоения точек эллиптической кривой;
- вычисление координат суммы различных точек эллиптической кривой;
- вычисление скалярного умножения точки эллиптической кривой на число, с использованием комбинации процедур удвоения и суммирования;
- определение порядка эллиптической кривой;
- поиск базовой точки группы точек эллиптической кривой заданного порядка;
- определение порядка точки эллиптической кривой.

С использованием этих функций осуществляются процедуры генерации ключей (открытого и закры-

ВЫВОДЫ

Проанализированы современные стандарты электронной цифровой подписи США и России. Эти стандарты являются программно совместимыми, т. к. вычислительные процедуры осуществляются над простым полем Галуа в отличие от Стандарта ЭЦП Украины. Программная реализация допускает использование обоих стандартов в одном пакете. Поскольку допустимый в данной реализации диапазон значений основных параметров алгоритмов ECDSA и ГОСТ Р 34.10-2001 существенно меньше требуемого стандартами, то предлагаемая реализация не может быть использована в коммерческих целях, но вполне пригодна для использования в учебном процессе, так как структура кода достаточно прозрачна и позволяет наглядно проиллюстрировать работу как стандартов ECDSA и ГОСТ Р 34.10-2001 в целом.

ПЕРЕЧЕНЬ ССЫЛОК

1. Про електронний цифровий підпис: Закон України № 852-IV від 22. 05. 2002 р. // Відомості Верховної Ради. – 2003. – № 36. – С. 276.

2. Бессалов А. В., Телиженко А. Б. Криптосистемы на эллиптических кривых: Учебное пособие. – К.: ІВЦ Видавництво «Політехніка», 2004. – 224 с.
3. Иванов М. А. Криптографические методы защиты информации в компьютерных системах и сетях. – М.: КУДИЦ-ОБРАЗ, 2001. – 368 с.
4. Математические и компьютерные основы криптологии: Учеб. пособие / Ю. С. Ханин, В. И. Берник, Г. В. Матвеев, С. В. Агиевич. – Мн.: Новое знание, 2003. – 382 с.
5. Горбенко И. Д., Качко Е. Г., Колесников П. В. Генерация параметров и ключей для цифровой подписи на эллиптических кривых для конечного простого поля // Радиотехника. – 2002. – № 125. – С. 44–49.

Надійшла 20.06.06
Після доробки 20.10.06

В статті проводиться аналіз сучасних стандартів електронного цифрового підпису на еліптичних кривих. Розроблено реалізацію стандартів ECDSA та ГОСТ Р 34.10-2001 в спеціалізованому математичному пакеті з написання окремих процедур для базових операцій над точками еліптичної кривої. Наведені відмінності у часі і процедурах формування і перевірки підпису.

The article deals with the analysis of electronic digital signature standards on elliptic curves. Several standards (ECDSA and ГОСТ Р 34.10-2001) are realized by means of special mathematical software. Testing of the realization and standard's analysis are also provided.

УДК 004.8

А. С. Горецкий

МУЛЬТИАГЕНТНЫЕ ТЕХНОЛОГИИ В СИСТЕМАХ УПРАВЛЕНИЯ ПЕРСОНАЛОМ

В работе рассмотрены особенности агентных технологий и перспективы их использования для разработки программных систем. Предложен агентный подход к созданию компьютерных систем управления персоналом. Рассмотрена модель системы управления профильного отбора персонала.

ВВЕДЕНИЕ

Кадровый учет и управление персоналом являются важными аспектами в деятельности любой организации. Для успешной работы в условиях рынка сегодня на предприятиях наравне с оценкой потребности в финансах, оборудовании и материалах определяется потребность в людских ресурсах.

Для контроля людских ресурсов наиболее эффективным решением является использование автоматизированной системы управления персоналом. Первоначально системы управления персоналом предназначались для автоматизации процесса управления фондом заработной платы и выплаты поощрений. Однако

со временем функциональные возможности таких систем существенно расширились. Система управления персоналом позволяет значительно облегчить и ускорить поиск квалифицированного персонала и максимально эффективно его использовать для достижения целей организации.

Имеющиеся на рынке автоматизированные системы управления персоналом по их функциональной направленности можно разделить на следующие группы [1]:

- многофункциональные экспертные системы, позволяющие проводить профориентацию, отбор, аттестацию конкретных работников;
- экспертные системы для группового анализа персонала, выявления тенденций развития подразделений и организации в целом;
- программы расчета зарплаты;
- комплексные системы управления персоналом, позволяющие формировать и вести штатное расписание, хранить полную информацию о сотрудниках, от-

ражать движение кадров внутри фирмы, рассчитывать зарплату.

Однако существующие автоматизированные системы управления персоналом обладают не достаточно эффективным функциональным аппаратом анализа персонала. Поэтому вопросы повышения эффективности в задачах профильного отбора, профориентации, анализа персонала являются достаточно актуальными на сегодняшний день. Решения этих вопросов можно достичь посредством усовершенствования методов анализа информации и внедрением технологий искусственного интеллекта в системах управления персоналом.

ПОСТАНОВКА ЗАДАЧИ

Целью данной работы является анализ концепции применения мультиагентных технологий для создания эффективных компьютерных систем управления персоналом.

В работе выбраны наиболее перспективные отрасли использования мобильных агентов в компьютерных системах управления персоналом – многофункциональные экспертные системы, позволяющие проводить профориентацию, отбор, аттестацию конкретных работников; системы получения информации из хранилищ, сортировки, классификации и фильтрации данных о персонале. Этот подкласс систем будут отличать базовые принципы поддержки принятия решений, основанные на элементах искусственного интеллекта.

Задача данного исследования – разработать прототип мультиагентной системы профильного отбора персонала. Система должна быть кроссплатформенной и удовлетворять условиям модульности, расширяемо-

сти, а также должна быть совместима со спецификациями международной организации FIPA (Foundation for Intelligent Physical Agents).

МОДЕЛЬ АГЕНТНОЙ ПОДСИСТЕМЫ ПО ПРОФИЛЬНОМУ ОТБОРУ ПЕРСОНАЛА

Автором разработана модель агентной подсистемы управления профильного отбора персонала [2]. Данная подсистема может существовать как полноценная независимая система из класса систем управления персоналом или входить в качестве функционального модуля в комплексную систему управления персоналом. Разработанная модель представлена на рис. 1, где:

- Internet – группа устройств, предоставляющих доступ к глобальной сети Интернет;
- Mobile Agent Server – сервер мобильных агентов;
- Local Network – устройства локальной сети, обеспечивающие связь серверов;
- DataBase Server – сервер БД (например, MS SQL Server 2000, Oracle9i, DB/2 Sybase, FireBird), который содержит информацию о персонале и т. п.;
- Agent DB, Agent Manager – программные агенты системы, функционирующие на сервере мобильных агентов;
- USER – компьютер пользователя системы, который использует мобильных агентов для получения информации, необходимой для принятия решения. Благодаря особенностям технологии мультиагентных систем, данный компьютер может находиться как в локальной сети организации, так и в удаленном офисе, и связываться с сервером мобильных агентов посредством Интернет.

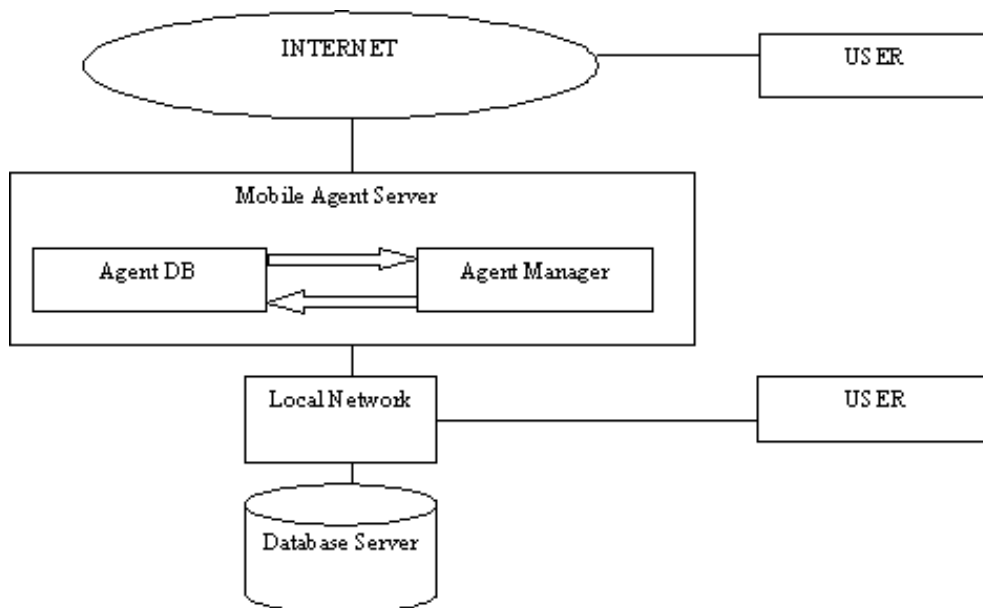


Рисунок 1 – Модель агентной системы управления по профильному отбору персонала

ВЫБОР ИНСТРУМЕНТАРИЯ РАЗРАБОТКИ СИСТЕМЫ

На данный момент разработано несколько агентных платформ, ориентированных на использование спецификации FIPA-2000: ZEUS компании BTextact Technologies (Великобритания), Comtec Agent Platform компании Comtec (Япония), JADE компании CSELT (Италия), AAR компании Fujitsu Labs (США), FIPA-OS компании Nortel Networks (Великобритания) [3].

В качестве программно-инструментального средства для программной реализации мультиагентной системы управления персоналом выбрана JADE как одна из агентных платформ, наиболее полно поддерживающих спецификацию FIPA-2000.

Выбор данной платформы обусловлен также тем, что она полностью написана на языке программирования Java с использованием таких продвинутых возможностей как Java RMI, Java CORBA IDL, Java Serialization и Java Reflection API, что позволяет обеспечить кроссплатформенность разрабатываемой системы управления персоналом.

JADE упрощает разработку мультиагентных систем, благодаря использованию FIPA-спецификаций. Эта агентная платформа может распространяться среди компьютеров с разными операционными системами, и ее можно конфигурировать посредством удаленного GUI-интерфейса. На рис. 2 представлено окно управления агентной платформой JADE.

Процесс конфигурирования этой платформы достаточно гибкий: ее можно изменить даже во время исполнения программ, для этого необходимо просто переместить агентов с одной машины на другую. Единственным требованием этой системы является наличие установленного программного обеспечения JRE (Java Runtime Environment) [4].

РЕАЛИЗАЦИЯ АГЕНТНОЙ СИСТЕМЫ ПО ПРОФИЛЬНОМУ ОТБОРУ ПЕРСОНАЛА

На основе предложенной модели средствами языка Java и агентной платформы JADE разработан прототип программной системы, выполняющий функции профильного отбора персонала.

Разработанный прототип может функционировать как в качестве независимого приложения, так и составе комплексной системы управления персоналом.

Основной функционал системы представлен агентами двух типов: Агент базы данных (Agent DB) и Агент-менеджер (Manager Agent). Данные агенты функционируют в качестве отдельных процессов на сервере мобильных агентов. Следует отметить, что благодаря использованию платформы JADE, данные агенты могут быть физически распределены между несколькими серверами, то есть агент-менеджер, запущенный в одном департаменте компании, может взаимодействовать с агентом базы данных, который запущен в другом департаменте, географически удаленном. Взаимодействие между программными агентами происходит посредством сообщений, которые кодируются в соответствии со спецификацией «FIPA ACL Message Structure Specification». ACL (Agent Communication Language) – это язык общения программных агентов, который является мировым стандартом. Формат ACL сообщений представлен в табл. 1. Количество элементов в одном сообщении может быть разным, в зависимости от ситуации, но всегда обязательными элементами сообщениями являются такие как performative, sender, receiver и content.

Если агент встречает в полученном сообщении неизвестные значения элементов сообщения, то он отвечает отправителю специальным сообщением non-understood.

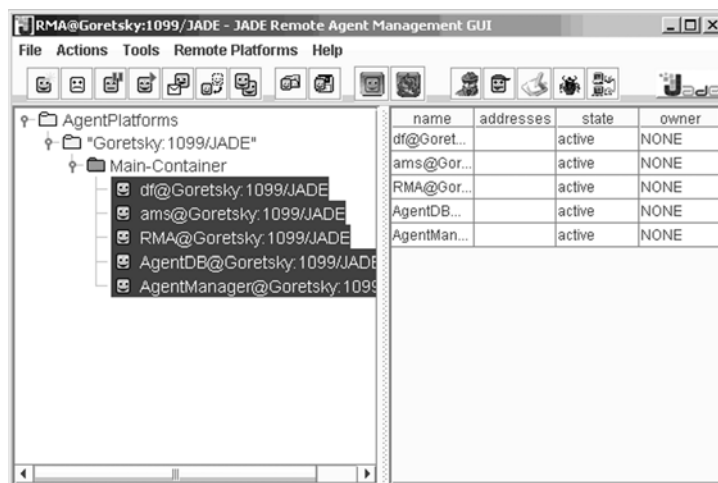


Рисунок 2 – GUI-интерфейс управления агентной платформой

Таблица 1 – Элементы сообщения FIPA ACL

Элемент	Категория элемента
Performative	тип речевого акта
Sender	агент-участник коммуникации
Receiver	агент-участник коммуникации
reply-to	агент-участник коммуникации
Content	содержание сообщения
Language	описание сообщения
Encoding	описание сообщения
Ontology	описание сообщения
Protocol	управление разговором
conversation-id	управление разговором
reply-with	управление разговором
in-reply-to	управление разговором
reply-by	управление разговором

При создании агентов системы профильного отбора персонала для описания их активности использовались классы JADE, в частности базовый класс Agent, который реализует функции взаимодействия с агентной платформой (регистрация, конфигурация, удаленное управление и др.). Для описания поведения агента использовались классы SimpleBehaviour и CyclicBehaviour, а также набор методов для организации поведения программного агента (например, послать/принять сообщения, использование стандартных протоколов взаимодействия, регистрация в нескольких каталогах и т. д.).

На языке Java описываются и переопределяются классы, реализующие логику поведения агентов. Схема взаимодействия программных агентов разрабатываемой системы представлены на рис. 3.

При запуске системы профильного отбора персонала на агентной платформе JADE одновременно создаются агент-менеджер и агент базы данных. Сразу после успешного запуска агенты переходят в режим ожидания запроса от пользователя.

Пользователь через графический интерфейс (см. рис. 4) делает запрос к агенту-менеджеру о требуемых качествах персонала. Агент-менеджер получает запрос и формирует на основе данных запроса критерии отбора. Затем эти критерии инкапсулируются в ACL-сообщение в элемент «content» и передаются агенту базы данных. После этого агент-менеджер переходит в режим ожидания до тех пор, пока не получит ответа от агента базы данных.

Сразу после создания агент базы данных переходит в режим ожидания запроса. При получении ACL-сообщения агент базы данных считывает критерии отбора, инкапсулированные в элементе «content», и формирует запрос к базе данных. Для доступа к базе данных используется технология JDBC, что позволяет агенту базы данных работать с любыми СУБД, для которых существует специальный jdbc-драйвер. После получения данных от СУБД агент базы данных инкапсулирует результаты в ACL-сообщение и передает его агенту-менеджеру.

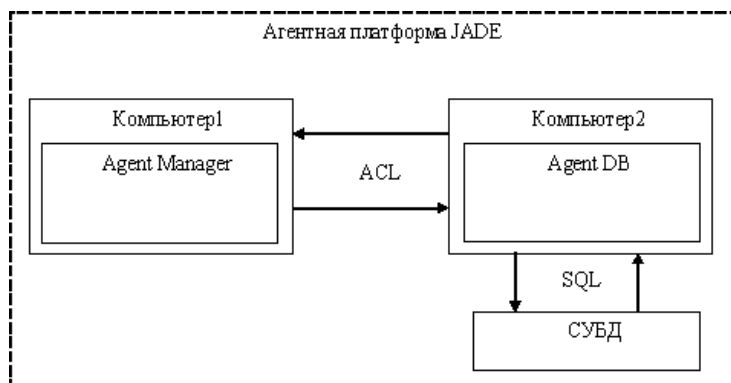


Рисунок 3 – Схема взаимодействия программных агентов в системе профильного отбора персонала

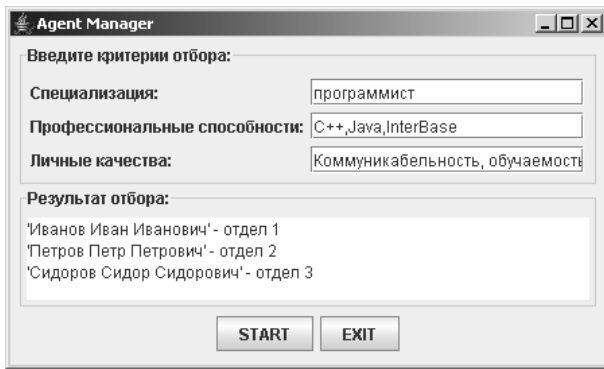


Рисунок 4 – Графический интерфейс агента менеджера

Агент-менеджер декодирует ACL-сообщение, полученное от агента базы данных, и выводит декодированные данные пользователю через графический интерфейс.

ЗАКЛЮЧЕНИЕ

В работе проанализированы теоретические и практические особенности использования агентных технологий в компьютерных системах управления персоналом. Рассмотрена система управления персоналом как мультиагентная система, разработана модель агентной системы управления профильного отбора персонала на базе агентной архитектуры JADE, на основе модели разработан прототип программной системы, выполняющий функции профильного отбора персонала. Разработанный прототип обладает на сегодняшний день базовыми функциями профильного отбора персонала в рамках системы управления предприятием. Но в перспективе система позволит охватить все этапы тести-

рования, аттестации и отбора персонала в рамках предприятия. Таким образом, применение технологий распределенного искусственного интеллекта позволит повысить эффективность принимаемых решений по управлению кадрами.

Данная работа является начальным этапом исследования концепции единого информационного пространства организации. В дальнейшем планируется исследование методов интеграции информационных систем и разработка методики создания единого информационного пространства на основе мультиагентного подхода.

ПЕРЕЧЕНЬ ССЫЛОК

1. Системы управления персоналом [Электронный ресурс] / О. Тютина – Электрон. журн. – Computerworld, № 18, М., 2000. – Режим доступа: <http://www.osp.ru/text/302/4672/> свободный. – Загл. с экрана.
2. Гаврилов Т. А., Хорошевский В. Ф. Базы знаний интеллектуальных систем. – СПб: Питер, 2000. – 384 с.
3. Келеберда И. Н., Лесная Н. С., Репка В. Б. Использование мультиагентного онтологического подхода к созданию распределенных систем дистанционного обучения // Educational Technology & Society. – 2004. – № 7(2). – С. 6–11.
4. Bellifemine F., Caire G., Trucco T., Rimassa G. JADE Programmer's Guide. – 2005. – 52 с.

Надійшла 18.01.07

У роботі розглянуто особливості агентних технологій і перспективи їх використання для розробки програмних систем. Запропоновано агентний підхід створення комп'ютерних систем управління персоналом. Розглянута модель системи управління профільного відбору персоналу.

In this work the special features of agent technologies and the prospects of it use for developed of program system were considered. The agent approach of computer system of Human Resource construction was proposed. The model of system of profile selection of the personnel was considered.

УДК 004.94

В. И. Дубровин, Ю. С. Афонин, Р. Н. Зинченко, С. С. Емельяненко,
Т. В. Манило

ПОДХОД К ПОСТРОЕНИЮ МОДЕЛИ ПО ЗВУКОВОМУ СИГНАЛУ С ИСПОЛЬЗОВАНИЕМ ВЕЙВЛЕТ-ПРЕОБРАЗОВАНИЯ

Представлен подход к построению модели по звуковому сигналу, являющийся комбинацией методов пакетного вейвлет-преобразования для обработки звукового сигнала и искусственной нейронной сети для оценки диагностируемых параметров. Эффективность предложенного подхода исследована на примере задачи диагностики износа режущего инструмента.

© Дубровин В. И., Афонин Ю. С., Зинченко Р. Н., Емельяненко С. С., Манило Т. В., 2007

ВВЕДЕНИЕ

Построение моделей диагностирования различных объектов и процессов является актуальной задачей. Многие объекты в процессе своей работы сопровождаются звуковым излучением, которое содержит инфор-

мацию о текущем состоянии диагностируемого объекта или процесса.

Целью работы является рассмотрение и анализ подхода к построению модели по звуковому сигналу на примере задачи диагностики износа режущего инструмента (РИ).

1 ЭТАПЫ ПОСТРОЕНИЯ МОДЕЛИ

Построение модели по звуковому сигналу исследуемого процесса состоит из следующих основных этапов:

- обработка звукового сигнала и последующее получение информативных характеристик сигнала;
- установление взаимосвязи между отобранными характеристиками и диагностируемыми параметрами.

В качестве основного инструмента обработки звукового сигнала предлагается использовать вейвлет-преобразование, а для определения взаимосвязи между результатами вейвлет-преобразования и диагностируемыми параметрами – искусственная нейронная сеть.

2 БЫСТРОЕ И ПАКЕТНОЕ ВЕЙВЛЕТ-ПРЕОБРАЗОВАНИЕ

Вейвлет-преобразование (ВП) является эффективным математическим аппаратом анализа нестационарного по своей природе звукового сигнала.

Построение базиса ВП состоит в использовании масштабного преобразования и смещений. Именно за счет изменения масштабов ВП способно выявить различия в характеристиках на разных шкалах, а путем сдвига проанализировать свойства сигнала в разных точках на всем изучаемом интервале [1]. Результатом ВП являются аппроксимирующие и детализирующие коэффициенты, которые в общем случае можно вычислить по формулам (1) и (2), соответственно.

$$c_{j,k} = \frac{1}{\sqrt{2^j}} \sum_{t=0}^L s(t) \psi\left(\frac{t-2^j k}{2^j}\right), \quad (1)$$

$$d_{j,k} = \frac{1}{\sqrt{2^j}} \sum_{t=0}^L s(t) \phi\left(\frac{t-2^j k}{2^j}\right), \quad (2)$$

где $c_{j,k}$, $d_{j,k}$ – коэффициенты аппроксимации и детализации, j – параметр масштаба, k – параметр сдвига, $s(t)$ – исходный сигнал, L – длина сигнала, ψ – вейвлет функция, ϕ – масштабирующая функция.

Модификациями ВП является быстрое вейвлет-преобразование (БВП) и пакетное вейвлет-преобразование (ПВП).

При БВП выполняются быстрые итерационные вычисления, в результате которых коэффициенты аппроксимации и детализации могут быть найдены, исполь-

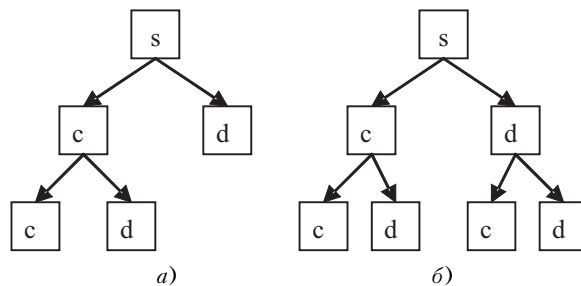


Рисунок 1 – Схемы вейвлет-преобразования:
а – БВП, б – ПВП

зуя аппроксимирующие коэффициенты предыдущего уровня [2]. Схема таких вычислений представлена на рис. 1, а.

В ПВП бинарное дерево разложения (рис. 1, б) «расщепляется» и для детализирующих коэффициентов любого уровня. Такое расщепление позволяет выделить как низкочастотные, так и высокочастотные составляющие сигнала.

3 ОБРАБОТКА ЗВУКОВОГО СИГНАЛА С ПОМОЩЬЮ ПВП

Обработка сигнала – процесс преобразования сигнала и оценивание его информативных характеристик.

Для анализа звукового сигнала выбрано ПВП, поскольку операция декомпозиции может применяться как к низкочастотной, так и высокочастотной компоненте разложения [3].

Рассмотрим звуковой сигнал процесса резания и его изменение при различной степени износа РИ (рис. 2).

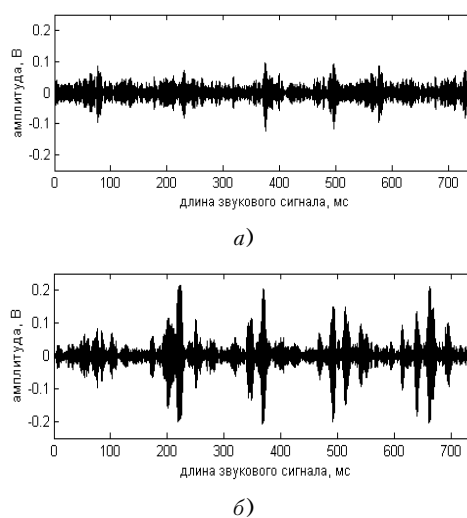


Рисунок 2 – Изменение звукового сигнала с увеличением износа инструмента:

а – низкая степень износа, б – высокая степень износа

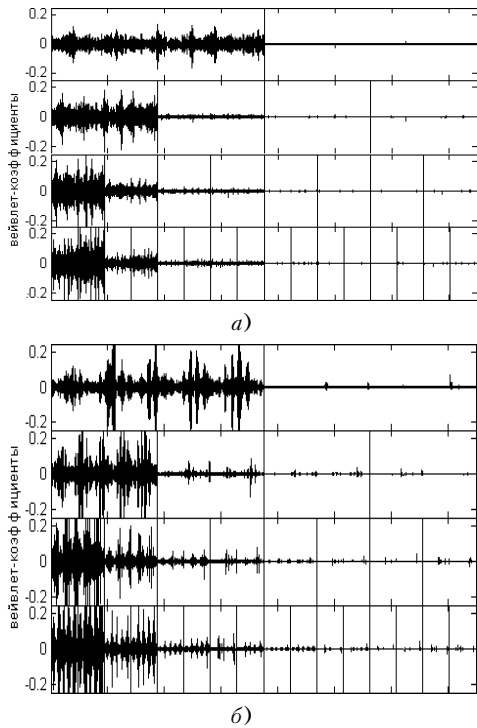


Рисунок 3 – Вейвлет-коэффициенты ПВП звуковых сигналов при:
а – низкой, б – высокой степени износа инструмента

На рис. 3 показаны результаты ПВП звуковых сигналов, представленных на рис. 2, соответственно.

На четвертом уровне разложения ПВП имеем 16 окон, в которых показаны декомпозированные результаты сигнала в частотных диапазонах.

Среднеквадратические значения (СКЗ) вейвлет-коэффициентов каждого окна СКЗ-1, СКЗ-2, ..., СКЗ-16 определяют энергетический уровень звукового сигнала в соответствующем частотном диапазоне.

В качестве информативных характеристик берутся СКЗ коэффициентов окон, наиболее чувствительных к износу. В данном случае такими являются СКЗ-2,

СКЗ-3, СКЗ-5, СКЗ-6, СКЗ-7, СКЗ-9. Для удобства представим их как x_1, x_2, \dots, x_6 , соответственно.

4 ОЦЕНКА ИЗНОСА С ПОМОЩЬЮ ИСКУССТВЕННОЙ НЕЙРОННОЙ СЕТИ

Для установления взаимосвязи между информативными характеристиками звукового сигнала и степенью износа РИ используется искусственная нейронная сеть (ИНС), поскольку ИНС позволяет аппроксимировать сложные зависимости [4].

Для построения модели была выбрана следующая структура нейронной сети: два слоя нейронов, в первом слое – 6 нейронов, на втором – 1 нейрон, сигмоидные функции активации для нейронов первого и второго слоя. Для обучения сети использовался метод Левенберга-Марквардта.

Для обучения ИНС на вход подавались значения информативных характеристик x_1, x_2, \dots, x_6 , а на выход – степень износа РИ.

В результате обучения нейронной сети была получена модель представленная в таблице 1.

В результате тестирования полученной модели точность оценки износа инструмента составила 97,3 %.

ЗАКЛЮЧЕНИЕ

Рассмотренный подход к построению модели представляет собой комбинацию метода пакетного вейвлет-преобразования для обработки звукового сигнала и искусственной нейронной сети для оценки диагностируемых параметров.

Модель износа инструмента, построенная на основе рассмотренного подхода, показала высокую точность оценки степени износа режущего инструмента, из чего можно сделать вывод об эффективности предложенного подхода.

Перспективным является интеграция представленного метода в интерактивную систему диагностики и контроля износа инструмента.

Таблица 1 – Нейросетевая модель оценки износа режущего инструмента

Слой	1						2
№ нейрона в слое	1	2	3	4	5	6	1
Весовые коэффициенты	13.3759	-3.018	-5.2558	3.4364	-9.009	2.804	-7.4461
	1.0645	0.92168	72.9234	-32.0328	133.1599	0.11889	-1.9372
	0.054435	-0.00956	-24.4895	-4.1369	-15.2134	-0.03587	58.6086
	0.066867	0.012423	-42.8313	1.7003	14.5045	0.019599	12.1296
	-0.03743	0.29665	23.0538	-68.7071	-63.4904	-0.01181	-48.686
	-0.01810	-0.04274	-1.8792	-0.50713	25.3585	-0.00397	-2.0132
Пороговые значения	433757	-4.9022	20.7449	-20.9038	21.848	6.0621	-2.6703

ПЕРЕЧЕНЬ ССЫЛОК

1. Новиков Л. В. Основы вейвлет-анализа сигналов. Учебное пособие. – СПб., МОДУС+, 1999. – 152с.
2. Переберин А. В. О систематизации вейвлет-преобразования // Вычислительные методы и программирование. – 2001. – Том 2. – С. 15–40.
3. Jaideva C. Goswami, Andrew K. Chan. Fundamentals of wavelets. Theory, algorithms and applications. – Wiley Interscience, 1999. – 324 p.
4. Круглов В. В., Дли М. И., Голунов Р. Ю. Нечеткая логика и искусственные нейронные сети: Учеб. пособие. – М.: Физматлит, 2001. – 224 с.

Надійшла 22.01.07

Представлений підхід до побудови моделі по звуковому сигналу, який являє собою комбінацію методів пакетного вейвлет-перетворення для обробки звукового сигналу та штучної нейронної мережі для оцінки параметрів, що діагностуються. Ефективність запропонованого підходу досліджена на прикладі задачі діагностики зносу ріжучого інструмента.

A way of model definition by acoustic signal is presented, which is the combination of wavelet packet transform for acoustic signal processing and artificial neural network for diagnostic parameter estimation. The effectiveness of proposed approach for the task of tool wear diagnostics is investigated.

УДК 001:002

І. О. Романенко, С. В. Алексєєв, І. В. Рубан, С. А. Войтович, В. В. Калачова

ДИСТАНЦІЙНЕ НАВЧАННЯ В ОСВІТНЬОМУ ПРОЦЕСІ ЗБРОЙНИХ СИЛ УКРАЇНИ

На основі аналізу структури існуючої системи дистанційного навчання (СДН) Міністерства освіти і науки України та положень Концепції створення системи дистанційного навчання у Збройних Силах України запропоновано варіант загальної структури інтегрованої СДН Міністерства освіти і науки та Міністерства оборони України.

ВСТУП

Глобальна інформатизація, що відбувається у сучасному світовому суспільстві, розвиток телекомунікаційних технологій та засобів обчислювальної техніки зумовлюють суттєві зміни форм освітнього процесу у всьому світі. При цьому зростає роль професійної та безперервної освіти без відриву від професійної діяльності, що сприяє розвитку та впровадженню дистанційних форм навчання [1–5].

Актуальність дослідження. Впровадження дистанційного навчання у процес підготовки та перепідготовки військових фахівців Збройних Сил України (ЗСУ) зумовлено низкою причин:

- інтенсивність розвитку військових наук потребує постійного вдосконалення професійних знань та навичок військовослужбовців;
- сучасне озброєння та військова техніка базуються на використанні новітніх електронних та комп'ютерних технологій;
- перехід ЗСУ до професійної армії на контрактній основі потребує якісної та гнучкої системи підготовки висококваліфікованих кадрів;

– тільки технології дистанційного навчання спроможні забезпечити своєчасне корегування змісту навчання військових фахівців за рахунок високої швидкості оновлення знань в інформаційно-освітньому середовищі;

– висока економічна ефективність дистанційного навчання.

Однією зі складових впровадження дистанційного навчання у ЗСУ є створення сучасного інформаційно-освітнього середовища для підготовки, перепідготовки й підвищення кваліфікації військових фахівців та надання військовослужбовцям, незалежно від місця проходження служби, рівних умов доступу до освітніх ресурсів на базі використання новітніх інформаційних технологій та телекомунікаційних мереж Міністерства оборони (МО) України, а також телекомунікаційної науково-освітньої мережі установ Міністерства освіти і науки (МОН) України з доступом до Інтернет [4].

Дистанційна форма навчання в ЗСУ може бути застосована:

- при отриманні військової освіти за всіма напрямками підготовки;
- при перепідготовці та підвищенні кваліфікації фахівців різних професійних напрямів і рівнів кваліфікації;
- при вивченні окремих навчальних курсів, які можуть викликати інтерес до вдосконалення знань і вмінь у військовослужбовців (ділова українська мова, іноземна мова, військове право, менеджмент, економіка тощо);

- при підготовці науково-педагогічних кадрів;
- при військовій підготовці студентів вищих навчальних закладів за програмою офіцерів запасу, особливо тих, що перебувають за межами розташування тих військових навчальних закладів, які мають право проводити таку підготовку;
- при попередній підготовці та тестуванні кандидатів до вступу до військових навчальних закладів;
- при навчанні іноземних військовослужбовців;
- для надання військовослужбовцям цивільної освіти.

Мета статті. Впровадження дистанційного навчання у ЗСУ повинно ґрунтуватися на дидактичних, методологічних, організаційних та інформаційно-телекомунікаційних засадах, що потребує створення системи дистанційного навчання (СДН), яка є частиною системи освіти України [4]. Таким чином необхідно розробити структуру військової системи дистанційного навчання, як складової загальнодержавної системи дистанційної освіти України. Для цього слід застосувати методи системного аналізу, агрегування, реструктуризації, деконпозиції, багатокритеріального синтезу. Завдяки такій структурі можливо досягнення значної економії коштів за рахунок використання високоякісних навчальних програм і технологій для широкого кола військовослужбовців, експлуатації телекомунікаційних та інформаційних ресурсів в межах єдиного інформаційного середовища на базі корпоративної

мережі, скорочення витрат на утримання будинків і гуртожитків, зменшення витрат на відрядження.

Задача дослідження полягає в аналізі існуючої структури системи дистанційного навчання Міністерства освіти і науки України та розробці організаційної структури СДН методами системного аналізу та багатокритеріального синтезу.

ОСНОВНИЙ МАТЕРІАЛ

У сфері цивільної освіти на цей час основним нормативно-правовим актом, що регламентує процес впровадження дистанційного навчання є «Положення про дистанційне навчання» [2], згідно якого структура СДН включає в себе (рис. 1):

- Координаційну раду з розвитку дистанційного навчання при Міністерстві освіти і науки, яка забезпечує координацію робіт з формування і реалізації державної політики щодо розвитку СДН;
- комісії при Координаційній раді за окремими напрямками забезпечення розвитку СДН, що вирішують питання розвитку СДН за такими основними напрямками: нормативно-правового і організаційного, науково-методичного забезпечення; системотехнічного забезпечення та стандартизації дистанційного навчання; матеріально-технічного, кадрового забезпечення; моніторингу якості, експертизи і сертифікації дистанційного навчання;

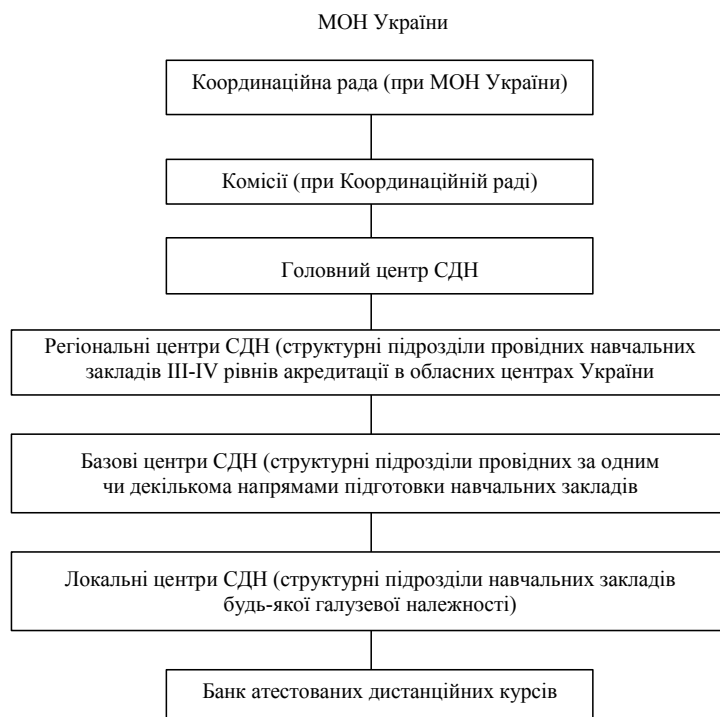


Рисунок 1 – Загальна структура системи дистанційного навчання Міністерства науки і освіти України

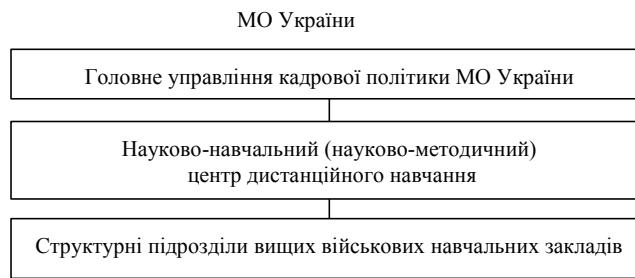


Рисунок 2 – Загальна структура системи дистанційного навчання Міністерства оборони України

– головний, регіональні, базові та локальні центри СДН;

– банк атестованих дистанційних курсів, що об'єднані між собою спеціалізованими інформаційно-комунікаційними мережами.

Функції кожного з елементів СДН визначені в [2].

Атестація дистанційних курсів здійснюється Координаційною радою у порядку, що визначається Міністерством освіти і науки України, за поданням авторів дистанційних курсів чи осіб, які репрезентують їх інтереси [2].

Для забезпечення впровадження дистанційної форми навчання, підтримки її функціонування та розвитку, доцільно включити до складу організаційної структури СДН ЗСУ [4] наступні компоненти (рис. 2):

– Головне управління кадрової політики МО України;

– науково-навчальний (науково-методичний) центр дистанційного навчання МО України;

– структурні підрозділи вищих військових навчальних закладів (ВВНЗ) з організації дистанційного навчання.

Загальне керівництво щодо впровадження дистанційної форми навчання у Збройних Силах України може здійснювати Головне управління кадрової політики МО України, яке буде забезпечувати [4]:

– формування нормативно-правової бази СДН ЗСУ;

– розробку механізму фінансування та створення матеріально-технічної бази СДН ЗСУ;

– формування і контроль за організаційною структурою СДН ЗСУ, включаючи всі її компоненти;

– координацію діяльності структурних складових СДН ЗСУ;

– координацію міжнародної діяльності СДН ЗСУ.

До функцій науково-методичного (науково-навчального) центру дистанційного навчання може належати [4]:

– підготовка проектів нормативно-правових документів СДН ЗСУ;

– розробка єдиних вимог щодо навчальних планів, програм і нормативів системи дистанційного навчання виходячи з державних стандартів освіти;

– розробка теоретичних і науково-психологічних засад дистанційного навчання;

– розробка рекомендацій щодо впровадження інформаційних технологій і дистанційних курсів у різні форми навчання;

– координація розробок та впровадження дистанційних курсів;

– створення розподіленої інформаційної структури СДН ЗСУ та системи її адміністрування;

– проведення навчання та перепідготовки кадрів в галузі дистанційної навчання: викладачів-т'юторів, адміністраторів, розробників курсів тощо;

– участь у створенні електронної бібліотеки дистанційних курсів;

– проведення попередньої експертизи усіх складових СДН ЗСУ, включаючи рекомендації щодо акредитації військових закладів освіти у реалізації дистанційної форми навчання і сертифікації окремих дистанційних курсів.

Структурні підрозділи ВВНЗ з організації дистанційного навчання створюються на базі вищих військових навчальних закладів, що мають доступ до телекомунікаційних мереж, сучасну комп'ютерну базу та підготовлений кадровий склад, та здійснюють навчання за дистанційними технологіями відповідно до ліцензованої освітньої діяльності. До функцій цих структурних підрозділів відносяться [4]:

– участь у підготовці проектів нормативно-правових документів СДН ЗСУ;

– розробка методик навчання за напрямками підготовки фахівців;

– розробка дистанційних курсів;

– створення електронної бібліотеки дистанційних курсів за напрямками підготовки;

– сприяння перепідготовці своїх кадрів для участі у СДН і розповсюдженню дистанційних технологій навчання в інших формах навчання.

Зазначимо, що наведений в [4] варіант СДН ЗСУ недостатньо повно відображає специфіку та структурованість військової освіти. Тому структура СДН ЗСУ, що буде впроваджуватися, повинна бути більш набли-

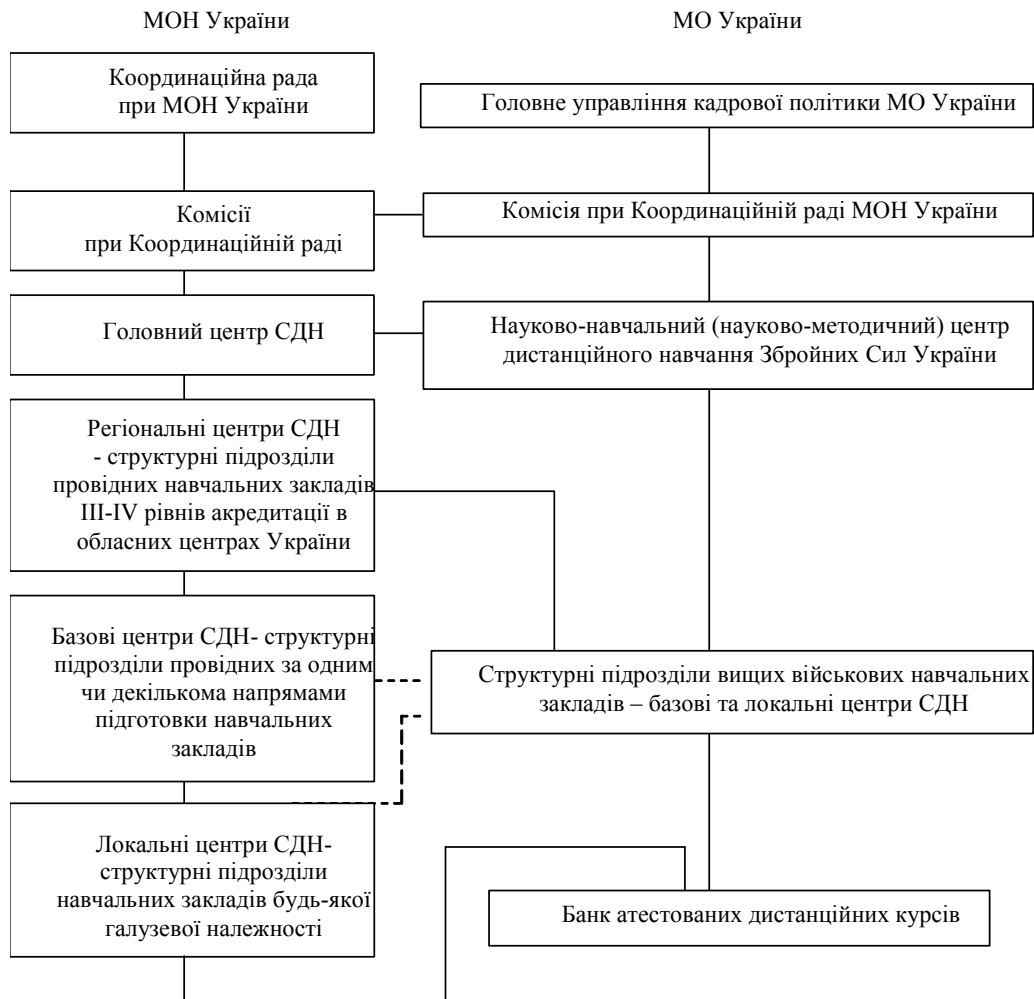


Рисунок 3 – Варіант загальної структури інтегрованої системи дистанційного навчання Міністерства освіти і науки та Міністерства оборони України

женою до структури, що затверджена як загальнодержавна [2].

Може бути запропонований наступний варіант загальної структури інтегрованої СДН Міністерства освіти і науки та Міністерства оборони України (рис. 3).

Виходячи з того, що Головне управління кадрової політики МО України є координуючим органом взаємодії Міністерства оборони України з Міністерством освіти і науки України доцільно та найбільш ефективно визначення його в якості виконавця головної ролі у військовій системі дистанційного навчання.

Враховуючи специфіку підготовки спеціалістів військовими навчальними закладами, при Координаційній раді МОН України доцільно створити окрему комісію з цього напрямку.

Слід зазначити, що вищі військові навчальні заклади МО України III–IV рівнів акредитації скоріш за все не зможуть на даний час виступати в якості регіональних центрів СДН. Це зумовлено, по-перше – ма-

лою кількістю таких закладів, по-друге – тим, що в тих регіональних центрах, де розташовані вищі військові навчальні заклади, існують ще декілька більш потужних цивільних вищих навчальних закладів з більш досвідченим науково-педагогічним складом. Тому можна припустити, що всі військові навчальні заклади (зокрема, факультети та кафедри військових навчальних закладів), навчальні частини МО України в загальнодержавній системі дистанційного навчання будуть займати рівні базових та локальних центрів СДН.

ВИСНОВКИ

В умовах подальшої інтеграції військової та цивільної освіти впровадження дистанційних форм навчання в освітній процес Збройних Сил України потребує створення обґрунтованої структури військової системи дистанційного навчання, яка повинна стати складовою загальнодержавної системи дистанційної освіти Ук-

раїни та відповідати вимогам нормативно-правових актів у даній сфері.

Наукова новизна роботи заключається в тому, що запропонований варіант загальної структури інтегрованої СДН Міністерства освіти і науки та Міністерства оборони України з одного боку відповідає затвердженій загальнодержавній структурі СДН, а з другого — враховує специфіку впровадження дистанційного навчання в освітній процес Збройних Сил України і може бути взятий за основу.

Перспективи подальших досліджень в даному напрямку полягають в деталізації структури, що запропонована, конкретизації функцій структурних елементів СДН, обґрунтуванні та розробці фізичної структури спеціалізованої інформаційно-телекомунікаційної системи, а також в вирішенні великої кількості вторинних задач стосовно організації процесу навчання: його інформаційного, програмного забезпечення тощо.

ПЕРЕЛІК ПОСИЛАНЬ

1. Концепція розвитку дистанційної освіти в Україні [Електронний ресурс]. – Електрон. текстові дан. – Режим доступу: <http://udec.ntu-kpi.kiev.ua/udec.nsf/Documentukr?OpenPage&AutoFramed>, свободний. – Загл. с екрана. – Яз. укр., англ., рус.
2. Положення про дистанційне навчання [Електронний ресурс]. – Електрон. текстові дан. – Режим доступу:

- <http://udec.ntu-kpi.kiev.ua/udec.nsf/polozzennyaukrDO?OpenPage&AutoFramed>, свободний. – Загл. с екрана.
3. Возможности дистанционного обучения в военном образовании [Электронный ресурс]: G & Sh Laboratories Inc. Публикации / Шварц Ю. С. // Вестник Военной Академии Республики Беларусь. – № 2 (3). – 10 июня 2004 г. – Электрон. текстовые дан. – Режим доступа: http://rezult.at.tut.by/il_edu_DL.html, свободный. – Загл. с экрана.
4. Концепція створення системи дистанційного навчання у Збройних Силах України. – К.: МО України, 2004. – 12 с.
5. Дистанционное обучение: Устойчивые структуры учебного материала [Электронный ресурс] / Лавров О. // Вопросы Интернет образования. – № 19. – Электрон. журн. – Режим доступа: http://www/vio.fio.ru/vio_19/cd_site/Articles/art_1_11.htm, свободный. – Загл. с экрана.

Надійшла 23.03.06
Після доробки 8.12.06

На основе анализа структуры существующей системы дистанционного обучения (СДО) Министерства образования и науки Украины и положений Концепции создания системы дистанционного обучения в Вооруженных Силах Украины предложен вариант общей структуры интегрированной СДО Министерства образования и науки и Министерства обороны Украины.

On the basis of analysis of the existent distance learning system (DLS) structure of the Ministry of Education and Science of Ukraine and positions of creation of the distance learning system Conception in Military Forces of Ukraine the variant of general structure of integrated DLS of the Ministry of Education and Science and Department of defense Ukraine is offered.

УДК 658.012.011.56

С. С. Ющенко

ІНФОРМАЦІЙНО-ПРАВОВІ СИСТЕМИ: РОСІЙСЬКИЙ ДОСВІД

Статтю, створену в рамках реалізації науково-технічної програми «Розробка методів та програмних засобів лінгвістичної підтримки та експертизи законодавчих процесів» (провідна організація – Український мовно-інформаційний фонд НАН України), присвячено аналізу ряду російських інформаційно-правових систем.

ВСТУП

Демократизаційні процеси, що відбуваються як на Україні, так і на всьому пострадянському просторі, розвиток прав людини зумовлює велику запитаність на актуальну правову інформацію серед широкого кола споживачів.

У наш час доступ до такої інформації не є проблемою – завдяки розвитку Інтернет-середовища, комп'ютерних технологій та технічної бази, органи влади та управління на власних сайтах мають змогу оперативно

інформувати про зміни, які відбуваються у законодавстві. Але на початку 90-х ситуація була зовсім іншою: з одного боку – безпрецедентна активність у галузі нормотворчості, з другого – відсутність електронних банків законодавчих документів та систем електронного документообігу в державних органах, не кажучи вже про зачатковий стан Інтернет-сервісу.

Попит, який на той час новоутворені держави були неспроможні задовольнити та нові ринкові відносини сприяли появі незалежних (недержавних) розповсюджувачів правової інформації, зорієнтувавши їх на створення специфічних програмних продуктів – інформаційно-правових систем (ІПС). У короткий термін виник конкурентний ринок із відповідними законами.

Об'єктом аналізу, який є змістом даної статті, не випадково обрані саме російські ІПС – адже їх розробники, діючи в умовах жорсткої конкуренції, яку

складають один одному та мають з боку органів державної влади (за даними довідкового web-сайту www.russianlaw.net, на російському інформаційно-правовому просторі діють приблизно 800 ІПС при існуванні близько 1500 повноцінних Інтернет-ресурсів органів державної влади та управління (www.medialaw.ru)), для привертання клієнтів використовують не лише маркетингові прийоми, а й постійно розширюють спектр послуг, розробляють та застосовують технологічні та організаційні підходи, будують нові наукові концепції. У цьому плані досвід російських розробників є показовим та край корисним для вітчизняних, які знаходяться в умовах менш сильного конкурентного тиску.

1 ХАРАКТЕРИСТИКА ІПС РОСІЙСЬКОЇ ФЕДЕРАЦІЇ

Системи, представлені на російському ринку інформаційно-правових послуг, розподіляються на такі групи:

- власне ІПС (бази даних (БД) містять довідково-правову інформацію: нормативно-правові акти, судові рішення, консультації спеціалістів, аналітичні матеріали тощо);

- вузькоспеціалізовані програмні продукти, спрямовані на автоматизацію роботи юристів, секретарів, суддів, архіваріусів, органів державної реєстрації та інших (БД містять збірники типових договорів з можливістю пошуку, автозаповнення та редагування, ведення обліку та інше);

- експертні юридичні системи;

- юридичні електронні словники.

Враховуючи те, що потужні ІПС у своєму складі містять вузькоспеціалізоване програмне забезпечення та словники, розподіл можна звести до двох груп, на яких ми зупинимось детальніше – умовно назвемо їх системи загального та спеціального призначення.

1.1 Системи загального призначення

До цієї групи відносимо професіональні правові системи, основним призначенням яких є розповсюдження правової інформації.

За рахунок гнучких тарифних планів (вартість користування системою залежить від обсягу обраного тематичного розділу), коло користувачів край широке – від потужних клієнтів, таких як органи різних гілок влади, корпорації, концерни, юридичні, фінансові аудиторські та консалтингові компанії, банки, великі підприємства, учбові та наукові заклади до представників малого та середнього бізнесу.

Як вже відмічалось, більшість розробників систем закладає в БД не тільки документи федерального та регіонального законодавства, а й різноманітні мате-

ріали, що стосуються міжнародного права, судової та правозастосовної практики, коментарі законодавства, типові зразки правових та звітних документів, консультаційну інформацію, електронні довідники нормативно-технічної та правової інформації, юридичні словники, стандарти, підручники та книги, навчальні посібники й багато іншого.

ІПС складаються з двох основних компонентів: розподіленої БД, яка постійно актуалізується та програмного комплексу. Оновлення системи замовника здійснюється через кур'єра або через Інтернет з потрібною періодичністю – від доби до тижня.

Сервісні послуги концентруються переважно навколо пошуку документів за запитом користувачів; передбачено можливість його здійснення за кількома критеріями: атрибутами, тематикою або контекстом. Результатом пошуку є всі документи, що відповідають зазначеним критеріям.

Розробниками ІПС, діючими незалежно один від одного, однаково ефективно було вирішено ряд взаємопов'язаних організаційних, програмних та технологічних задач. Наприклад, для формування інформаційного наповнення здійснюється співробітництво у загальноросійському масштабі з органами влади федерального та регіонального рівнів, місцевого самоврядування, іншими державними інститутами та суспільними організаціями, закладами науки, ЗМІ. Організація багатокритеріального пошуку потребувала систематизації, тематичної розмітки та індексування документів; оновлення системи через мережу Інтернет – організації розподілених систем, розробки та адаптації засобів флеш-взіємодії з великими масивами даних та інше.

На даний час у Росії найбільш потужними за складом та кількістю клієнтів, обсягом БД, спектром послуг, що надаються, кількістю сервісних центрів та дистриб'юторів є такі системи, як «КонсультантПлюс», «КОДЕКС: ПРАВО» та «Гарант».

Кожна з вказаних систем має версію у вільному доступі, розраховану на широкі верстви населення, що позитивно впливає на інформованість громадян стосовно змін, які відбуваються у законодавстві. Крім того, на сайтах розробників систем можна отримати відомості про поточні надходження документів, аналітичний огляд найбільш важливих документів законодавства, судової практики, консультації з правових та юридичних питань та іншу корисну інформацію.

Інформаційно-правова система КонсультантПлюс. Розробником цієї, без перебільшення, загальноросійської системи є компанія «КонсультантПлюс» (адреса серверу www.consultant.ru). 300 регіональних інформаційних центрів та 400 додаткових сервісних підрозділів компанії утворюють інформаційно-правову мережу. Обсяг БД повнотекстових документів сягає близько 2 млн.

До складу системи входять спеціалізовані системи: КонсультантБухгалтер, КонсультантФінансист, КореспонденціяСчетов.

Особливості: БД містить розділ міжнародного права з посиланням на сайти національних законодавств, довідково-правові системи інших країн, архівні НПА.

Інформаційно-правова система «КОДЕКС: ПРАВО». ІПС «КОДЕКС: ПРАВО» є розробкою ГП «Центр комп'ютерних розробок» (Санкт-Петербург). З метою її просування та утвердження на ринку, було створено консорціум «Кодекс», до складу якого увійшли більш ніж 200 компаній, що займаються розробкою та розповсюдженням програмних продуктів під торгівельною маркою «Кодекс» (www.wkkodex.ru).

Для зручності користувачів БД розподілена по блоках, кожний з яких містить певні види структурованої правової інформації.

Особливості: програмний комплекс «Кодекс-Мастер», на якому створені всі інформаційні розділи БД «Кодекс», надає користувачеві інструментарій для створення, розвитку та актуалізації власних розробок банків даних, довідників, електронних підручників тощо.

Інформаційно-правова система «Гарант». ІПС «Гарант» є продуктом ТОВ НВП «Гарант-Сервіс» (адреса серверу garant.park.ru).

БД містить більш ніж 900 тис. нормативних актів федерального та регіонального рівнів та допоміжну інформацію. Поповнення бази – щотижневе, приблизна кількість документів, що додаються, сягає 4 тис.

Інформаційний банк системи «Гарант», побудований по модульному принципу, складається з 40 блоків, вмістом яких є законодавство суб'єктів РФ та 25 спеціалізованих блоків з інформаційними рішеннями для ряду потенційних клієнтів – будівельних організацій, юристів, фармацевтичних та медичних організацій, керівників та менеджерів; бухгалтерів, аудиторів, фінансистів.

Особливості:

- наявність у базі матеріалів провідних економічних видань, книг по податкам та бухгалтеру, електронної версії Практичної Енциклопедії Бухгалтера, інформації, що стосується податкових судових справ;

- можливість багатоетапної юридичної обробки документів, що містяться у створенні коментарів та посилань на НПА, які цитуються.

1.2 Системи спеціального призначення

Значне підвищення за останні кілька років кількості звертань до експертів-юрислінгвістів [1] зумовило сплеск інтересу до використання інформаційних технологій при здійсненні різного роду експертиз НПА; результатом такої інтеграції стали системи, віднесені нами до цієї групи.

На відміну від систем, функція яких зводиться до підтримки повнотекстового пошуку, що ґрунтується на статистичних методах, у системах спеціального призначення використовуються технології інтелектуальної обробки текстів, створені у співробітництві зі спеціалістами у різних галузях знань. Дана група отримала назву «експертні» (що, на нашу думку, треба пов'язувати більш із буквальною семантикою – призначенням таких систем є інформаційний супровід проведення експертиз, аніж зі здатністю виконувати функції, визначені в класичній теорії експертних систем [2, 3]).

Програмно-методичний комплекс «Проект НПА». Комплекс «Проект НПА», як і технологія, на якій він базується – прикладна система АРМ «Юрист», є розробками юридично-інформаційного агентства INTRALEX.

Призначення комплексу – інформаційний супровід процесу створення НПА – від законодавчої ініціативи до проекту, який представляється у відповідний орган законодавчої влади, визначає коло користувачів – ними є суб'єкти права, які безпосередньо займаються розробкою НПА та органи державної влади і управління.

Джерелами інформації є БД нормативних актів ІПС «Консультант Плюс» та «Гарант».

Залучення комп'ютерних технологій до такої суто інтелектуальної діяльності, як правотворчість, дозволяє проводити роботу над текстом НПА на більш високому організаційному та логічному рівні. Координація зусиль різних розробників однієї теми суттєво спрощує контроль за ходом розробки з боку замовника та підрозділів, відповідних за проходження проекту.

В основу функціональної частини комплексу покладено дерево цілей, яке будується із врахуванням конкретних задач та особливостей нормотворчих процедур.

Комплекс передбачає зберігання усіх варіантів тексту та результатів його аналізу на будь-якій стадії підготовки певного акту, дозволяє фіксувати авторство творчих колективів із метою запобігання дублювання роботи.

Проект «Минерва». На даний час цей проект знаходиться у стадії розвитку; розробником є НВЦ «ІНТЕЛЕК ПЛЮС».

Мета проекту – створення системи, яка дозволить в автоматичному режимі проводити перевірку НПА на непротиречивість.

Вказану процедуру вбачається здійснювати шляхом порівняння семантичних описів документів, що надходять до БД системи, із семантичними описами, які вже зберігаються.

Організація такого процесу потребувала перегляду та вдосконалення технології автоматичного формування семантичного опису, яка бере початок ще з розробок систем спілкування природною мовою [2].

Похідний текстовий документ конвертується в еквівалентний формальномовний вираз (метамовою слугує об'єктна мова власної розробки) в результаті послідовного проходження через основні блоки системи, а саме – морфологічний, синтаксичний та семантичний аналізатори. Не вважаємо доцільним окремо зупинятися на функціях блоків – вони інтуїтивно зрозумілі з назв та детально описані у літературі [2, 3].

Протиріччям вважається відсутність у конкретного об'єкта певних, чітко визначених властивостей та (або) невідповідність моделі поведінки об'єкта набору раніше побудованих моделей.

Якщо у результаті співставлення описів має місце розбіжність, користувач отримує документи, в яких згадується об'єкт, для того, щоб мати можливість позиціонуватися на фрагментах, що, за припущенням, містять протиріччя.

Слабким місцем розробки, на нашу думку, є те, що кожна відсутня у базі, але наявна в об'єкта властивість або модель поведінки буде викликати зупин програми. Оскільки будь-який об'єкт має незлічену кількість характеристик, робота з системою може значним чином ускладнитися. Але зважаючи на те, що вдосконалення системи триває, будь-яке критичне судження щодо неї є передчасним, тим більш, що розробник намагається подолати вказану ваду шляхом закладення у систему потенційної спроможності розвитку при взаємодії з користувачем – передбачається, що словник предметної галузі (у даному випадку – юридичної), необхідний для зняття омонімії і база знань, в якій зберігаються шаблони (ними є графи речень, які за необхідністю перетворюються у формальномовний вираз і навпаки), будуть оперативно поповнюватися.

Інформаційно-правовий проект «Мониторинг правового пространства в правоприменительной практике в СФ РФ». Виконавцем проекту є НВЦ «ІНТЕЛТЕК ПЛЮС». Мета проекту – створення системи підтримки законопроектної експертизи, спрямованої на виявлення конфліктної взаємодії норм права.

Шукати протиріччя в текстах нормативно-правових актів передбачається за допомогою онтологічної моделі, побудованої, як і семантична мережа, з базових понять та відношень поміж ними. Розробник визначає «онтологію» як об'єднання усіх семантичних представлень текстів предметної галузі в єдину мережу.

Протиріччям вважається невідповідність між онтологічними моделями, що зберігаються у бібліотеці та моделлю – онтологічним представленням тексту, який надійшов на експертизу.

Ідея використовувати онтології в якості засобу лексико-семантичних представлень також не є новою. Вона виникла у середині 50-х років минулого сторіччя в рамках лінгвістичних теорій контекстно-незалежних граматик, що розглядали синтаксис як особливий аспект, не пов'язаний із семантикою. Але на той час була

практично нездійсненною, зважаючи на те, що швидкодія тогочасних ЕОМ складала десятки тисяч операцій на секунду, а ємність пам'яті мала розмірність «слово» [4].

Сучасний рівень розвитку обчислювальної техніки та інформаційних технологій уможливило реалізацію багатьох фундаментальних розробок.

Документальна інформаційна система організаційно-правового захисту інтелектуальної власності у вищій школі «ПраЗИС». Дана система є спільною розробкою кафедри «Електронні обчислювальні машини та системи» МДТУ ім. Баумана та НВЦ ІНТЕЛТЕК. Унікальність її полягає в тому, що вона поєднує в собі системи трьох класів: інформаційно-довідкового (призначення), гіпертекстового (обробка текстів) та експертного (спеціальні функції) [5].

Метою створення є інформаційно-правова підтримка представників означеної галузі, але завдяки нестандартному технологічному рішенню розробників, система має ряд функцій, що надають користувачам додаткові (у порівнянні з традиційними ІПС) можливості, серед яких: формування запитів для отримання необхідної консультативно-довідкової інформації у довільній формі; проведення аналізу взаємної та внутрішньої непротиворіччівості законодавчих та інструктивних документів; генерування форм вихідних документів та деякі інші.

Розробці системи передувала копійка робота експертів та інженерів знань, а також цілий ряд розробок – структур даних та моделі представлення знань, засобів логічного виводу, спеціальної об'єктної СУБД для експертної частини системи, діалогового інтерфейсу користувача, мови запитів експертної системи, підсистеми обробки запитів тощо.

Відмітимо, що інформаційно-правові проекти НВЦ «ІНТЕЛТЕК ПЛЮС» мають спільну рису – значну наукову цінність.

ВИСНОВКИ

Конкурентоспроможність ІПС визначається зовсім не кількістю доступних документів, частотою оновлення або платою за користування, як це може здатися на перший погляд. На наше переконання, основною перевагою є зручність для користувача.

Не потрібно забувати, що користувачі – суб'єкти права, які платять гроші за послуги ІПС, роблять це не заради підвищення своєї правової культури або задоволення цікавості – для цього існують численні державні та недержавні правові проекти у вільному доступі. Звертаючись до ІПС, користувачі шукають не тексти документів – вони шукають відповіді на питання, що стосуються попередження або розв'язання з найменшою для себе шкодою (в перш за все – матеріальною) певних конфліктних ситуацій. З цієї точки

зору, на даний час жодна з існуючих російських ІПС не має задовільного «навігаційного» інструменту, який дозволив би зорієнтуватися у непростій системі національного законодавства, до того ж ускладненої численними посиланнями.

Розв'язок проблеми вбачається у розробці комп'ютерних інструментів, орієнтованих на інтелектуальне опрацювання природної мови і в цьому напрямку російські колеги, безумовно, досягли певних успіхів, а досвід, якого вони надбали, заслуговує на вивчення.

ПЕРЕЛІК ПОСИЛАНЬ

1. Широков В. А., Сидоренко О. О. Наукові засади лінгвістичної експертизи політичних та правових документів: постановка проблеми // Мовознавство. – 2005. – № 2. – С. 18–31.
2. Попов Э. В. Общение с ЭВМ на естественном языке. – М.: Наука, 1982. – 360 с.
3. Попов Э. В. Экспертные системы: Решение неформализованных задач в диалоге с ЭВМ. – М.: Наука. Гл. ред. физ.-мат. лит., 1987. – 288 с.
4. Информатика в понятиях и терминах: Кн. для учащихся ст. классов сред. шк. / А. Г. Бордовский, В. А. Из-

возчиков и др.; Под ред. В. А. Извозчикова. – М.: Просвещение, 1991. – 208 с.: ил.

5. Смирнов Ю. М., Андреев А. М., Березкин Д. В., Друшляков Г. И. Документальная информационная система организационно-правовой защиты интеллектуальной собственности в высшей школе «ПраЗИС» // Вестник МГТУ. – 1998. – № 2.

Надійшла 31.07.06

Після доробки 16.09.06

Статья, созданная в рамках реализации научно-технической программы «Разработка методов и программных средств лингвистической поддержки и экспертизы законодательных процессов» (ведущая организация – Украинский языково-информационный фонд НАН Украины), посвящена анализу ряда российских информационно-правовых систем.

In this article, which is created in the framework of a scientific technical program, called «Development of methods and software for the linguistic support and expertise for lawmaking procedures» realized by the Ukrainian Lingual Informational Fund of the National Academy of Sciences of Ukraine, there is an analysis of a series of Russian Informational Legal Systems.

ТЕОРІЯ І МЕТОДИ АВТОМАТИЧНОГО УПРАВЛІННЯ

ТЕОРИЯ И МЕТОДЫ АВТОМАТИЧЕСКОГО УПРАВЛЕНИЯ

THEORY AND METHODS OF AUTOMATIC CONTROL

УДК 504.03:338.24

Н. М. Ветрова

ЭКОЛОГИЧЕСКОЕ УПРАВЛЕНИЕ: ВИДЫ И ОСОБЕННОСТИ СТРАТЕГИЙ ЭКОЛОГИЧЕСКОЙ НАПРАВЛЕННОСТИ

В данной статье рассматриваются проблемы учета экологического фактора при разработке стратегии управления деятельностью на микроуровне и предлагается классификация базовых стратегий экологической направленности и мероприятия по их реализации.

ПОСТАНОВКА ПРОБЛЕМЫ

На современном этапе перехода к рыночным отношениям основой успеха любого предприятия на рынке являются предпринимательский талант руководителя и гибкая система управления. При этом стратегическое управление является важнейшей составляющей антикризисного управления [1], поскольку позволяет вовремя выявить внешние опасности и возможности предприятия, внутренние сильные и слабые стороны его производственно-хозяйственной деятельности и на основе оценки ситуации и внутрифирменной культуры разработать систему стратегий. В зависимости от целей и приоритетов компании в систему стратегий могут войти следующие составляющие: маркетинговая стратегия (развитие рынка продукта, экспорт, проникновение на новые рынки); интеграционные стратегии (система отношений с поставщиками и потребителями);

стратегия выхода на внешний рынок; логистическая стратегия (рационализация производства, ассортимента продукции); стратегия эффективности; стратегия сокращения деятельности [2], хотя в теории разработаны подходы к классификации стратегий по различным принципиальным схемам [3, 4].

Основы теории стратегического управления разрабатывалась в трудах многих ведущих западных (Ф. Котлер [3], М. Портер [4], Г. Минцберг, Б. Альстрэнд, Дж. Лэмפל [5], А. Томпсон и Дж. Стрикланд [6], И. Ансофф [7] и др.), российских и отечественных (А. П. Градов [8], А. Д. Юданов [9], А. В. Козаченко, А. Э. Воронкова [10], В. С. Ефремов [11], Д. М. Загирняк [12]) и других ученых, хотя не всегда охватывались отдельные важные стороны функционирования субъектов, к числу которых относятся экологические аспекты стратегического управления. Именно поэтому методология стратегического управления на современном этапе требует дальнейшего развития и цель данной статьи – рассмотреть виды и особенности экологически направленных стратегий предприятия в условиях необходимости обеспечения экологической устойчивости системы жизнедеятельности общества.

© Ветрова Н. М., 2007

РЕЗУЛЬТАТЫ ИССЛЕДОВАНИЯ

Для Украины стратегическое управление деятельностью предприятия является сравнительно новым процессом, необходимость его использования была вызвана переходом от плановой экономики с централизованным управлением к рыночной, в которой предприятия существуют как самостоятельные хозяйственные единицы. При этом руководство должно постоянно совершенствовать существующие системы управления, поскольку нарастает динамичность и неопределенность среды функционирования предприятия, включая изменения в состоянии окружающей среды.

Формирование и реализация стратегии – сложный процесс, требующий учета множества факторов внешней и внутренней среды функционирования предприятия и в последнее время, особое значение приобретает экологический фактор, учет которого пока все же не стал первоочередной сферой при формировании стратегии фирмы. При этом стратегическое управление, учитывающее влияние экологической составляющей, способно сформулировать подходы к решению одновременно двух проблем осуществления деятельности: снижение отрицательного воздействия производственной деятельности на окружающую природную среду и обеспечение уровня конкурентоспособности предприятия через достижение требуемых экологических характеристик товаров (услуг, благ) и соблюдение экологических требований к потенциалу предприятия.

В последние годы экологический фактор, как составляющая внешней среды фирмы, начинает занимать важное место в функционировании украинских предприятий, поскольку экологическая ситуация в Украине ухудшается: резко активизировались оползневые процессы, число которых достигло 20 тыс. оползней (2003 г.), что на 45 % больше, чем в 80-е годы; 18 % земель поражено эрозией; на 17 % земель отмечаются природное и техногенное подтопление (практически полностью нарушена система природного дренажа равнинной части территории, что является значительной экологической проблемой); резко расширились процессы абразивного разрушения побережья Черного и Азовского морей [13]. Большой вред экологии наносят отходы производства и жизнедеятельности. Так, в настоящее время в Украине значительно загрязненные территории составляют 61 тыс. км², высоко загрязненные территории – почти 116 тыс. км², а загрязненные – 121 тыс. км², а в целом площади занятые отходами производства достигают 1 % территории Украины [14]. В результате свыше 35 % населения страны дышат воздухом, насыщенным токсичными веществами, в два и более раз превышающими предельно допустимую норму. Возрастает загрязнение водоемов, грунтовых вод. По оценкам специалистов, первым и важнейшим элементом, влияющим на загрязнение и деградацию

окружающей среды, выступают промышленные предприятия [15]. Согласно принятому на Всемирном экономическом форуме в Давосе в 2001 году рейтингу 122 государств по «индексу экологической устойчивости», который агрегирует 22 комплексных индикатора по 67 параметрам, дающим возможность оценивать прогресс стран на пути к экологически безопасному развитию, Украина оказалась на 110 месте [16]. В таких условиях актуальность разработки и реализации мероприятий по преодолению экологических проблем постоянно возрастает как на макро- и мезоуровнях управления, так и для отдельных предприятий.

В свою очередь, экологические проблемы вносят определенные ограничения в деятельность субъектов хозяйствования. Экологические требования потребителей, поставщиков, государства, формирующих внешнюю среду предприятия, заставляют предпринимательские структуры при решении экологических проблем корректировать многие приоритеты. Поэтому руководство украинских предприятий призвано решать задачу разработки стратегии перехода от техногенного, природоразрушающего управления к управлению, предупреждающему вредное воздействие на окружающую природную среду. Однако следует отметить, что пока данный подход все таки не нашел широкого распространения. Так, для оценки значимости управления по обеспечению экологической безопасности проводился опрос 280 руководителей и специалистов предприятий – членов Донецкой торгово-промышленной палаты. Почти 40 % опрошенных считают, что действующая система управления предприятием слабо взаимосвязана с проблемами экологического развития, 50 % отмечают, что управление предприятием с точки зрения учета экологических целей и критериев неэффективно [17]. Таким образом, можно признать наличие проблемы учета экологических факторов при разработке как стратегических, так и оперативных мероприятий на предприятиях, несмотря на то, что в Украине существует методологическая основа для широкого внедрения инструментов экологического управления – Законы Украины «Про охрану окружающей природной среды», «Про экологический аудит», стандарты серии ДСТУ-1S014001-97 и ДСТУ-1S014004-97, система подзаконных, ведомственных актов и инструкций.

Основными причинами неразвитости системы стратегических решений в рамках экологического управления на украинских предприятиях можно считать:

- переход предприятий к современной рыночной системе управления (включая стратегический уровень) происходит медленно;
- многие руководители не осознали преимущества, которые может принести система экологического управления для предприятий: повышение конкурентоспособности за счет соблюдения экологических норм, как при производстве, так и при разработке товаров (про-

дукции) может обеспечить стабильный спрос на продукцию;

- недостаточный уровень экологического сознания населения страны;

- государственное регулирование охватывает не все реальные экологические цели общества, что влияет на уровень контроля экологической работы на предприятиях;

- несовершенство системы стимулирования природоохранных мер приводит к ситуации, при которой национальным товаропроизводителям выгоднее осуществлять экологические платежи, чем планировать и осуществлять внедрение экологобезопасных систем и совершенствовать производство, разработку стратегий экологической направленности.

При рассмотрении возможных стратегий, которые могут быть целесообразны с учетом роли процессов экологизации, следует учитывать сущность процессов производства, потребления и их связи с состоянием окружающей природной среды. Так, потребление общества связано не только с внутренними мотивационными, но и производственными процессами, а также с формируемыми культурными ценностями, как в социальных установках, так и экономической сфере. Следовательно, экологизация хозяйственной деятельности связана с экологизацией сознания в общем понимании. С другой стороны, экологические проблемы создают основу для изменения поведенческих схем, что проявляется в принимаемых решениях при организации деятельности. В этом процессе возникает особое видение стратегических действий, которые формируются с учетом экологического мировоззрения и поэтому имеют определенные отличия от общепризнанных базовых стратегий [3, 4, 6, 7, 8].

С учетом сущности стратегического управления применительно к экологической сфере базисными экологическими стратегиями могут быть признаны следующие (рис. 1).

1. *Стратегия снижения производственного давления на природные комплексы*, в рамках которой прекращается или ограничивается производство отдельных продуктов для оптимизации материалоемкости продукции, для чего реализуются программы замены продуктов, требующих дефицитных природных ресурсов, на продукты менее материалоемкие, обеспечения вторичного использования отходов сырья, отработанных или потребленных продуктов. В этой стратегии реализуется принцип сокращения привлекаемых природных ресурсов из среды.

2. *Стратегия повышения экологической эффективности использования природных ресурсов*, предполагающая совершенствование всех технологических процессов с целью существенного увеличения коэффициента использования сырьевых ресурсов и коэффициента полезного действия.

3. *Стратегия перехода к чистому производству*, в рамках которой обеспечивается минимизация негативного воздействия отдельного предприятия на окружающую природную среду путем:

- радикального изменения технологии: создание так называемого «чистого производства» (низко- или безотходная технология), «чистой» номенклатуры выпускаемой продукции и применяемого сырья (продукция не несет экологической опасности при ее потреблении и ликвидации в конце срока эксплуатации);

- формирования между предприятиями производственных цепей, реализующих кругооборот потоков сырья и отходов с учетом экотенциала окружающей природной среды;



Рисунок 1 – Характеристика базовых стратегий предприятия экологической направленности

– сокращения отрицательного воздействия на состояние природной среды со стороны группы предприятий за счет кооперирования при совместном решении экологических проблем, в ресурсосбережении и минимизации отходов.

Выделенные виды стратегий могут являться основой экологического управления фирмы, однако, для полного учета экологического аспекта в стратегии фирмы необходим всесторонний стратегический анализ особенностей отрасли и предприятия. Следовательно, руководству отечественных предприятий необходимо включить в систему стратегического управления, следующие составляющие:

– всесторонний анализ и контроль экологических последствий природопользования при производстве продукции, выполнении принимаемых управленческих решений;

– разработку и внедрение инновационных экологически безопасных производственных процессов;

– анализ возможности совершенствования технологии производства в части снижения его аварийности, степени технологического риска, энерго- и материалоемкости, количества и токсичности выбросов;

– анализ возможности использования биопозитивных технологических решений, способствующих сохранению природы;

– по результатам экологического мониторинга своевременное выявление объектов, наносящих ущерб окружающей природной среде, и принятие соответствующих решений, направленных на предотвращение или снижение ущерба;

– обеспечение экологической совместимости всех производств поликомпании, анализ возможности вторичного применения отходов производства;

– применение основных инструментов экологической политики, используемых в мировой практике (система экологического управления, экологический аудит, экологический маркетинг).

В долгосрочной перспективе повсеместная реализация предприятиями экологических направлений в стратегическом управлении должна способствовать улучшению экологической обстановки в стране.

ВЫВОДЫ

Экологическая направленность стратегического управления деятельностью фирмы, которая строится на обосновании наиболее целесообразных базовых экостратегий и мероприятия по их реализации может способствовать повышению конкурентоспособности предприятий не только на внутреннем, но и на международном рынке. Кроме этого, что можно считать не менее важной задачей, учет экологического аспекта в управлении, в частности, промышленными предприяти-

ями позволит преодолеть негативные тенденции в экологической обстановке в стране.

ПЕРЕЧЕНЬ ССЫЛОК

1. Шубин А. Стратегическое управление как основополагающая часть системы менеджмента предприятия // Журнал Европейской экономики. – 2003. – Т. 2, № 4. – С. 305–311.
2. Фурса В. Стратегия развития предприятий строительной индустрии // Бизнес Информ. – 1998. – № 8. – С. 33–38.
3. Котлер Ф. Основы маркетинга: Пер. с англ. / Общ. ред. и вступ. ст. Е. М. Пеньковой. – М.: Прогресс, 1990. – 736 с.
4. Портер М. Стратегія конкуренції: Пер. з англ. А. Олійник, Р. Скільський. – К.: Основи, 1998. – 390 с.
5. Минсберг Г., Альстрэнд Б., Лэмпл Дж. Школы стратегий: Пер. с англ. – СПб.: Издательство «Питер», 2001. – 688 с.
6. Томпсон А. А., Стрикланд А. Дж. Стратегический менеджмент. Искусство разработки и реализации стратегии: Учебник для вузов / Пер. с англ. под ред. Л. Г. Зайцева, М. И. Соколовой. – М.: Банки и биржи, ЮНИТИ, 1998. – 576 с.
7. Ансофф И. Новая корпоративная стратегия: Пер. с англ. – СПб.: ПитерКом, 1999. – 416 с.
8. Экономическая стратегия фирмы: Учеб. пособие / Под ред. А. П. Градова. – 2-е изд., испр. и доп. – СПб.: Специальная Литература, 1999. – 589 с.
9. Юданов А. Ю. Конкуренция: теория и практика – М.: Изд-во «АКАЛИС», 1996. – 272 с.
10. Козаченко Г. В., Воронкова А. Е., Коренев Е. Н. Основы корпоративного управления: Навч. посібник – Луганськ: Видавництво СНУ, 2001. – 480 с.
11. Ефремов В. С. Стратегическое планирование в бизнес-системах. – М.: Издательство «Финпресс», 2001. – 240 с.
12. Загірняк Д. М. Організаційне забезпечення стратегічного вибору підприємства. – Луганськ: Вид-во Східноукраїнського національного університету імені В. Даля, 2004. – 192 с.
13. Національна доповідь про стан техногенної та природної безпеки в Україні у 2003 році / МНС України, Мін природи України, НАН України. – К.: ДП «Агентство інформації, міжнародного співробітництва та розвитку «Чорнобильінтерінформ». – 436 с.
14. Буравльов Є. П., Гетьман В. В. Загальнодержавний моніторинг техногенної безпеки // Екологія і ресурси 36. наук. праць Інституту проблем національної безпеки. – К.: ІПНБ, 2005. – № 11. – С. 48–58.
15. Шмандий В. М. Определение степени техногенной опасности машиностроительных предприятий // Прогрессивные технологии и системы машиностроения: Междунар. сб. научн. трудов. – Донецк: ДонГУ. – 1998. – Вып. 6, т. 3. – С. 287–290.
16. Веклич О. Современное состояние и эффективность экономического механизма экологического регулирования // Экономика Украины – 2003. – № 10. – С. 62–70.
17. Садеков А. А. Экологический вызов и проблемы устойчивого развития Прометей: региональный сборник научных трудов по экономике / Донецкий экономико-гуманитарный институт; Институт экономико-правовых исследований Национальной академии наук Украины. – Донецк: ООО «Юго-Восток, Лтд», 2004. Вып. 14. – С. 18–24.

Надійшла 29.09.06

У даній статті розглядаються проблеми обліку екологічного чинника при розробці стратегії управління діяльністю на мікрорівні, і пропонується класифікація базових екостратегій і перелік заходів, необхідних для її вирішення.

In the article the problems of account of ecological factor in strategy formation process are examined, and classification of basic ecological strategies is offered, and measures for their realization are proponed.

И. А. Орловский, А. А. Синявский

РАЗРАБОТКА МОДЕЛЕЙ НЕЛИНЕЙНЫХ ЭЛЕКТРОТЕХНИЧЕСКИХ ОБЪЕКТОВ В ВИДЕ СТЕПЕННЫХ РЕКУРРЕНТНЫХ НЕЙРОННЫХ СЕТЕЙ

Для создания моделей нелинейных объектов предложены степенные рекуррентные нейронные сети (СРНС). Разработана методика расчета этих сетей. Проведена проверка способности СРНС отображать модели нелинейных электротехнических объектов.

Большинство электротехнических объектов являются нелинейными с изменяющимися в процессе работы параметрами. Качественное управление нелинейными объектами даже при известных значениях их параметров является сложной и не до конца решенной задачей.

В настоящее время развиваются идеи настройки регуляторов в системах управления электротехническими объектами с использованием генетических алгоритмов и градиентных алгоритмов обучения (при реализации регуляторов на искусственных нейронных сетях (НС)). Поиск параметров регуляторов в таких системах требует многократного запуска объекта и, как правило, выполняется не на реальном объекте, а на его модели. Отсюда возникает задача получения достаточно точных моделей объекта в процессе его работы.

Модели объектов могут находиться с использованием классических методов, например: методом наименьших квадратов, методом наблюдателей состояния (идеи фильтрации Калмана), а также развиваемыми в последнее время методами с использованием искусственных НС. Вероятностные методы не обеспечивают точности и быстродействия при изменении параметров объекта. Для применения метода наименьших квадратов необходимо описание объекта идентификации алгебраическими уравнениями. Применение расширенного фильтра Калмана требует предварительной настройки и дает строгое решение только для линейных объектов. Нейросетевые модели обладают свойством аппроксимировать любую функцию, имеют алгоритмы обучения, но требуют времени для выполнения обучения. Снизить время обучения можно, выполнив их предварительный расчет.

В научной литературе вопросы получения моделей электротехнических объектов с помощью НС рассматриваются достаточно часто, например [1–6]. В работе [5] для получения модели асинхронного двигателя

и оценки электромагнитного момента, амплитуды и фазы магнитного потока используется НС прямого пространства с двадцатью нейронами в скрытом слое и нелинейными активационными функциями (гиперболический тангенс). Данная сеть обеспечивает удовлетворительную точность только для обучающего набора данных. При изменении входных наборов способность сети отображать объект отсутствует.

В работе [6] адаптивные регуляторы системы управления настраиваются на модели электропривода постоянного тока гребного винта подводного аппарата, выполненной в виде адаптивных дискретных фильтров (представленных в виде «матриц состояний»). Параметры матриц определяются решением матричных уравнений методом наименьших квадратов.

В статьях [2, 3] показана возможность получения с высокой точностью линейной модели объекта с помощью рекуррентных нейронных сетей (РНС) на примере тиристорного электропривода постоянного тока. В [2] при получении модели двигателя постоянного тока с помощью РНС Джордана предлагается выполнять линеаризацию объекта на малых интервалах времени. Отмечается, что в этом случае модель на РНС является нелинейной глобально, но может рассматриваться линейной локально. При локальной линеаризации принимается, что на коротком интервале времени (десятки – сотни тактов счета) внутренние параметры объекта являются неизменными. Обучение РНС (нахождение матриц, описывающих объект в пространстве состояний) выполняется на наборах данных измеренных на коротких интервалах.

Модель линейного объекта с помощью РНС можно получить аналитически, выполнив расчет сети [3], если известны параметры объекта, либо по результатам измерений всего вектора состояний в процессе работы объекта в течение необходимого числа тактов счета. Если параметры объекта неизвестны или нет необходимого количества измерений, то определение весовых коэффициентов РНС может выполняться с помощью алгоритмов обучения (градиентного [1] или генетического [4]).

При реализации модели нелинейного объекта в виде адаптивного дискретного фильтра [6], либо РНС с линейной функцией активации [1–4] точность отображения объекта не высока. Расчет моделей НС с нелинейными функциями активации в литературе отсутствует, а обучение таких сетей требует больших затрат времени и не всегда позволяет получить заданную точность модели [1]. Кроме того, как правило, отсутствует возможность получения параметров объекта (идентификации) по значениям весовых коэффициентов этих моделей.

Нелинейный объект описывается нелинейными разностными уравнениями. На коротком интервале времени нелинейности разностных уравнений с высокой точностью могут быть представлены в виде степенного ряда, тогда значения переменных состояния объекта определяются суммой взвешенных значений степеней входных сигналов и взвешенных значений степеней переменных состояния объекта на предыдущем такте счета. Такое представление модели объекта можно выполнить с помощью РНС, которую для краткости будем называть степенной РНС (СРНС). По весовым коэффициентам (коэффициентам при степенях переменных) такой СРНС возможна идентификация внутренних параметров объекта. Если нелинейность объекта представляет степенную зависимость, то СРНС описывает такой объект во всем его диапазоне работы.

Целью статьи является разработка методики расчета весовых коэффициентов СРНС, позволяющей с высокой точностью отображать модель нелинейного объекта, а также разработка программы расчета СРНС и проверка полученных аналитических результатов методом математического моделирования на примерах получения моделей нелинейных электротехнических объектов.

Статья построена следующим образом. В разделе 1 рассмотрена аппроксимация нелинейных матриц объекта степенными зависимостями любого порядка от вектора состояния и входных сигналов. Для объекта с такими матрицами получены выражения расчета весовых коэффициентов СРНС, представляющей модель этого объекта. В разделе 2 даны общие выражения для расчета весовых коэффициентов СРНС, моделирующих нелинейный объект, по результатам измерений объекта. В разделе 3 рассмотрены два примера расчета моделей, выполненных на СРНС, нелинейных электротехнических объектов; а именно, нелинейной электрической цепи и тиристорного привода постоянного тока с двигателем последовательного возбуждения.

1 ОПРЕДЕЛЕНИЕ СТРУКТУРЫ И ОБЩИХ УРАВНЕНИЙ РАСЧЕТА ВЕСОВЫХ КОЭФФИЦИЕНТОВ СРНС

Представим модель объекта в виде системы уравнений, записанной в матричном виде в форме пространства состояний [1]

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

где x – вектор состояния объекта, $x = [x_1, x_2, \dots, x_q]^T$; u – вектор входных сигналов, $u = [u_1, u_2, \dots, u_k]^T$; A – матрица коэффициентов размера $q \times q$; B – матрица коэффициентов размера $q \times k$. Начальными условиями является вектор начальных состояний объекта $x(0) = [x_1(0), x_2(0), \dots, x_q(0)]^T$.

Для нелинейного объекта матрицы A и B являются нелинейными. Нелинейность матриц A и B может быть аппроксимирована в виде степенных зависимостей от вектора состояния и вектора входных сигналов соответственно. Задачей раздела является получение аналитических выражений для расчета весовых коэффициентов СРНС, представляющей модель нелинейного объекта, из матриц объекта, аппроксимированных в виде степенных зависимостей. В начале рассмотрим линейный объект, когда матрицы A и B неизменны, то есть $\partial A/dx_i = 0$, $i = (1, \dots, q)$ и $\partial B/du_j = 0$, $j = (1, \dots, k)$.

Если в дискретной модели объекта представить производную вектора состояния в виде

$$\dot{x}_n = (x_n - x_{n-1})/T, \quad (2)$$

где T – такт счета, x_n и x_{n-1} – значения вектора состояния на n и $n-1$ такте счета, то систему уравнений (1) можно записать следующим образом [3]:

$$x_n = (AT + E)x_{n-1} + BTu_n, \quad (3)$$

где E – единичная матрица размера $q \times q$.

Обозначим коэффициенты уравнения следующим образом:

$$W^{oc} = AT + E, \quad W^{in} = BT, \quad (4)$$

тогда уравнение (3) примет вид

$$x_n = W^{oc}x_{n-1} + W^{in}u_n. \quad (5)$$

Уравнение (5) представляет собой описание РНС, где W^{oc} , W^{in} матрицы весовых коэффициентов сигналов обратной связи и входных сигналов соответственно, вычисляемые по уравнениям (4). Если матрицы A

и B не известны, то их расчет можно выполнить из необходимого числа уравнений (3) полученных в разные моменты времени, например смещенных во времени на один такт счета [3].

В случае нелинейного объекта его линеаризация, т. е. допущение, что на расчетном интервале времени параметры объекта практически неизменны, возможна при незначительном изменении этих параметров. Однако при изменении параметров на десятки и сотни процентов ошибка линейной модели становится значительной. Снизить эту ошибку можно, используя более сложную структуру РНС, в которой принимаются линейные изменения матрицы A от вектора состояния объекта и матрицы B – от входного вектора ($\partial A/dx_i = \text{const}$, $i = (1, \dots, q)$ и $\partial B/du_j = \text{const}$, $j = (1, \dots, k)$). Расчет параметров такой РНС выполняется следующим образом.

Введем обозначение диагональной матрицы состояния X размера $q \times q$, и диагональной матрицы U входных сигналов размера $k \times k$

$$X = \begin{bmatrix} x_1 & 0 & \dots & 0 \\ 0 & x_2 & \dots & 0 \\ \dots & \dots & \dots & \dots \\ 0 & 0 & \dots & x_q \end{bmatrix}, \quad U = \begin{bmatrix} u_1 & 0 & \dots & 0 \\ 0 & u_2 & \dots & 0 \\ \dots & \dots & \dots & \dots \\ 0 & 0 & \dots & u_k \end{bmatrix}. \quad (6)$$

При линейной зависимости от вектора состояния x и вектора входных сигналов u соответственно, матрицы A и B представляются следующими уравнениями:

$$A = A_1 X + A_0, \quad B = B_1 U + B_0. \quad (7)$$

где A_1 и A_0 матрицы коэффициентов размера $q \times q$, B_1 и B_0 матрицы коэффициентов размера $q \times k$. В этом случае уравнение (1) примет вид

$$\dot{x} = (A_1 X + A_0)x + (B_1 U + B_0)u. \quad (8)$$

Значения весовых коэффициентов РНС также изменяются линейно относительно поступающих на них сигналов и выражения (4) примут вид

$$W^{oc} = A_1 X T + A_0 T + E, \quad W^{in} = B_1 U T + B_0 T. \quad (9)$$

Уравнения (9) удобно представить в виде

$$W^{oc} = W_1^{oc} X + W_0^{oc}, \quad W^{in} = W_1^{in} U + W_0^{in}, \quad (10)$$

тогда весовые коэффициенты РНС определяются выражениями

$$\begin{aligned} W_1^{oc} &= A_1 T, \quad W_0^{oc} = A_0 T + E, \\ W_1^{in} &= B_1 T, \quad W_0^{in} = B_0 T. \end{aligned} \quad (11)$$

Разностное дифференциальное уравнение объекта (5) при подстановке выражения (10) принимает следующий вид:

$$\begin{aligned} x_n &= W_1^{oc} X_{n-1} x_{n-1} + \\ &+ W_0^{oc} x_{n-1} + W_1^{in} U_n u_n + W_0^{in} u_n. \end{aligned} \quad (12)$$

В частном случае для скалярного объекта ($q = 1$) и скалярного входного сигнала ($k = 1$) матрицы A и B вырождаются до вида

$$A = a_1 x + a_0, \quad B = b_1 u + b_0. \quad (13)$$

Тогда весовые коэффициенты единственного нейрона определяются выражениями

$$\begin{aligned} W_1^{oc} &= a_1 T, \quad W_0^{oc} = a_0 T + 1, \\ W_1^{in} &= b_1 T, \quad W_0^{in} = b_0 T \end{aligned} \quad (14)$$

и разностное уравнение принимает вид

$$x_n = a_1 T x_{n-1}^2 + (a_0 T + 1)x_{n-1} + b_1 T u_n^2 + b_0 T u_n. \quad (15)$$

Вернемся к рассмотрению многомерной системы. Более точная аппроксимация нелинейной функции (а следовательно, и модели нелинейного объекта) может быть выполнена с использованием полинома второго порядка, при этом $\partial^2 A/dx_i^2 = \text{const}$, $i = (1, \dots, q)$ и $\partial^2 B/du_j^2 = \text{const}$, $j = (1, \dots, k)$. В этом случае матрицы объекта описываются выражениями

$$A = A_2 X^2 + A_1 X + A_0, \quad B = B_2 U^2 + B_1 U + B_0. \quad (16)$$

Уравнение (3) примет следующий вид:

$$\begin{aligned} x_n &= (A_2 X_{n-1}^2 + A_1 X_{n-1} + A_0)x_{n-1} + \\ &+ (B_2 U_n^2 + B_1 U_n + B_0)u_n. \end{aligned} \quad (17)$$

Значения весовых коэффициентов РНС, согласно (4), определяются уравнениями

$$\begin{aligned} W^{oc} &= A_2 T X_{n-1}^2 + A_1 T X_{n-1} + A_0 T + E, \\ W^{in} &= B_2 T U_n^2 + B_1 T U_n + B_0 T. \end{aligned} \quad (18)$$

Эту зависимость, аналогично (10), можно представить в виде

$$\begin{aligned} W^{oc} &= W_2^{oc} X_{n-1}^2 + W_1^{oc} X_{n-1} + W_0^{oc}, \\ W^{in} &= W_2^{in} U_n^2 + W_1^{in} U_n + W_0^{in}, \end{aligned} \quad (19)$$

тогда весовые коэффициенты РНС определяются выражениями

$$W_2^{oc} = A_2T, W_1^{oc} = A_1T, W_0^{oc} = A_0T + E, \\ W_2^{in} = B_2T, W_1^{in} = B_1T, W_0^{in} = B_0T. \quad (20)$$

Аналогично, при использовании полинома m -го порядка для матрицы A и полинома r -го порядка для матрицы B объекта ($\partial^m A / dx_i^m = \text{const}$, $i = (1, \dots, q)$) и $\partial^m B / du_j^m = \text{const}$, $j = (1, \dots, k)$), выражения матриц A и B имеют вид

$$A + A_m X^m + A_{m-1} X^{m-1} + \dots + A_1 X + A_0, \\ B = B_r U^r + B_{r-1} U^{r-1} + \dots + B_1 U + B_0. \quad (21)$$

Тогда разностное уравнение процесса примет вид

$$x_n = (A_m X_{n-1}^m + A_{m-1} X_{n-1}^{m-1} + \dots + A_1 X_{n-1} + A_0) x_{n-1} + \\ + (B_r U_n^r + B_{r-1} U_n^{r-1} + \dots + B_1 U_n + B_0) u_n. \quad (22)$$

Значения весовых коэффициентов РНС, аналогично (4), определяются уравнениями

$$\left. \begin{aligned} W^{oc} &= A_m T X_{n-1}^m + A_{m-1} T X_{n-1}^{m-1} + \dots + A_0 T + E \\ W^{in} &= B_r T U_n^r + B_{r-1} T U_n^{r-1} + \dots + B_1 T U_n + B_0 T \end{aligned} \right\}. \quad (23)$$

Если представить уравнение (23) в виде

$$\left. \begin{aligned} W^{oc} &= W_m^{oc} X_{n-1}^m + W_{m-1}^{oc} X_{n-1}^{m-1} + \dots + \\ &\quad + W_1^{oc} X_{n-1} + W_0^{oc}, \\ W^{in} &= W_r^{in} U_n^r + W_{r-1}^{in} U_n^{r-1} + \dots + W_1^{in} U_n + W_0^{in}, \end{aligned} \right\} \quad (24)$$

тогда весовые коэффициенты РНС определяются выражениями

$$\left. \begin{aligned} W_m^{oc} &= A_m T, W_{m-1}^{oc} = A_{m-1} T, \dots, W_1^{oc} = A_1 T, \\ W_0^{oc} &= A_0 T + E; \\ W_r^{in} &= B_r T, W_{r-1}^{in} = B_{r-1} T, \dots, W_1^{in} = B_1 T, \\ W_0^{in} &= B_0 T. \end{aligned} \right\} \quad (25)$$

В общем виде СРНС, построенная с использованием рассмотренных выше полиномов, представлена на рис. 1. Блоки Б1 и Б2 (рис. 1) формируют необходимое количество выходов, значения которых равны степеням входных сигналов, блоки z^{-1} выполняют задержку сигнала на один такт счета T .

2 ОПРЕДЕЛЕНИЕ ВЕСОВЫХ КОЭФФИЦИЕНТОВ СРНС ПО РЕЗУЛЬТАТАМ ИЗМЕРЕНИЙ

В предложенной методике для определения коэффициентов СРНС должны быть известны векторы состояний объекта и входных сигналов. Запишем уравнение (22) в виде системы разностных дифференциальных уравнений с использованием зависимостей (23) и (24) следующим образом:

$$x_n = \\ = (W_m^{oc} X_{n-1}^m + W_{m-1}^{oc} X_{n-1}^{m-1} + \dots + W_1^{oc} X_{n-1} + W_0^{oc}) x_{n-1} + \\ + (W_r^{in} U_n^r + W_{r-1}^{in} U_n^{r-1} + \dots + W_1^{in} U_n + W_0^{in}) u_n. \quad (26)$$

Неизвестными в данной системе являются матрица $W^{oc} = [W_m^{oc}, W_{m-1}^{oc}, \dots, W_1^{oc}, W_0^{oc}]$ размера $1 \times (m+1)$, каждый элемент которой является матрицей размера $q \times q$ и матрица $W^{in} = [W_r^{in}, W_{r-1}^{in}, \dots, W_1^{in}, W_0^{in}]$ размера $1 \times (r+1)$, каждый элемент которой является матрицей размера $q \times k$. Общее количество неизвестных составляет

$$N = (m+1)q^2 + (r+1)qk. \quad (27)$$

Для вычисления коэффициентов СРНС число измерений объекта (число уравнений) должно быть не меньше числа неизвестных N . Следовательно, минимальное количество тактов счета N_T при вычислении

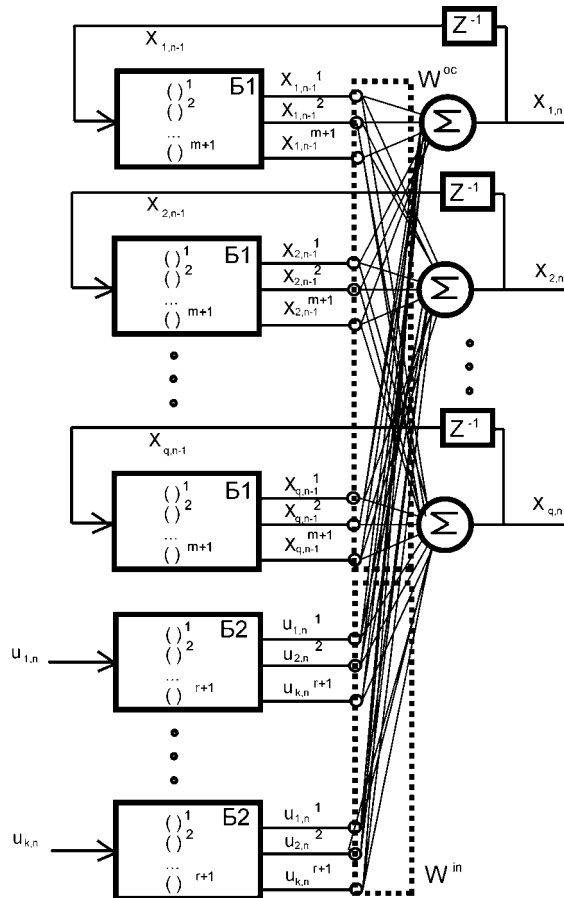


Рисунок 1 – Общая структурная схема СРНС

производной по формуле (2) должно быть на единицу больше числа неизвестных, т. е.

$$N_T = (m + 1)q^2 + (r + 1)qk + 1. \quad (28)$$

Пусть имеется $M + 1$ ($M > N$) последовательных измерений входного вектора и вектора состояния объекта. Это позволяет составить следующую систему уравнений:

$$\left\{ \begin{aligned} x_n &= W_m^{oc} X_{n-1}^m x_{n-1} + W_{m-1}^{oc} X_{n-1}^{m-1} x_{n-1} + \dots + \\ &+ W_1^{oc} X_{n-1} x_{n-1} + W_0^{oc} x_{n-1} + W_r^{in} U_n^r u_n + \\ &+ W_{r-1}^{in} U_n^{r-1} u_n + \dots + W_1^{in} U_n u_n + W_0^{in} u_n; \\ x_{n-1} &= W_m^{oc} X_{n-2}^m x_{n-2} + W_{m-1}^{oc} X_{n-2}^{m-1} x_{n-2} + \dots + \\ &+ W_1^{oc} X_{n-2} x_{n-2} + W_0^{oc} x_{n-2} + W_r^{in} U_{n-1}^r u_{n-1} + \\ &+ W_{r-1}^{in} U_{n-1}^{r-1} u_{n-1} + \dots + W_1^{in} U_{n-1} u_{n-1} + W_0^{in} u_{n-1}; \\ &\dots \\ x_{n-M+1} &= W_m^{oc} X_{n-M}^m x_{n-M} + W_{m-1}^{oc} X_{n-M}^{m-1} x_{n-M} + \\ &+ \dots + W_1^{oc} X_{n-M} x_{n-M} + W_0^{oc} x_{n-M} + \\ &+ W_r^{in} U_{n-M+1}^r u_{n-M+1} + W_{r-1}^{in} U_{n-M+1}^{r-1} u_{n-M+1} + \\ &+ \dots + W_1^{in} U_{n-M+1} u_{n-M+1} + W_0^{in} u_{n-M+1}. \end{aligned} \right. \quad (29)$$

Обозначим значения вектора в левой части системы (29) в течение M тактов в виде матрицы X_M , размерностью $(M \times 1)$, каждый элемент которой имеет размерность $(q \times 1)$ тогда матрица X_M имеет вид

$$X_M = \begin{bmatrix} x_n \\ x_{n-1} \\ \dots \\ x_{n-M+1} \end{bmatrix}. \quad (30)$$

Обозначим через $X_{M,u}$, согласно уравнению (29), матрицу степеней сигналов состояния объекта и степеней входных сигналов системы размера $((m + k + 2) \times M)$, элементы которой, определяемые через u , имеют

размерность $(q \times 1)$, элементы, определяемые через u , имеют размерность $(k \times 1)$

тогда систему (29) можно представить в виде

$$X_M = (W \cdot X_{M,u})^T, \quad (32)$$

где матрица весовых коэффициентов W размера $(1 \times (m + k + 2))$ имеет вид

$$W = [W_m^{oc}, W_{m-1}^{oc}, \dots, W_1^{oc}, W_0^{oc}, W_r^{in}, W_{r-1}^{in}, \dots, W_1^{in}, W_0^{in}]. \quad (33)$$

Элементы матрицы W , имеющие индекс «oc» представляют собой матрицы размера $q \times q$, элементы с индексом «in» представляют собой матрицы размера $q \times k$. Матрица W на расчетном интервале времени принимается неизменной.

Если число уравнений равно числу искомым коэффициентов ($M = N$), то матрица весовых коэффициентов СРНС определяется через обратную матрицу $X_{M,u}^{-1}$ уравнением

$$W = (X_M)^T \cdot X_{M,u}^{-1}. \quad (34)$$

В реальных условиях возможны ситуации, когда изменения вектора состояния объекта за время T не превышает точности измерения датчиков. В этом случае отсутствует обратная матрица $X_{M,u}^{-1}$. Из-за неточности измерений датчиков (если даже обратная матрица существует) расчет последовательно на каждом такте счета матрицы весовых коэффициентов W приводит к существенному изменению значений ее составляющих, которые вычислены на последовательных тактах счета. Для устранения этих недостатков число измерений необходимо взять больше минимального (например, в нес-колько раз), тогда определение матрицы W выполняется нахождением минимального среднеквадратичного отклонения для всех уравнений с помощью псевдообратной матрицы $X_{M,u}^+$ [7] по выражению

$$W = (X_M)^T \cdot X_{M,u}^+. \quad (35)$$

Обеспечение точной работы модели в первые моменты времени моделирования достигается заданием вектора

$$X_{M,u} = \begin{bmatrix} X_{n-1}^m x_{n-1} & X_{n-1}^{m-1} x_{n-1} & \dots & X_{n-1} x_{n-1} & x_{n-1} & U_n^r u_n & U_n^{r-1} u_n & \dots & U_n u_n & u_n \\ X_{n-2}^m x_{n-2} & X_{n-2}^{m-1} x_{n-2} & \dots & X_{n-2} x_{n-2} & x_{n-2} & U_{n-1}^r u_{n-1} & U_{n-1}^{r-1} u_{n-1} & \dots & U_{n-1} u_{n-1} & u_{n-1} \\ \dots & \dots & \dots & \dots & \dots & \dots & \dots & \dots & \dots & \dots \\ X_{n-M}^m x_{n-M} & X_{n-M}^{m-1} x_{n-M} & \dots & X_{n-M} x_{n-M} & x_{n-M} & U_{n-M+1}^r u_{n-M+1} & U_{n-M+1}^{r-1} u_{n-M+1} & \dots & U_{n-M+1} u_{n-M+1} & u_{n-M+1} \end{bmatrix}^T, \quad (31)$$

начальных состояний СРНС $x_0 = [x_{01}, x_{02}, \dots, x_{0q}]^T$ размера $(q \times 1)$, значения которого устанавливаются во временных задержках обратных связей СРНС. Расчет этого вектора выполняется по следующей зависимости:

$$x_0 = x(0) - W^{in} \cdot u_0, \quad (36)$$

где $u_0 = [U^r(0)u(0), U^{r-1}(0)u(0), \dots, U(0)u(0), u(0)]^T$ – матрица входных сигналов в первом такте моделирования, составленная из степенных значений элементов вектора входных сигналов в этом же такте размера $((r+1) \times 1)$, каждый элемент которой является матрицей размера $(k \times 1)$. В общем случае векторы x_0 и $x(0)$ не равны друг другу, так как выходной вектор СРНС в начальный момент времени формируется в виде суммы взвешенных сигналов с временных задержек и входных.

3 ПРИМЕРЫ МАТЕМАТИЧЕСКОГО МОДЕЛИРОВАНИЯ РАСЧЕТА МОДЕЛЕЙ НЕЛИНЕЙНЫХ ЭЛЕКТРОТЕХНИЧЕСКИХ ОБЪЕКТОВ В ВИДЕ СРНС

Реализация СРНС в стандартных средствах математического моделирования не предусмотрена, в связи с чем была разработана программа, позволяющая выполнять расчет весовых коэффициентов СРНС с любого места измеряемой последовательности сигналов на заданном числе тактов счета с любым количеством степеней входных сигналов и сигналов обратных связей. Программа также позволяет выполнять моделирование динамики рассчитанных СРНС для примеров 1 и 2.

Пример 1. Пусть объект имеет источник питания с ЭДС E , усилитель с нелинейным коэффициентом усиления $K = K(E)$, нагрузку, состоящую из последовательного соединения индуктивности L и нелинейного активного сопротивления $R = R(i)$. Ток в такой схеме определяется из уравнения:

$$\dot{i} = KL^{-1}E - RL^{-1}i. \quad (37)$$

В этом случае, согласно выражению (1), матрицы коэффициентов A и B определяются уравнениями

$$A = -RL^{-1}, \quad B = KL^{-1}. \quad (38)$$

СРНС, представляющая модель такого объекта, имеет один нейрон ($q = 1$) и один входной сигнал ($k = 1$). Для проверки приведенных выше выражений расчета весовых коэффициентов СРНС рассмотрим случаи, когда нелинейность коэффициента усиления K представляет различные степенные зависимости от э.д.с. источника питания ($K = k_0$, $K = k_1E + k_0$, $K = k_2E^2 + k_1E + k_0$) и нелинейность сопротивления R представляет степенные зависимости от протекающего через него тока ($R = R_0$, $R = R_1i + R_0$, $R = R_2i^2 + R_1i + R_0$), где k_0, k_1, k_2, R_1, R_2 – постоянные коэффициенты, R_0 – значение активного сопротивления при токе, равном нулю.

Для расчета моделей рассматриваемого объекта в виде СРНС при различных видах нелинейностей необходимо составить расчетные матрицы X_M и $X_{M,u}$ (уравнения (30)–(31)). Количество строк в этих матрицах определяется числом тактов счета, на которых выполняются измерения. Минимальное количество тактов счета, определяемое по формуле (28), для семи вариантов объекта представлено в таблице 1.

Моделирование выполнялось при следующих значениях параметров объекта: $E = 2 + 2\sin 5t$; $R_0 = 1$ Ом; $R_1 = 0,5$ Ом/А; $R_2 = 0,02$ Ом/А²; $k_0 = 1$; $k_1 = 0,3$; $k_2 = 0,05$; такт счета $T = 0,005$ с. Расчет коэффициентов СРНС (матрица W , уравнение (33)) выполнялся по уравнению (35) на различном количестве тактов счета с использованием псевдообратной матрицы $X_{M,u}^+$. Структура СРНС (порядок степенных зависимостей) выбиралась в точном соответствии с порядком степенных

Таблица 1

№ варианта	Уравнения коэффициента усиления K	Уравнения активного сопротивления R	Минимальное число тактов счета
1	$K = k_0$	$R = R_0$	3
2	$K = k_0$	$R = R_1i + R_0$	4
3	$K = k_1E + k_0$	$R = R_1i + R_0$	5
4	$K = k_0$	$R = R_2i^2 + R_1i + R_0$	5
5	$K = k_1E + k_0$	$R = R_2i^2 + R_1i + R_0$	6
6	$K = k_2E^2 + k_1E + k_0$	$R = R_2i^2 + R_1i + R_0$	7
7	$K = k_2E^2 + k_1E + k_0$	$R = R_1i + R_0$	6

зависимостей объекта. Результаты расчета весовых коэффициентов на 400 тактах счета для рассматриваемых семи вариантов и максимальная относительная ошибка моделей приведены в таблице 2. Незаполненные ячейки таблицы свидетельствуют об отсутствии связи в СРНС. Максимальная относительная ошибка определялась как максимальное значение отношения мгновенной абсолютной ошибки тока Δi (разность мгновенных значений токов в объекте и в модели на СРНС) к мгновенному значению тока в объекте на всем интервале времени моделирования.

Переходные процессы токов в объекте и ошибки моделей на СРНС для вариантов 1, 3 и 6 приведены на рис. 2. На рис. 2, а, рис. 2, в, рис. 2, д показаны изме-

нения токов i в объекте для вариантов 1, 3, 6 соответственно. На рис. 2, б, рис. 2, г, рис. 2, е приведены изменения ошибок тока Δi моделей на СРНС для вариантов 1, 3, 6 соответственно.

Результаты показали высокую точность рассчитанных моделей СРНС. Максимальная относительная ошибка для всех вариантов, согласно таблице 2, не превышала 0,66 %, что подтверждает правильность приведенных выше аналитических выражений для расчета коэффициентов СРНС.

Пример 2. Разработаем модели в виде СРНС и оценим их точность для тиристорного привода постоянного тока последовательного возбуждения при непрерывном токе якоря с различными нелинейностями, отлич-

Таблица 2

№ вар.	Значения весовых коэффициентов СРНС						Максим. относ. ошибка моделей на СРНС (%)
	W_2^{oc}	W_1^{oc}	W_0^{oc}	W_2^{in}	W_1^{in}	W_0^{in}	
1			0,995			0,00500	0,66
2		-0,00135	0,995			0,00502	0,46
3		-0,00140	0,995		0,00259	0,00473	0,26
4	-0,000559	-0,000553	0,994			0,00503	0,42
5	-0,000289	-0,00120	0,995		0,00260	0,00472	0,18
6	-0,000271	-0,00128	0,995	0,000117	0,00252	0,00479	0,16
7		-0,00140	0,995	0,000108	0,00256	0,00475	0,25

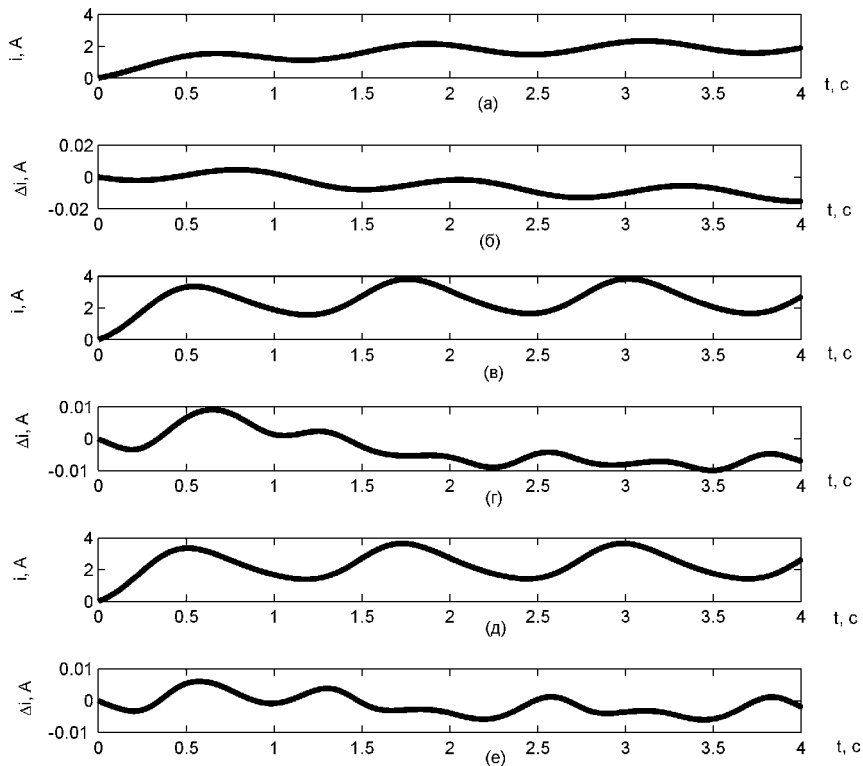


Рисунок 2 – Переходные процессы в нелинейном объекте и ошибки в моделях, выполненных в виде СРНС для примера 1

ными от степенных. Двигатель постоянного тока последовательного возбуждения описывается следующей системой уравнений [8]:

$$\begin{cases} E_d = L \frac{di}{dt} + E + Ri, & E = c\Phi(i)\omega, \\ J \frac{d\omega}{dt} = M - M_c, & M = c\Phi(i)i, \end{cases} \quad (39)$$

где E_d – ЭДС тиристорного преобразователя, i – ток двигателя, ω – угловая частота вращения вала (скорость) двигателя, M – электромагнитный момент двигателя, M_c – статический момент сопротивления, L , R – эквивалентные индуктивность и активное сопротивление электропривода, c – конструктивный параметр двигателя, $\Phi(i)$ – текущее значение магнитного потока, зависящее от тока двигателя, J – момент инерции двигателя, E – ЭДС двигателя.

Если в системе импульсно-фазового управления тиристорного преобразователя используется опорное линейное напряжение, то напряжение на выходе преобразователя для непрерывного тока определяется из уравнения [8]:

$$E_d = E_{d0} \cos(U_y / U_{y\max}), \quad (40)$$

где E_{d0} – максимальное значение ЭДС на выходе преобразователя, определяемое его схемой и входным напряжением, U_y – управляющее напряжение на входе преобразователя, $U_{y\max}$ – максимальное значение линейного (пилообразного) опорного напряжения. Коэффициент передачи тиристорного преобразователя является нелинейной зависимостью от U_y и определяется выражением

$$k_n = \frac{E_d}{U_y} = \frac{E_{d0} \cos(U_y / U_{y\max})}{U_y}. \quad (41)$$

Активное сопротивление якорной цепи изменяется в зависимости от температуры. Упрощенно принята зависимость этого сопротивления от модуля тока двигателя $R = R(i)$.

Зависимость магнитного потока от тока $\Phi(i)$ задается кривой намагничивания. Для рассматриваемого двигателя кривая намагничивания, в относительных единицах, задана таблицей 3 [9].

Таблица 3

i , (о.е.)	0,4	0,6	0,8	1	1,2
Φ , (о.е.)	0,65	0,8	0,91	1	1,08

Рассмотрим рабочий механизм, для которого момент инерции зависит от частоты вращения $J = J(\omega)$ и описывается следующим выражением:

$$J = \begin{cases} J_2, & \text{если } \omega > \omega_2, \\ k_J \omega, & \text{если } \omega_1 < \omega < \omega_2, \\ J_1, & \text{если } \omega < \omega_1. \end{cases} \quad (42)$$

На первом этапе для упрощения модели привода примем постоянную времени преобразователя и момент нагрузки привода M_c равными нулю. Такое упрощение позволяет исследовать возможности получения модели нелинейного объекта в виде СРНС. Значения параметров в выражениях (39)–(42) принимались следующими: $L = 0,334$ Гн, $R_{\text{ном}} = 0,476$ Ом, $c\Phi_{\text{ном}} = 0,67$ В·с, $J_{\text{ном}} = 0,144$ кг·м², $k_J = 0,0125$ с (кг·м²)⁻¹, $E_{d0} = 300$ В, $U_{y\max} = 10$ В.

СРНС, формирующие модели привода, имели три рекуррентных нейрона ($q = 3$) и один входной сигнал ($k = 1$). Модель привода выполнялась на СРНС третьей ($m = 2, r = 2$), шестой ($m = 5, r = 5$), и десятой ($m = 9, r = 9$), степеней. Число искоемых весовых коэффициентов для таких СРНС, согласно формуле (27), составляет 36, 72 и 120 соответственно. Для обеспечения устойчивости степенных РНС по экспериментальным данным установлена необходимость обеспечения значений входных и выходных сигналов в интервале от -1 до 1 . Для этого определяются максимальные значения входных и выходных сигналов объекта на расчетном интервале времени, которые в дальнейшем используются как нормировочные коэффициенты. Сеть рассчитывается и работает на нормированных данных. Модель объекта представляется рассчитанной СРНС с входными и выходными нормировочными коэффициентами. Результаты расчета весовых коэффициентов СРНС третьей степени приведены в таблице 4.

Таблица 4

Значения весовых коэффициентов СРНС			
степень	W_1^{in}	W_2^{in}	W_3^{in}
1	0.23273	0.002091	0.002982
2	-0.50198	-0.00323	-0.00695
3	0.27659	0.001152	0.004114
	W_1^{oc}	W_2^{oc}	W_3^{oc}
1	0.99105	0.008089	-0.00013
2	-0.00231	6.82E-05	0.000198
3	-0.00157	-1.32E-05	-1.44E-05
1	-0.04763	0.99844	-0.00029
2	0.010515	0.000792	0.001772
3	-0.00866	4.51E-05	-0.00029
1	0.039461	-0.00012	0.99917
2	-0.12566	-0.00072	0.000573
3	0.069677	0.000464	-0.00014

Результаты моделирования СРНС и абсолютные ошибки полученных моделей приведены на рис. 3–6.

На рисунке 3 отображены графики скорости, тока двигателя и ЭДС преобразователя на интервале 0–2 с. На рис. 4–6 отображены ошибки СРНС по отношению к объекту в интервале расчетных данных (0–2 с) со степенями переменных СРНС 3, 6 и 10 соответственно.

Относительные ошибки моделей для СРНС различных степеней приведены в таблице 5. Управляющим воздействием на преобразователь U_y являлось пилообразное напряжение с амплитудой 2π В и периодом

0,2 секунды. Такое задающее напряжение объединяет плавно нарастающее изменение сигнала и скачкообразное. Расчет СРНС выполнялся на всем наборе данных.

СРНС, согласно таблице 5, позволяют получить достаточно точную модель тиристорного привода постоянного тока последовательного возбуждения при непрерывном токе якоря с различными нелинейностями, которая может быть использована для поиска оптимальных параметров регуляторов в системах управления этими объектами.

Таблица 5

Степень СРНС	Максим. относит. ошибка E_d , %	Максим. относит. ошибка i , %	Максим. относит. ошибка ω , %
3	24	13	5,3
6	5,6	2,2	0,90
10	0,00022	0,16	0,54

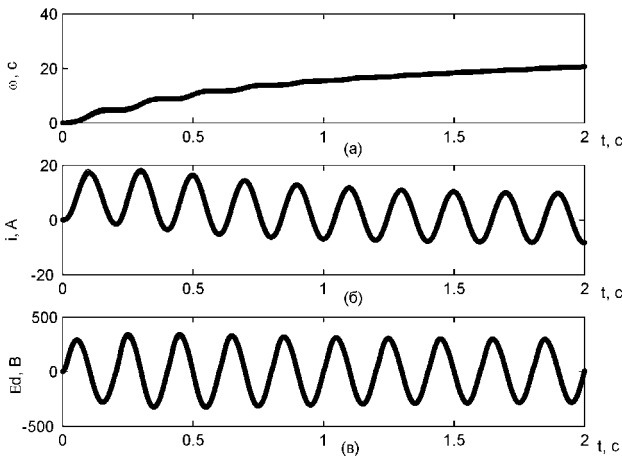


Рисунок 3 – Переходные процессы в нелинейном объекте

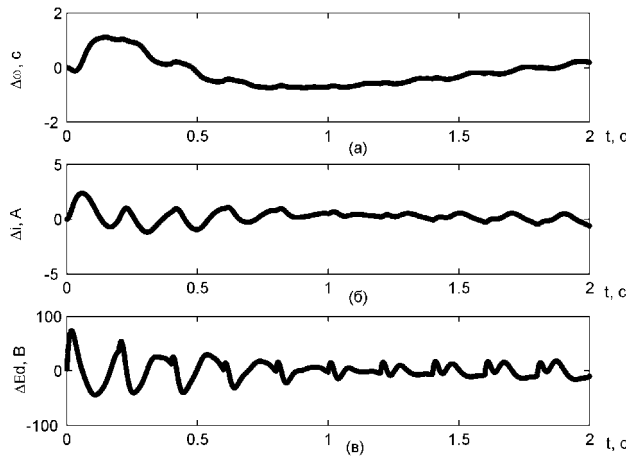


Рисунок 4 – Графики ошибок СРНС с $m = 2$ и $r = 2$

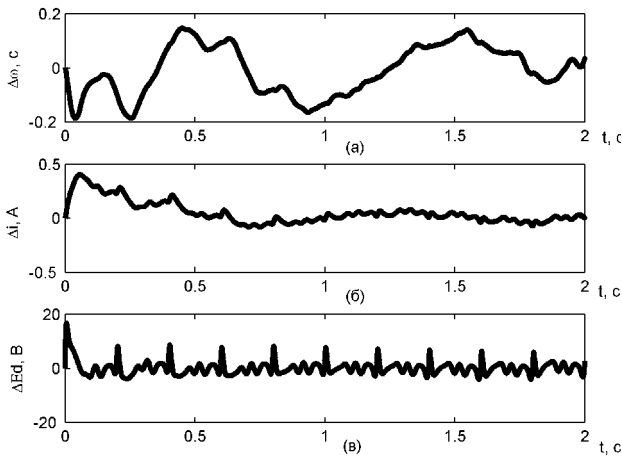


Рисунок 5 – Графики ошибок СРНС с $m = 5$ и $r = 5$

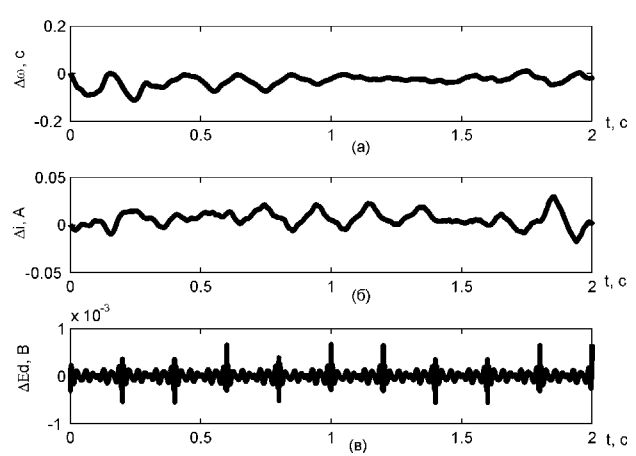


Рисунок 6 – Графики ошибок СРНС с $m = 9$ и $r = 9$

ВЫВОДЫ

1. Приведенные в табл. 2 результаты подтверждают правильность предложенной методики расчета значений весовых коэффициентов СРНС, представляющих модели нелинейного объекта со степенными зависимостями матриц объекта. Относительная ошибка моделей не превышала 0,66 %.

2. При получении моделей на СРНС сложного объекта с произвольными нелинейностями (на примере тиристорного привода с двигателем постоянного тока последовательного возбуждения) установлено, что СРНС на рассматриваемых интервалах времени позволяют получить достаточно точную модель произвольного нелинейного объекта.

Авторы выражают благодарность профессору Запорожского национального технического университета Потапенко Е. М. за ценные рекомендации, во многом способствующие появлению статьи, а также полезные замечания при ее подготовке.

ПЕРЕЧЕНЬ ССЫЛОК

1. Хайкин С. Нейронные сети: полный курс, 2-е издание: Пер. с англ. – М.: Издательский дом «Вильямс», 2006. – 1104 с.
2. Baruch I. S., Palacios I., Flores J. M., Garrido R. An adaptive neural control system of a DC motor drive // 15th Triennial World Congress of the International Federation of Automatic Control, Barcelona, Spain, 2002, Session Slot: T-Th-M04, Area Code: За.

3. Орловский И. А. Модель электропривода постоянного тока на рекуррентных нейронных сетях // Радиоэлектроника, информатика, управління. – 2006. – № 1. – С. 151–159.
4. Орловский И. А., Бут Ю. С. Поиск генетическими алгоритмами весовых коэффициентов в моделях тиристорного электропривода на рекуррентных нейронных сетях // Искусственный интеллект. – 2006. – № 3. – С. 314–326.
5. P. Marino, M. Milano, F. Vasca. Linear Quadratic State Feedback and Robust Neural Network. Estimator for Field-Oriented-Controlled Induction Motors // IEEE Transactions on Industrial Electronics. – 1999. – Vol. 46, No. 1, February. – Pp. 150–161.
6. Палис Ф., Филаретов В. Ф., Цепковский Ю. А., Юхимец Д. А. Синтез прогнозирующих систем с нейрофаззи сетями для управления нелинейными динамическими объектами с переменными параметрами. Электромашиностроение та електрообладнання. – Київ: Техніка. – 2006. – Вип. 66. – С. 357–360.
7. Гантмахер Ф. Р. Теория матриц. – 4-е изд. – М.: Наука. Гл. ред. физ.-мат. лит., 1988. – 522 с.
8. Перельмутер В. М., Сидоренко В. А. Системы управления тиристорными электроприводами постоянного тока. – М.: Энергоатомиздат, 1988. – 304 с.
9. Булгар В. В. Теорія електроприводу: збірник задач / ОНПУ. – Одеса: Поліграф, 2006. – 408 с.

Надійшла 19.01.07

Для отримання моделей нелінійних об'єктів запропоновані степені рекурентні нейронні мережі (СРНМ). Розроблена методика розрахунку таких мереж. Проведена перевірка здібностей СРНМ відображати моделі нелінійних електротехнічних об'єктів.

Degree recurrent neural networks (DRNN) for reception of nonlinear objects models are offered. Methodic of calculation of these networks are received. Checks of ability DRNN are executed to reflect models of nonlinear electro technical objects.

УДК 681.514.6+681.513.6

А. А. Худяев

ВЛИЯНИЕ НЕСЛУЧАЙНОЙ СОСТАВЛЯЮЩЕЙ ПОЛЕЗНОГО СИГНАЛА НА ТОЧНОСТЬ ИТЕРАЦИОННОЙ ДВУХКАНАЛЬНОЙ ВОСПРОИЗВОДЯЩЕЙ СИСТЕМЫ ПРИ СЛУЧАЙНЫХ ВОЗМУЩЕНИЯХ

Рассмотрено влияние неслучайной составляющей полезного сигнала на точность итерационной двухканальной воспроизводящей системы с эталонной настройкой каналов управления при случайных возмущениях. Графически определен закон самонастройки полюсы пропускания второго, компенсирующего, канала на максимально возможную точность работы системы при изменении характеристик входных возмущающих воздействий.

ВВЕДЕНИЕ

Повышение динамической точности воспроизводящих САУ с типовой настройкой контуров управления, работающих при случайных и неслучайных возмущениях, является одной из центральных задач разработки современных унифицированных автоматизированных электромеханических систем. Значительный эффект при этом может быть достигнут за счет введения

многоканальности воспроизведения полезного сигнала $x(t)$ и применения итерационного принципа взаимодействия каналов управления многоканальной САУ [1, 2].

1 ПОСТАНОВКА ЗАДАЧИ

В работах [3–6] введен критерий эффективности и выполнен анализ качества воспроизведения в классе так называемых *неравноточных* итерационных двухканальных систем с эталонной настройкой каналов при случайных входных воздействиях $x(t)$, $f(t)$, $\varphi(t)$. Показано, что при различном задании эталонных операторов автономных каналов управления и параметров принятых типовых характеристик случайных воздействий [3] *выигрыш в точности* G двухканальной системы, получаемый за счет введения и использования второго, точного (или *компенсирующего*), канала W_2^* , тем больше, чем больше показатель $\mu = \frac{\overline{\delta_x^2}}{\overline{\delta_f^2}}$ и меньше параметр $\lambda_0 = S_\varphi(0)/S_f(0)$. Здесь $\overline{\delta_x^2}$ и $\overline{\delta_f^2}$ – дисперсии ошибок первого, грубого (или *основного*), канала W_1^* от действия на него соответственно полезного сигнала $x(t)$ и помехи $f(t)$, приведенной ко входу этого канала; $S_f(0)$ и $S_\varphi(0)$ – начальные (при $\omega = 0$) значения спектральных плотностей помехи $f(t)$ и помехи $\varphi(t)$, приведенной к входу второго, точного, канала.

В [4–6] сделан важный вывод о том, что при узкополосном случайном полезном сигнале $x(t)$ и широкополосных помехах $f(t)$, $\varphi(t)$ (что характерно для большинства воспроизводящих систем) составляющей ошибки итерационной двухканальной системы от полезного сигнала можно пренебречь. При этом оптимальные значения $r_{2\text{опт}}$ перестраиваемого безразмерного *формирующего параметра* (масштабного множителя) точного канала r_2 определяются, в основном, соотношением интенсивностей помех $S_\varphi(0)/S_f(0)$ в точном и грубом каналах, а сама зависимость

$$r_{2\text{опт}}(\lambda_0) = r_1^* z_{\text{опт}}^{-1}(\lambda_0) \tag{1}$$

определяет *закон самонастройки* реальной полосы пропускания компенсирующего канала

$$\omega_{2\text{опт}}(\lambda_0) = 2\pi\Omega_{\varphi 2} r_{2\text{опт}}^{-1}(\lambda_0) \tag{2}$$

на максимально возможные показатели качества работы системы с принятой эталонной настройкой каналов. Здесь $z = r_1/r_2$ – оптимизируемый нормированный формирующий параметр двухканальной системы; $r_1 = r_1^* = \text{const}$ – заданное значение безразмерного формирующего параметра (масштабного множителя)

основного канала; $\Omega_{\varphi 2}$ – эталонная собственная частота компенсирующего канала.

В работах [7, 8] показано, что в неравноточных итерационных многоканальных системах установившаяся динамическая точность повышается с ростом номера i канала воспроизведения W_i^* , $i = \overline{1, N}$ и может быть существенно повышена по сравнению с одноканальной системой W_i^* . Для двухканальной ($N = 2$) системы введен критерий эффективности воспроизведения системой неслучайной составляющей (или математического ожидания) $m_x(t)$ полезного сигнала $x(t)$ в виде *выигрыша в динамической точности*

$$g_d(t) = \frac{\delta_d^2(t)}{\varepsilon_d^2(t)}, \tag{3}$$

равного отношению квадратов установившихся значений динамических ошибок одноканальной и двухканальной систем. Выполнена оценка динамической точности при синусоидальном входном сигнале $m_x(t)$.

Вместе с тем, вопросы учета влияния возможных динамических ошибок от неслучайной составляющей полезного сигнала $m_x(t)$ на оптимальные значения полосы пропускания точного канала $\omega_{2\text{опт}}$ (2) и на суммарный выигрыш в точности G_Σ двухканальной системы при случайных мешающих возмущениях $f(t)$ и $\varphi(t)$ остались нерассмотренными и требуют анализа при решении задачи разработки адаптивной итерационной двухканальной системы с эталонной настройкой каналов.

2 СУММАРНЫЙ ПОКАЗАТЕЛЬ ЭФФЕКТИВНОСТИ (ВЫИГРЫШ В ТОЧНОСТИ) ДВУХКАНАЛЬНОЙ СИСТЕМЫ

Пусть эталонные операторы первого W_1^* и второго W_2^* замкнутых каналов управления итерационной двухканальной воспроизводящей системы (рис. 1) различны и соответствуют колебательному и аperiodическому звеньям с одинаковыми постоянными времени, равными эталонной постоянной времени T_φ . Тогда реальные операторы каналов $W_1^*(p)$, $W_2^*(p)$, где $p \triangleq \frac{d}{dt}$ – оператор дифференцирования по времени t , и спектральные плотности случайных входных возмущений (помех) $f(t)$, $\varphi(t)$ могут быть представлены в виде [4]

$$\begin{aligned} W_1^*(p) &= \frac{1}{(r_1^*)^2 T_\varphi^2 p^2 + 2\xi_1 r_1^* T_\varphi p + 1}, \\ W_2^*(p) &= \frac{1}{r_2 T_\varphi p + 1} = \frac{z}{r_1^* T_\varphi p + z}; \end{aligned} \tag{4}$$

$$S_f(\omega) = \frac{S_f(0)}{\gamma_f^2(r_1^*)^2 T_3^2 \omega^2 + 1},$$

$$S_\varphi(\omega) = \frac{S_\varphi(0)}{\gamma_\varphi^2(r_1^*)^2 T_3^2 \omega^2 + 1}, \quad (5)$$

где

$$z = \frac{r_1^* T_3}{r_2^* T_3} = \frac{r_1^*}{r_2^*} = \frac{\Omega_2}{\Omega_1^*} = \frac{\omega_2}{\omega_1^*}, \quad \omega_1^* = 2\pi\Omega_1^*,$$

$$\omega_2 = 2\pi\Omega_2, \quad \xi_1 = \xi_3; \quad (6)$$

$$\gamma_f = \frac{T_f}{r_1^* T_3} = \frac{\Omega_3}{r_1^* \Omega_f} = \frac{\Omega_1^*}{\Omega_f} = \frac{\omega_1^*}{\omega_f},$$

$$\Omega_f = T_f^{-1}, \quad \omega_f = 2\pi\Omega_f; \quad (7)$$

$$\gamma_\varphi = \frac{T_\varphi}{r_1^* T_3} = \frac{\Omega_3}{r_1^* \Omega_\varphi} = \frac{\Omega_1^*}{\Omega_\varphi} = \frac{\omega_1^*}{\omega_\varphi},$$

$$\Omega_\varphi = T_\varphi^{-1}, \quad \omega_\varphi = 2\pi\Omega_\varphi. \quad (8)$$

Здесь T_3 – эталонная постоянная времени каналов управления, с; T_f и T_φ – постоянные времени корреляции стационарных случайных процессов $f(t)$ и $\varphi(t)$, с; ξ_3 – эталонный коэффициент демпфирования первого канала; $\Omega_1^* = \Omega_3(r_1^*)^{-1}$ и $\Omega_2 = \Omega_3 r_2^{-1}$ – реальные собственные частоты, определяющие реальные полосы пропускания ω_1^* и ω_2 , соответственно заданного (первого) и оптимизируемого (второго) каналов с эталонной настройкой, с⁻¹.

Из (6)–(8) следует, что безразмерный нормированный параметр z характеризует отношение полос пропускания точного и грубого каналов двухканальной системы и определяется отношением их собственных

частот; безразмерные *параметры* γ_f и γ_φ – отношение полосы пропускания основного (грубого) канала к ширине спектров помех ω_f и ω_φ соответственно.

По аналогии с показателем эффективности (выигрышем в точности) G двухканальной системы, характеризующим эффективность введения и использования второго, компенсирующего, канала W_2^* (см. рис. 1) при случайных входных возмущениях $x(t)$, $f(t)$, $\varphi(t)$ [3, 5], при учете неслучайной составляющей полезного сигнала $m_x(t)$ введем *суммарный выигрыш в точности* G_Σ , равный отношению квадратов суммарных ошибок одноканальной $\delta_\Sigma^2(t)$ и двухканальной $\varepsilon_\Sigma^2(t)$ систем:

$$G_\Sigma = \frac{\delta_\Sigma^2}{\varepsilon_\Sigma^2} = \frac{\delta_\Sigma^2 + \overline{\delta_{\text{сл}}^2}}{\varepsilon_\Sigma^2 + \overline{\varepsilon_{\text{сл}}^2}}, \quad (9)$$

где $\delta_\Sigma^2 \equiv \delta_\Sigma^2(t)$, $\varepsilon_\Sigma^2 \equiv \varepsilon_\Sigma^2(t)$ и $\overline{\delta_{\text{сл}}^2} = \overline{\delta_f^2}$, $\overline{\varepsilon_{\text{сл}}^2} = \overline{\varepsilon_f^2} + \overline{\varepsilon_\varphi^2}$ – соответственно квадраты динамической и случайной составляющих ошибок основного канала и итерационной системы в целом; $\overline{\varepsilon_f^2}$ и $\overline{\varepsilon_\varphi^2}$ – дисперсии ошибок двухканальной системы, обусловленные воздействиями помех $f(t)$ и $\varphi(t)$ соответственно.

Выигрыш в точности G_Σ (9) после преобразования удобно представить в виде

$$G_\Sigma = \frac{\delta_\Sigma^2 + \overline{\delta_{\text{сл}}^2}}{\varepsilon_\Sigma^2 + \overline{\varepsilon_{\text{сл}}^2}} =$$

$$= (1 + \mu_\Delta) \frac{1}{\mu_\Delta g_\Delta^{-1} + g_f^{-1} + \lambda g_\varphi^{-1}} = G_{\Sigma 0} \left(\frac{1}{1 - \Delta G_\Sigma} \right), \quad (10)$$

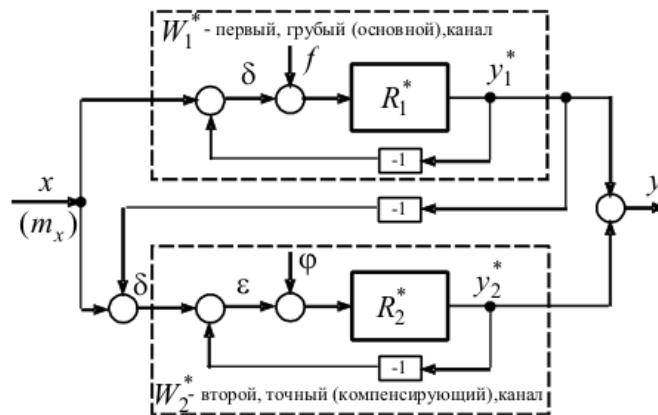


Рисунок 1 – Вариант структурной схемы итерационной двухканальной воспроизводящей системы:

R_1^* и R_2^* – первый и второй автономные разомкнутые каналы управления; y_1^* , y_2^* и y – выходные (управляющие) сигналы соответственно первого автономного канала, второго автономного канала и двухканальной системы в целом

где

$$G_{\Sigma 0} = (1 + \mu_d) \frac{1}{\mu_d g_d^{-1} + g_{0f}^{-1} + \lambda_0 g_{0\phi}^{-1}}; \quad (11)$$

$$\Delta G_{\Sigma} = \frac{1}{1 + \mu_d} \left[\frac{G_{\Sigma 0}}{g_{0f}} \Delta g_f + \frac{G_{\Sigma 0}}{g_{0\phi}} \lambda_0 (\Delta \lambda + \Delta g_{\phi} - \Delta \lambda \Delta g_{\phi}) \right]; \quad (12)$$

$$g_d = \frac{\delta_d^2}{\varepsilon_d^2}, \quad g_f = \frac{\delta_f^2}{\varepsilon_f^2} = g_{0f} \left(\frac{1}{1 - \Delta g_f} \right),$$

$$g_{\phi} = \frac{\delta_{\phi}^2}{\varepsilon_{\phi}^2} = g_{0\phi} \left(\frac{1}{1 - \Delta g_{\phi}} \right); \quad (13)$$

$$\mu_d = \frac{\delta_d^2}{\delta_f^2}, \quad \lambda = \frac{\delta_{\phi}^2}{\delta_f^2} = \lambda_0 (1 - \Delta \lambda). \quad (14)$$

В (10)–(14) обозначены: $\frac{\delta_d^2}{\varepsilon_d^2}$ – дисперсия ошибки основного (грубого) канала W_1^* при условном воздействии на него помехи $\phi(t)$, приложенной ко входу компенсирующего (точного) канала W_2^* ; g_f – показатель эффективности (выигрыш в точности), характеризующий эффективность использования точного канала для компенсации ошибки по помехе $f(t)$; g_{ϕ} и μ_d , λ – показатели, определяющие соответственно нормированную дополнительную ошибку, вносимую помехой $\phi(t)$, и относительную точность первого, грубого, канала; $G_{\Sigma 0}$ и ΔG_{Σ} – первое приближение и относительная поправка к показателю эффективности G_{Σ} ; g_{0f} , $g_{0\phi}$, λ_0 и Δg_f , Δg_{ϕ} , $\Delta \lambda$ – соответственно первые приближения и поправки к показателям g_f , g_{ϕ} , λ , значения которых для случая (4), (5) получены в работе [4]. Из (10), (13), (14) видно, что значениями поправок можно пренебречь, если их модули достаточно малы по сравнению с единицей.

Обычно в воспроизводящих САУ полезный сигнал $m_x(t)$ узкополосен, а помехи $f(t)$, $\phi(t)$ близки к процессам типа «белого шума», т. е. $T_f \ll r_1^* T_{\Sigma}$, $T_{\phi} \ll r_1^* T_{\Sigma}$, и $\gamma_f \ll 1$, $\gamma_{\phi} \ll 1$. Кроме того, в итерационных воспроиз-

водящих системах, как правило, $z \geq 1$, что обеспечивается закономерным расширением полосы пропускания более быстродействующего компенсирующего канала по сравнению с основным. Тогда для (4)–(8), (13), (14) в наиболее *неблагоприятном случае*, когда помехи $f(t)$ и $\phi(t)$ представляют собой «белый шум», выполняются равенства

$$\gamma_f = \gamma_{\phi} = 0, \quad \Delta g_f = \Delta g_{\phi} = \Delta \lambda = 0; \quad (15)$$

и

$$g_f = g_{0f} = z^2 + 2\xi_{\Sigma} z + 1, \quad g_{\phi} = g_{0\phi} = \frac{1}{2\xi_{\Sigma} z},$$

$$\lambda = \lambda_0. \quad (16)$$

Из (10)–(12), (15), (16) получим $\Delta G_{\Sigma} = 0$ и для оценки выигрыша в точности G_{Σ} окончательно найдем

$$G_{\Sigma} = G_{\Sigma 0} = \frac{1 + \mu_d}{2\xi_{\Sigma} \lambda_0} \times$$

$$\times \frac{z^2 + 2\xi_{\Sigma} z + 1}{\left(\frac{\mu_d g_d^{-1}}{2\xi_{\Sigma} \lambda_0} + z \right) (z^2 + 2\xi_{\Sigma} z + 1) + \frac{1}{2\xi_{\Sigma} \lambda_0}}. \quad (17)$$

На практике часто случайная помеха $f(t)$, воздействующая на основной канал W_1^* , может иметь относительно узкополосный спектр $S_f(\omega)$, ширина которого $\omega_f = 2\pi T_f^{-1}$ сопоставима с полосой пропускания основного канала $\omega_1^* = 2\pi(r_1^* T_{\Sigma})^{-1}$. Полагая в этом случае $\gamma_f \neq 0$ и $\gamma_{\phi} = 0$, вместо (15), (16) имеем [4]

$$\Delta g_f \neq 0, \quad \Delta \lambda \neq 0, \quad \Delta g_{\phi} = 0$$

и

$$g_f = g_{0f} \left(1 - \frac{c_{\gamma_f} z}{\gamma_f z + 1} \right)^{-1}, \quad \lambda = \lambda_0 \left(1 + \frac{c_{\gamma_f}}{2\xi_{\Sigma}} \right),$$

$$g_{\phi} = g_{0\phi}, \quad (18)$$

где

$$c_{\gamma_f} = \frac{2\xi_{\Sigma} \gamma_f^2}{1 + 2\xi_{\Sigma} \gamma_f}.$$

Тогда из (10), (18) с учетом (16) для выигрыша в точности G_{Σ} получим

$$G_{\Sigma} = \frac{1 + \mu_d}{2\xi_{\Sigma} \lambda_0 \left(1 + \frac{c_{\gamma_f}}{2\xi_{\Sigma}} \right)} \cdot \frac{(z^2 + 2\xi_{\Sigma} z + 1)(\gamma_f z + 1)}{\left[\frac{\mu_d g_d^{-1}}{2\xi_{\Sigma} \lambda_0} \left(1 + \frac{c_{\gamma_f}}{2\xi_{\Sigma}} \right)^{-1} + z \right] (z^2 + 2\xi_{\Sigma} z + 1)(\gamma_f z + 1) + \left[2\xi_{\Sigma} \lambda_0 \left(1 + \frac{c_{\gamma_f}}{2\xi_{\Sigma}} \right) \right]^2 [(\gamma_f - c_{\gamma_f})z + 1]} \quad (19)$$

Заметим, что полученные соотношения (17) и (19) для суммарного выигрыша в точности G_Σ с погрешностью до их динамических составляющих совпадают с соответствующими выражениями для выигрыша в точности G , полученными в работе [5] для случая (4), (5) при чисто случайных входных возмущениях. Кроме того, из (19) при выполнении условия $\gamma_f = 0$ получаем выражение (17).

3 НОРМИРОВАННОЕ ОТНОСИТЕЛЬНОЕ УСКОРЕНИЕ НЕСЛУЧАЙНОЙ СОСТАВЛЯЮЩЕЙ ПОЛЕЗНОГО СИГНАЛА

Не нарушая общности, будем полагать, что при расчете динамических ошибок двухканальной системы можно ограничиться только ошибками по скорости

$v_{m_x}(t) = \frac{dm_x(t)}{dt}$ и ускорению $a_{m_x}(t) = \frac{d^2m_x(t)}{dt^2}$ входного сигнала $m_x(t)$. Тогда для случая (4) с учетом (6) и результатов работы [8] получим следующие выражения для оценки динамических ошибок основного канала (при $0 < \xi_\vartheta \ll 1$) $\delta_d(t)$ и итерационной двухканальной системы в целом $\varepsilon_d(t)$:

$$\delta_d(t) = \delta_v(t) + \delta_a(t) = \frac{r_1^*}{k_{v1}^*} v_{m_x}(t) + \frac{(r_1^*)^2}{k_{v1}^*} \cdot \frac{T_\vartheta}{2\xi_\vartheta} a_{m_x}(t); \quad (20)$$

$$\varepsilon_d(t) = \varepsilon_a(t) = \frac{r_1^* r_2^*}{k_a^*} a_{m_x}(t) = \frac{r_1^* r_2^*}{k_{v1}^* k_{v2}^*} a_{m_x}(t), \quad (21)$$

где $\delta_v(t)$ и $\delta_a(t)$ – составляющие динамической ошибки основного канала по скорости и ускорению;

$$k_{v1}^* = \frac{1}{2\xi_\vartheta T_\vartheta}, \quad k_{v2}^* = \frac{1}{T_\vartheta}, \quad (22)$$

– эталонные коэффициенты усиления по скорости соответственно основного W_1^* и компенсирующего W_2^* каналов. Для ошибки $\varepsilon_d(t)$ из (21), (22) найдем

$$\varepsilon_d(t) = \frac{2\xi_\vartheta (r_1^* T_\vartheta)^2}{z} a_{m_x}(t)$$

или с учетом (20)

$$\varepsilon_d(t) = \frac{1}{z} q \delta_v(t), \quad (23)$$

где

$$q \equiv q(t) = \frac{\Delta v_{m_x}(t)}{v_{m_x}(t)} = \frac{T_1 a_{m_x}(t)}{v_{m_x}(t)},$$

$$\delta_v(t) = 2\xi_\vartheta T_1 v_{m_x}(t), \quad T_1 = r_1^* T_\vartheta. \quad (24)$$

Здесь $\Delta v_{m_x}(t) = T_1 a_{m_x}(t)$ – приращение скорости входного сигнала $m_x(t)$ за период T_1 собственных колебаний основного (грубого) канала, рад/с; q – относительное приращение скорости.

Оценим величину q . Пусть $m_x(t)$ – синусоидальный сигнал с периодом T_{m_x} . Тогда $m_x(t) = A_m \sin 2\pi T_{m_x}^{-1} t$. В этом случае из (24) получим для максимальных значений скорости и ускорения

$$q = 2\pi \frac{T_1}{T_{m_x}} = 2\pi \frac{\Omega_{m_x}}{\Omega_1^*}.$$

При сигнале $m_x(t)$, частота которого $\Omega_{m_x} = T_{m_x}^{-1}$ на полтора – три порядка меньше собственной частоты основного канала Ω_1^* , имеем $q \approx (0,005 \div 0,125)$.

Ошибку основного канала $\delta_d(t)$ (20) с учетом (22), (24) можно представить в виде

$$\delta_d(t) = \delta_v(t) \left(1 + \frac{1}{2\xi_\vartheta} q\right). \quad (25)$$

Следовательно, для выигрыша в динамической точности $g_d(t)$ из (3), (23), (25) получим

$$g_d(t) \equiv g_d = \frac{\left(1 + \frac{1}{2\xi_\vartheta} q\right)^2}{q^2} z^2. \quad (26)$$

Для дисперсии ошибки $\overline{\delta_f^2}$ от широкополосной ($\gamma_f = 0$) помехи $f(t)$ с учетом (4)–(7) найдем [3]

$$\overline{\delta_f^2} = \frac{S_f(0)}{4\xi_\vartheta r_1^* T_\vartheta}, \quad (27)$$

где [1]

$$S_f(0) = 2\sigma_f^2 T_f; \quad (28)$$

σ_f – среднее квадратическое отклонение случайного процесса $f(t)$. Тогда для показателя μ_d относительной динамической точности основного канала из (14), (25), (27) получим

$$\begin{aligned} \mu_d \equiv \mu_d(t) &= \frac{\delta_d^2(t)}{\overline{\delta_f^2}} = \frac{\delta_v^2(t)}{\overline{\delta_f^2}} \left(1 + \frac{1}{2\xi_\vartheta} q\right)^2 = \\ &= \mu_{d,v}(t) \left(1 + \frac{1}{2\xi_\vartheta} q\right)^2, \end{aligned} \quad (29)$$

где, с учетом (24),

$$\mu_{д,в}(t) = \frac{\delta_v^2(t)}{\delta_f^2} = (2\xi_3 r_1^* T_3) \frac{32v_{m_x}^2(t)}{S_f(0)} \quad (30)$$

или, с учетом (28),

$$\mu_{д,в}(t) = (2\xi_3 r_1^* T_3) \frac{3v_{m_x}^2(t)}{T_f \sigma_f^2}.$$

Оценим показатели $\mu_{д,в}(t)$ и $\mu_d(t)$. Пусть $\xi_3 = 0,5$; $r_1^* = 0,1$; $T_3 = 1$ с. При синусоидальном сигнале $m_x(t) = A_m \sin 2\pi T_{m_x}^{-1} t$, период T_{m_x} которого на полтора – три порядка больше периода $T_1 = r_1^* T_3$ собственных колебаний основного канала, для максимальных значений скорости $v_{m_x}(t)$ и ускорения $a_{m_x}(t)$ найдем (при $A_m = 1$)

$$\begin{aligned} v_{m_x} &= 2\pi T_{m_x}^{-1} \approx (0,06 \div 1,4), \text{ рад/с;} \\ a_{m_x} &= 4\pi^2 T_{m_x}^{-2} \approx (0,00125 \div 0,5), \text{ рад/с}^2. \end{aligned} \quad (31)$$

Тогда из (29)–(30) при $S_f(0) = 10^{-6}$ рад²·с и $q = 0,05$ для максимальных значений показателей $\mu_{д,в}(t)$ и $\mu_d(t)$ получим

$$\begin{aligned} \mu_{д,в} &= 2 \cdot 10^3 \cdot v_{m_x}^2 \approx (7 \div 4 \cdot 10^3), \\ \mu_d &= \mu_{д,в} (1+q)^2 \approx (7,7 \div 4,4 \cdot 10^3). \end{aligned} \quad (32)$$

Так, например, при $v_{m_x} = 0,6$ рад/с и $q = 0,05$ из (32) имеем $\mu_{д,в} = 720$, $\mu_d \approx 800$.

Составляющую $\mu_d g_d^{-1}$ выигрыша в точности G_Σ в (17), (19) на основании (26), (29) можно представить следующим образом:

$$\mu_d g_d^{-1} = \mu_{д,в} q^2 \frac{1}{z^2} = \aleph_{m_x} z^{-2}, \quad (33)$$

где, с учетом (24), (30),

$$\aleph_{m_x} \equiv \aleph_{m_x}(t) = \mu_{д,в}(t) q^2 = 16\xi_3^3 (r_1^* T_3) \frac{5a_{m_x}^2(t)}{S_f(0)}. \quad (34)$$

Обозначая

$$a_f^2 = \frac{S_f(0)}{16\xi_3^3 (r_1^* T_3)^5}, \quad (35)$$

для \aleph_{m_x} согласно (34) окончательно найдем

$$\aleph_{m_x} = \frac{a_{m_x}^2}{a_f^2}. \quad (36)$$

Из (33)–(35) с учетом (13), (14) нетрудно видеть, что размерность величины a_f соответствует размерности ускорения a_{m_x} , а ее значение численно равно тому значению ускорения полезного сигнала $a_{m_x}(t) = d^2 a_{m_x}(t)/dt^2$, при котором для $z = 1$ выполняется равенство $\varepsilon_d^2 = \overline{\delta_f^2}$. В частности, при $\xi_3 = 0,5$; $r_1^* = 0,1$; $T_3 = 1$ с; $S_f(0) = 10^{-6}$ рад²·с из (35) найдем: $a_f \approx 0,224$ рад/с² $\approx 12,81$ угл. град/с².

Безразмерную величину \aleph_{m_x} (36), определяющую на основании (33) влияние параметров компенсирующего канала W_2^* на динамическую точность итерационной системы, можно рассматривать как *нормированное относительное ускорение неслучайной составляющей полезного сигнала $m_x(t)$* .

4 ЭКСТРЕМАЛЬНЫЕ ХАРАКТЕРИСТИКИ СУММАРНОГО ПОКАЗАТЕЛЯ ЭФФЕКТИВНОСТИ

Определим теперь из (17), (33) выигрыш в точности G_Σ в виде

$$G_\Sigma = \frac{1 + \mu_d}{2\xi_3 \lambda_0} \times \frac{z^2 g_{0f}(z, \xi_3)}{z^3 g_{0f}(z, \xi_3) + \frac{1}{2\xi_3 \lambda_0} [z^2 + \aleph_{m_x} g_{0f}(z, \xi_3)]}, \quad (37)$$

где

$$g_{0f}(z, \xi_3) = z^2 + 2\xi_3 z + 1. \quad (38)$$

Из (32), (34) при $q = 0,05$ зададим примерный диапазон изменения относительного ускорения \aleph_{m_x} : $\aleph_{m_x} \approx (0 \div 10)$. На рис. 2, а, б приведены графики зависимости величины $(1 + \mu_d)^{-1} G_\Sigma$ от параметров z и \aleph_{m_x} при $\xi_3 = 0,5$ и заданных значениях $\lambda_0 = 0,3$ и $\lambda_0 = 1$ соответственно. Графики иллюстрируют влияние учета динамической ошибки, характеризуемой относительным ускорением полезного сигнала $\aleph_{m_x} = a_{m_x}^2/a_f^2$, на вид экстремальных характеристик выигрыша в точности G_Σ (37). Линии, соединяющие точки экстремумов, определяют оптимальные с точки зрения качества двух-

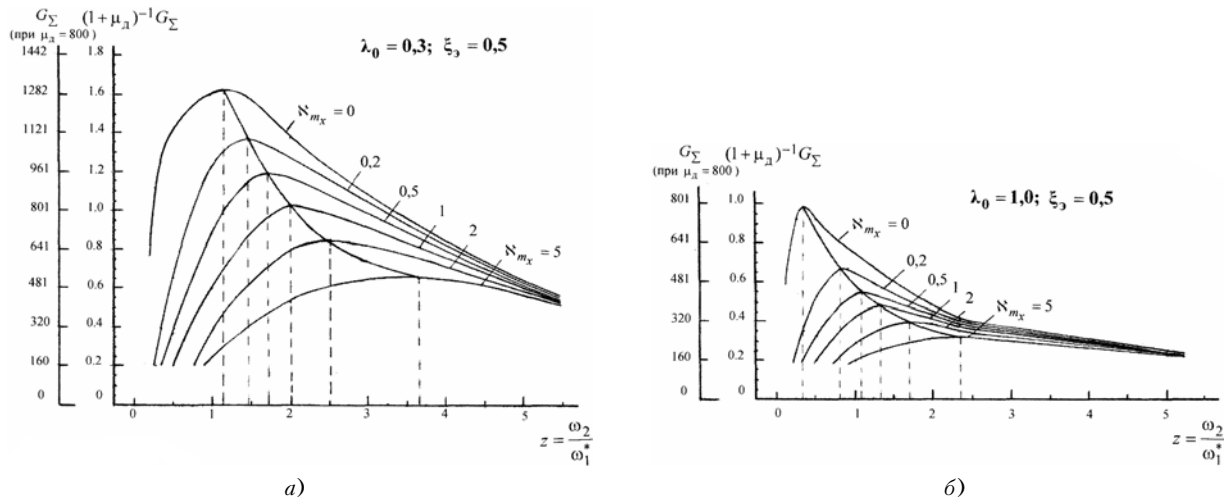


Рисунок 2 – Графики зависимости величины $(1 + \mu_d)^{-1} G_\Sigma$ от параметров z и N_{m_x} при $\xi_3 = 0,5$ и заданных значениях $\lambda_0 = 0,3$ и $\lambda_0 = 1$ и при $\gamma_f = \gamma_\phi = 0$

канальной системы значения $z_{\text{опт}}$ и $(1 + \mu_d)^{-1} G_\Sigma$ max для различных значений N_{m_x} .

Из графиков на рис. 2, а, б видно, что более глубокий экстремум выигрыша в точности G_Σ имеет место при достаточно малых N_{m_x} ($0 \leq N_{m_x} < 1$), при которых оптимальное значение нормированного параметра $z_{\text{опт}} < 2$. С ростом ускорения a_{m_x} ($N_{m_x} > 0$) уменьшаются крутизна и максимальное значение характеристики $G_\Sigma(z)$, система становится менее чувствительной к изменению полосы пропускания компенсирующего канала ω_2 . Вместе с тем, при неизменной интенсивности случайных помех в каналах ($\lambda_0 = \text{const}$) с ростом N_{m_x} ($N_{m_x} \geq 1$), т. е. с ростом удельного веса динамической составляющей ошибки в суммарном функционале качества G_Σ (37), соответствующие оптимальные значения полосы пропускания $\omega_{2\text{опт}} = \omega_1^* z_{\text{опт}}$ существенно увеличиваются по сравнению с оптимальным значением, полученным при отсутствии учета влияния на выигрыш в точности G_Σ динамической составляющей $g_d(t)$ (3) ($N_{m_x} = 0$), что физически очевидно.

Оценим удельный вес динамической ошибки. Для отношения квадратов динамической и случайной составляющих ошибки итерационной системы с учетом (9), (13), (14) получим

$$\frac{\varepsilon_d^2}{\varepsilon_{\text{сл}}^2} = \frac{\varepsilon_d^2}{\varepsilon_f^2 + \varepsilon_\phi^2} = \frac{\mu_d g_d^{-1}}{g_f^{-1} + \lambda g_\phi^{-1}}$$

или, согласно (15), (16), (33),

$$\frac{\varepsilon_d^2}{\varepsilon_{\text{сл}}^2} = \frac{N_{m_x}}{2\xi_3\lambda_0} \cdot \frac{g_0 f(z, \xi_3)}{z^2 \left[z \cdot g_0 f(z, \xi_3) + \frac{1}{2\xi_3\lambda_0} \right]}. \quad (39)$$

Отсюда и из рис. 2, а, б нетрудно видеть, что при $N_{m_x} = \text{const}$ с расширением полосы пропускания компенсирующего канала ω_2 ($\omega_2 > \omega_{2\text{опт}}$) удельный вес динамической ошибки ε_d двухканальной системы уменьшается, и одновременно увеличивается негативное влияние на точность системы некомпенсируемой высокочастотной помехи $\phi(t)$ в этом канале. При $\xi_3 = 0,5$ и $z = 1$, когда с учетом (6) оба канала имеют практически равные полосы пропускания, из (38), (39) найдем

$$\frac{\varepsilon_d^2}{\varepsilon_{\text{сл}}^2} = \frac{N_{m_x}}{\lambda_0} \cdot \frac{z^2 + z + 1}{z^2 \left(z^3 + z^2 + z + \frac{1}{\lambda_0} \right)} = \frac{N_{m_x}}{\lambda_0 + \frac{1}{3}}. \quad (40)$$

При изменении соотношения интенсивностей случайных помех в каналах λ_0 от нуля [$S_\phi(0) = 0$] до единицы [$S_\phi(0) = S_f(0)$] из (40) имеем $\frac{\varepsilon_d^2}{\varepsilon_{\text{сл}}^2} = \left(3 \div \frac{3}{4} \right) N_{m_x}$.

В частности, при $z = 1$ и $\lambda_0 = \frac{2}{3} \approx 0,67$ отношение

$\frac{\varepsilon_d^2}{\varepsilon_{\text{сл}}^2} = N_{m_x}$, и значения величины $\sqrt{N_{m_x}}$ определяют

удельный вес динамической ошибки итерационной системы при наличии случайных помех в ее каналах.

Из (37), (38) после дифференцирования и преобразования результата к удобному для дальнейшего анализа виду для $\frac{\partial G_{\Sigma}}{\partial z}$ получим

$$\frac{\partial G_{\Sigma}}{\partial z} = \frac{1 + \mu_d}{2\xi_3} \times \frac{z[P(z, \xi_3, \lambda_0) + L(z, \xi_3, \aleph_{m_x})]}{\left\{ \lambda_0 z^3 g_{0f}(z, \xi_3) + \frac{1}{2\xi_3} [z^2 + \aleph_{m_x} g_{0f}(z, \xi_3)] \right\}^2}, \quad (41)$$

где

$$P(z, \xi_3, \lambda_0) = \lambda_0 z^3 P_0(z, \xi_3, \lambda_0) = -z^3 \left[\lambda_0 (z^2 + 2\xi_3 z + 1)^2 - \frac{1}{\xi_3} z - 1 \right], \quad (42)$$

$$L(z, \xi_3, \aleph_{m_x}) = \frac{\aleph_{m_x}}{\xi_3} (z^2 + 2\xi_3 z + 1)^2. \quad (43)$$

Экстремальные значения параметра z могут быть определены из решения уравнения

$$z[P(z, \xi_3, \lambda_0) + L(z, \xi_3, \aleph_{m_x})] = 0.$$

Отсюда находим $z_{\text{экт},1} = 0$, что не представляет практического интереса, так как с учетом (4), (6) соответствует случаю сужения полосы пропускания второго, компенсирующего, канала ω_2 до нуля и переходу от двухканальной системы к одноканальной. При этом $G_{\Sigma} = 0$. Экстремальное значение $z_{\text{экт},2} \equiv z_{\text{опт}} = z_{\text{опт}}(\xi_3, \lambda_0, \aleph_{m_x})$, при котором (для заданных $\xi_3, \lambda_0, \aleph_{m_x}$) достигается максимальное значение $G_{\Sigma\text{max}}$, может быть определено из уравнения

$$P(z, \xi_3, \lambda_0) = -L(z, \xi_3, \aleph_{m_x}), \quad (44)$$

в котором параметры λ_0 и \aleph_{m_x} функционально разделены, что удобно для его графического решения.

На рис. 3 приведены результаты графического решения уравнения (44) относительно формирующего параметра итерационной системы z ($z_{\text{опт}} = z_{\text{опт}}(\lambda_0, \aleph_{m_x})$), полученные из (42), (43) с учетом результатов работы [1] при $\xi_3 = 0,5$ для значений $\lambda_0 = 0, 1 \div 1$ и $\aleph_{m_x} = 0 \div 10$. Из рис. 3 видно, что с уменьшением λ_0 и увеличением \aleph_{m_x} значения $z_{\text{опт}}$ возрастают. Учет влияния ускорения a_{m_x} неслучайной составляющей полезного сигнала приводит (при $\lambda_0 = \text{const}$) к необходимости расширения полосы пропускания точного канала ω_2 в несколько раз. Так, например, при $\aleph_{m_x} = 1$ для $0,3 \leq \lambda_0 \leq 0,5$ значения $\omega_{2\text{опт}} = \omega_1^* z_{\text{опт}}$ увеличиваются примерно в два

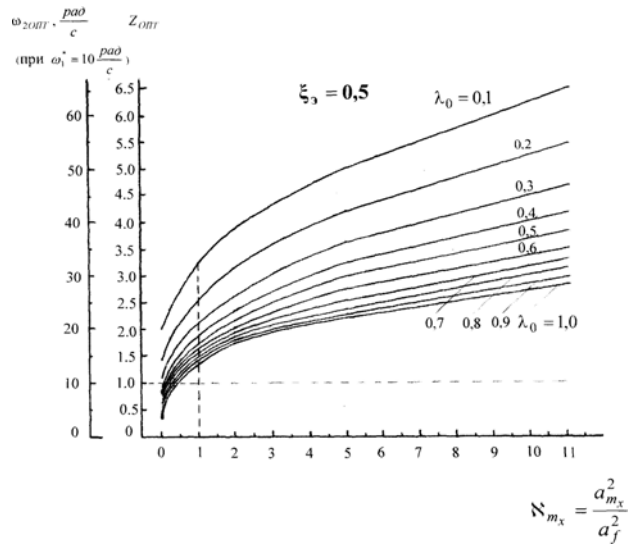


Рисунок 3 – График зависимости оптимальных значений $z_{\text{опт}} = z_{\text{опт}}(\lambda_0, \aleph_{m_x})$ при $\xi_3 = 0,5$ и $\gamma_f = \gamma_{\phi} = 0$

раза, а для $0,8 \leq \lambda_0 \leq 1,0$ более чем в три раза, по сравнению со случаем, когда составляющая динамической ошибки не учитывается ($\aleph_{m_x} = 0$).

Если отношение $\aleph_{m_x}(t) = a_{m_x}^2(t)/a_f^2$ может быть прямо или косвенно измерено, то при известных неизменяемых параметрах основного канала (ξ_3, r_1^*) зависимость

$$r_{2\text{опт}}(\lambda_0, \aleph_{m_x}, t) = r_1^* z_{\text{опт}}^{-1}(\lambda_0, \aleph_{m_x}, t) \quad (45)$$

определяет закон самонастройки формирующего параметра r_2 компенсирующего канала на максимум суммарного показателя эффективности G_{Σ} итерационной двухканальной системы.

Подставляя полученные из графиков на рис. 3 (при $\xi_3 = 0,5$) оптимальные значения $z_{\text{опт}} = z_{\text{опт}}(\lambda_0, \aleph_{m_x})$ в (37), найдем

$$G_{\Sigma\text{max}} = \frac{1 + \mu_d}{\lambda_0} \times \frac{z_{\text{опт}}^2 g_{0f}^*(z_{\text{опт}})}{z_{\text{опт}}^3 g_{0f}^*(z_{\text{опт}}) + \frac{1}{\lambda_0} [z_{\text{опт}}^2 + \aleph_{m_x} g_{0f}^*(z_{\text{опт}})]}, \quad (46)$$

где

$$g_{0f}^*(z_{\text{опт}}) = z_{\text{опт}}^2 + z_{\text{опт}} + 1. \quad (47)$$

На основании (46), (47) построим графики величины $(1 + \mu_d)^{-1} G_{\Sigma\text{max}} = F^*(\lambda_0, \aleph_{m_x})$ (рис. 4), определяющей максимальные значения $G_{\Sigma\text{max}}$ суммарного выигрыша в точности (37) при $\xi_3 = 0,5$ и заданном

значении показателя μ_d . Из графиков видно, что с увеличением параметров λ_0 и \aleph_{m_x} максимальные значения $G_{\Sigma \max}$ уменьшаются. Так, для $\lambda_0 = 0,3$ при $\aleph_{m_x} = 0$ и $\aleph_{m_x} = 5$ имеем $(1 + \mu_d)^{-1} G_{\Sigma \max} \approx 1,58$ и $(1 + \mu_d)^{-1} \times G_{\Sigma \max} \approx 0,65$; а уже для $\lambda_0 = 1,0$ при тех же значениях \aleph_{m_x} соответственно получим $(1 + \mu_d)^{-1} G_{\Sigma \max} \approx 0,97$ и $(1 + \mu_d)^{-1} G_{\Sigma \max} \approx 0,31$. Эти результаты совпадают с результатами, приведенными на рис. 2.

Из (9), (46), (47) и из графиков на рис. 4 может быть рассчитано максимальное значение отношения суммарных ошибок итерационной системы $\frac{\delta_{\Sigma}}{\varepsilon_{\Sigma}}$:

$$\left(\frac{\delta_{\Sigma}}{\varepsilon_{\Sigma}}\right)_{\max} = \sqrt{G_{\Sigma \max}},$$

при этом для диапазона значений $\lambda_0 > 0,1$ и $0 \leq \aleph_{m_x} < 1$ выполняется неравенство

$$\left(\frac{\delta_{\Sigma}}{\varepsilon_{\Sigma}}\right)_{\max} > \frac{1}{2} \sqrt{\frac{1 + \mu_d}{\lambda_0}}.$$

Например, при $\lambda_0 = 0,3$; $\mu_d = 800$ и $\aleph_{m_x} = 1$ имеем $\left(\frac{\delta_{\Sigma}}{\varepsilon_{\Sigma}}\right)_{\max} = 28,44$, т. е. суммарная ошибка итерационной двухканальной системы $\varepsilon_{\Sigma} = \sqrt{\varepsilon_d^2 + \varepsilon_{\text{сл}}^2}$ при оптимальной настройке параметров компенсирующего канала уменьшается более чем в 28 раз по сравнению с сум-

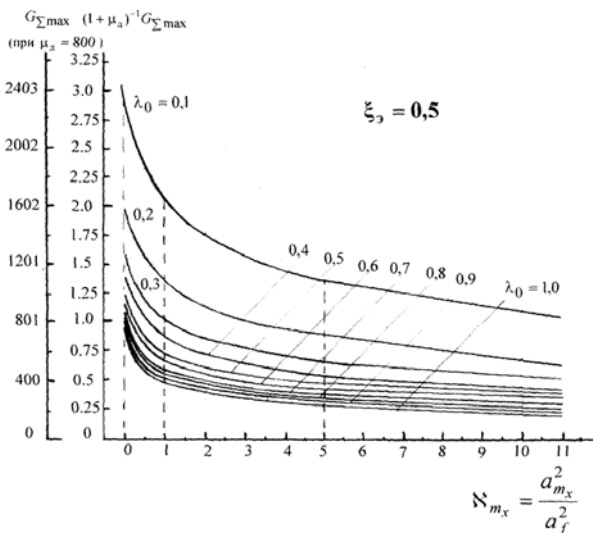


Рисунок 4 – Графики зависимости максимальных значений $(1 + \mu_d)^{-1} G_{\Sigma \max} = F^*(\lambda_0, \aleph_{m_x})$ при $\xi_3 = 0,5$ и $\gamma_f = \gamma_{\phi} = 0$

марной ошибкой основного канала $\delta_{\Sigma} = \sqrt{\delta_d^2 + \delta_{\text{сл}}^2}$. Вместе с тем, это требует расширения полосы пропускания второго канала ω_2 примерно в 2 раза по сравнению с полосой пропускания первого ω_1^* (при $\lambda_0 = 0,3$, $z_{\text{опт}} \approx 2,08$; см. рис. 3).

Отметим, что в случае узкополосной помехи $f(t)$, когда для возмущений системы выполняются соотношения $\gamma_f \neq 0$, $\gamma_{\phi} = 0$, из (19), (33) найдем зависимость

$$(1 + \mu_d)^{-1} G_{\Sigma} = F(z, \lambda_0, \gamma_f, \aleph_{m_x}),$$

характер изменения кривых которой для заданных значений $\gamma_f \neq 0$ аналогичен характеру кривых на рис. 2. Однако, при $\gamma_f = \text{const} > 0$ значения ординат зависимости $(1 + \mu_d)^{-1} G_{\Sigma \max} = F^*(\lambda_0, \gamma_f, \aleph_{m_x})$ увеличиваются по сравнению со значениями ординат аналогичных кривых (при $\gamma_f = 0$), приведенных на рис. 4, что физически очевидно. При $\gamma_f = \text{var}$ ($0 < \gamma_f \leq 1$) и $\lambda_0 = \text{const}$ значения характеристик величины $(1 + \mu_d)^{-1} G_{\Sigma} = F(z, \gamma_f, \aleph_{m_x})$ отличаются друг от друга незначительно! Это подтверждает выводы, сделанные в работе [6], и позволяет не учитывать параметр γ_f в законе самонастройки (45).

ВЫВОДЫ

1. Получены экстремальные характеристики суммарного показателя эффективности G_{Σ} , и построены графики закона самонастройки (45) формирующего параметра r_2 второго, компенсирующего, канала, определяющие влияние неслучайной составляющей (математического ожидания) полезного сигнала $m_x(t)$ на точность итерационной двухканальной системы при наличии ошибок от случайных возмущений (помех) в каналах с принятой типовой настройкой.

2. Показано, что при учете $m_x(t)$ оптимальное значение полосы пропускания второго канала $\omega_{2\text{опт}}$ зависит как от соотношения $\lambda_0 = S_{\phi}(0)/S_f(0)$ интенсивностей широкополосных помех $\phi(t)$ и $f(t)$ в компенсирующем (точном) W_2^* и основном (грубом) W_1^* каналах, так и от величины нормированного относительного ускорения $\aleph_{m_x} = a_{m_x}^2/a_f^2$ неслучайного полезного сигнала. При этом для заданного $\lambda_0 = \text{const}$ значения $\omega_{2\text{опт}} = 2\pi\Omega_3 r_{2\text{опт}}^{-1}$ тем больше, чем больше значение параметра \aleph_{m_x} .

Для логического завершения выполненных исследований необходима разработка структурно-функциональной схемы адаптивной системы, позволяющей

перейти к рассмотрению вопросов практической реализации итерационных двухканальных воспроизводящих систем с самонастройкой второго, компенсирующего, канала управления.

ПЕРЕЧЕНЬ ССЫЛОК

1. Осмоловский П. Ф. Итерационные многоканальные системы автоматического управления. – М.: Сов. Радио, 1969. – 256 с.
2. Многоканальные итерационные системы управления / Б. И. Кузнецов, А. А. Худяев, И. Н. Богаенко и др. – К.: НПК «КИА», 1998. – 224 с.
3. Худяев А. А. Критерии эффективности итерационных двухканальных воспроизводящих систем при случайных воздействиях // Электромашиностроение и электрооборудование. – 2002. – № 58. – С. 92–96.
4. Худяев А. А. Влияние параметров случайных воздействий и полосы пропускания точного канала на качество итерационной двухканальной системы с эталонной настройкой // Радиотехника. Информатика. Управление. – 2002. – № 2(8). – С. 148–156.
5. Худяев А. А. Суммарные функционалы качества итерационных двухканальных систем управления с эталонной настройкой и при случайных воздействиях // Радиотехника. Информатика. Управление. – 2003. – № 1(9). – С. 142–149.
6. Худяев А. А. Экстремальные характеристики качества и законы самонастройки точного канала итерационной двухканальной системы управления при случайных

- воздействиях // Радиотехника. Информатика. Управление. – 2003. – № 2(10). – С. 132–138.
7. Худяев А. А. Коэффициенты ошибок неравноточных итерационных многоканальных систем управления // Электромашиностроение и электрооборудование. – 2003. – Вып. 60. – С. 54–60.
8. Худяев А. А., Гвоздева Е. В., Момот А. П. Оценка установившейся динамической точности итерационной многоканальной воспроизводящей системы // Вестник НТУ «ХПИ». Сборник научных трудов. Тематический выпуск «Проблемы автоматизированного электропривода. Теория и практика». В 2-х томах. – Харьков: НТУ «ХПИ», 2003. – Вып. 10. – Т. 2. – С. 351–355.

Надійшла 4.09.06

Розглянуто дію невинувачливої складової корисного сигналу на точність ітераційної двоканальної відтворювальної системи з еталонним настроюванням каналів керування при випадкових збуреннях. Графічно визначено закон самонастроювання смуги пропускання другого, компенсуючого каналу на максимально можливу точність роботи системи при змінюванні характеристик вхідних збурювальних впливів.

The influence of setting (useful) signal nonrandom component on the precision of the iterative type two-channel control system with basic standard tuning of channels by random noises is considered. The law of second, compensate, channel bandwidth self-adjusting on the most possible precision of control system operation upon changing of input actions characteristics is graphically defined.

УДК 681.5.015: 681.511.4

Е. А. Шушляпин, Е. В. Виноградов

ОБОБЩЕНИЕ МЕТОДА КОНЕЧНОГО СОСТОЯНИЯ НА НЕЛИНЕЙНЫЕ СИСТЕМЫ С ПЕРЕМЕННЫМИ ЗАПАЗДЫВАНИЯМИ

Для нелинейных динамических систем со многими переменными во времени запаздываниями решена задача терминального управления методом конечного состояния.

ВВЕДЕНИЕ

Системы с запаздываниями являются достаточно распространенным видом управляемых систем, особенно при управлении на большом расстоянии либо при использовании сложной обработки измерительной информации. Имеются и другие области, где адекватными математическими моделями являются именно системы с запаздываниями. В арсенале современной теории автоматического управления имеется ряд подходов и методов, позволяющих управлять системами данного класса, но практически все они либо накладывают большие ограничения на модель объекта управления и параметры запаздывания, либо являются узко направленными

и предназначены для решения конкретной задачи. Одно из направлений – использование методов управления и анализа для систем без запаздываний. Здесь можно выделить два подхода: замена звеньев модели, содержащих запаздывание, апериодическими звеньями с подходящей постоянной времени либо игнорирование малых запаздываний. В работах [1–3] рассматриваются подходы, при которых строятся специальные прогнозирующие устройства или наблюдатели, но при этом классы рассматриваемых систем имеют ряд существенных ограничений.

Среди достаточно общих методов, пригодных для синтеза нелинейных систем, выделим линеаризационный метод А. Исидори (linearization I/O design method). Его обобщение на системы с одним постоянным запаздыванием и скалярным управлением предложено в [4], где, наряду с теоретической частью, приведен пример задачи терминального управления производи-

© Шушляпин Е. А., Виноградов Е. В., 2007

тельностью химического реактора с рециркуляцией вещества.

Предлагаемое в настоящей статье исследование посвящено развитию альтернативного метода управления нелинейными системами с запаздываниями, основанного на так называемом «методе конечного состояния».

ПОСТАНОВКА ЗАДАЧИ

В ряде работ нами предложен и развит метод управления, названный «методом конечного состояния», который предназначен для решения непрерывных, дискретных и смешанных терминальных задач для нелинейных систем (см., например, [5–6]). Метод основан на представлении систем так называемыми «моделями конечного состояния», где в качестве переменных (переменных конечного состояния, сокращенно ПКС) используются зависящие от текущего момента времени и состояния в этот момент конечные состояния неуправляемой системы.

Рассмотрим нелинейную дифференциальную систему с запаздываниями

$$\begin{aligned} \frac{dx(t)}{dt} &= \Phi(t, X(t)) + B(t) \cdot u(t), \\ X(t) &= \text{colon}(x^0(t), x^1(t), \dots, x^N(t)), \\ x^m(t) &= x(t - h_m(t)), \quad m = 0, \dots, N, \quad h^0(t) \equiv 0, \\ t &\in [t_0, t_f], \quad x(t_0) = x(L_{t_0}), \quad L_{t_0} = (-\infty, t_0], \end{aligned} \quad (1)$$

где $u(t) - (r \times 1)$ – вектор управляющих воздействий; h_m – неотрицательные непрерывные и дифференцируемые функции времени; $x(L_{t_0})$ – n -мерный вектор начальных условий, заданный на полубесконечном интервале. Переменные конечного состояния $\bar{x}(t_f, \tau, x(\tau))$ как функции первого аргумента, а также формула связи между конечным состоянием, начальными условиями и внешними воздействиями $u(t)$, полученные в работе [7] на основе формулы В. М. Алексеева [8] и приведенных в одном из отчетов о НИР результатах Ю. И. Томина (для линейных нестационарных систем), определяются выражениями

$$\begin{aligned} x(t_f) &= \bar{x}(t_f, \tau, x(\tau)) + \int_{\tau}^{t_f} W(t_f, t, x(L_{t_f})) B(t) u(t) dt, \\ \frac{d\bar{x}(t, \tau, x(L_{t_f}))}{dt} &= \Phi(t, \bar{X}(t, \tau, x(L_{t_f}))), \end{aligned} \quad (2)$$

$$\begin{aligned} \frac{dW(t, \tau, x(L_{t_f}))}{dt} &= \sum_{m=0}^N A_m(t, \bar{X}(t, \tau, x(L_{t_f}))) \cdot W_m(t, \tau, x(L_{t_f})), \\ A_m(t, \bar{X}(t, \tau, x(L_{t_f}))) &= \left. \frac{\partial \Phi(t, X)}{\partial x^m} \right|_{X = \bar{X}(t, \tau, x(L_{t_f}))}, \end{aligned}$$

$$\begin{aligned} W_m(t, \tau, x(L_{t_f})) &\equiv W(t - h_m(t), \tau), \quad t \in [\tau, t_f], \\ W(\tau, \tau, x(L_{t_f})) &= I, \quad W(t, \tau, x(L_{t_f}))|_{t < \tau} = 0, \\ \bar{x}(\tau, \tau, x(L_{t_f})) &= x(L_{t_f}), \end{aligned} \quad (3)$$

где I – единичная матрица.

Для целей управления важны, следующие из (2), уравнения, определяющие ПКС как функцию второго аргумента

$$\frac{d\bar{x}(t_f, \tau, x(L_{t_f}))}{d\tau} = W(t_f, \tau, x(L_{t_f})) B(\tau) u(\tau), \quad (4)$$

которые мы и называем моделью конечного состояния. Использование в наименованиях ПКС и модели (4) понятия «конечного состояния» обусловлено тем, что $\bar{x}(t_f, \tau, x(L_{t_f}))$, как это следует из (2), есть конечное состояние неуправляемой на участке $[\tau, t_f]$ системы, возбужденной в момент τ начальным условием в виде текущего состояния $x(L_{t_f})$.

Использование моделей конечного состояния применительно к задачам терминального [9] и оптимального [10] управления оказалось достаточно плодотворным, что позволило получить ряд новых алгоритмов.

Для метода конечного состояния центральную роль играют соотношения вида (2) вместе с определениями ПКС как функций второго аргумента, т. е. моделями конечного состояния. Заметим, что соотношения вида (2) для линейных систем разного рода широко известны и используются в различных задачах (преимущественно как средство вывода математических соотношений). Так, для линейных дифференциальных систем соответствующая формула носит название «формула Коши» (Коши – Лагранжа, Грина). Известны подобные формулы для линейных непрерывно-дискретных систем, систем с постоянными запаздываниями, интегро-дифференциальных систем [11, 12].

Настоящая работа посвящена решению терминальной задачи вида

$$\begin{aligned} J &= J(x(t_f)) \rightarrow J^*, \\ \frac{dx(t)}{dt} &= \Phi(t, X(t)) + B(t) \cdot u(t), \\ X(t) &= \text{colon}(x^0(t), x^1(t), \dots, x^N(t)), \\ x^m(t) &= x(t - h_m(t)), \quad m = 0, \dots, N, \quad h^0(t) \equiv 0, \\ t &\in [t_0, t_f], \quad x(t_0) = x(L_{t_0}), \quad L_{t_0} = (-\infty, t_0], \end{aligned} \quad (5)$$

где J – целевая функция критерия, зависящая от вектора состояния, J^* – желаемое значение целевой функции.

ОБОБЩЕНИЕ МЕТОДА КОНЕЧНОГО СОСТОЯНИЯ НА НЕЛИНЕЙНЫЕ СИСТЕМЫ С ПЕРЕМЕННЫМИ ЗАПАЗДЫВАНИЯМИ

Рассмотрим терминальную задачу (5) и определим управление на основании обобщения формулы Алексева (2). Для этого получим дифференциальное уравнение для переменной конечного состояния как функции второго временного аргумента. С этой целью продифференцируем (2) по аргументу τ . Поскольку левая часть от τ не зависит, слева будет нуль. Справа же, применяя правило дифференцирования по нижнему пределу определенного интеграла, получим

$$0 = \frac{d\bar{x}(t_f, \tau, x(L_\tau))}{d\tau} - W(t_f, \tau, x(L_\tau))B(\tau)u(\tau).$$

Переставив члены и заменив τ на t , получаем

$$\frac{d\bar{x}(t_f, t, x(L_t))}{dt} = W(t_f, t, x(L_t))B(t)u(t). \quad (6)$$

Систему уравнений относительно переменных конечного состояния как функций второго аргумента можно трактовать как модель системы управления – «модель конечного состояния». Ее особенность – отсутствие ПКС в правой части (6). В то же время терминальная задача (5) может быть сформулирована и через ПКС, так как $J(x(t_f)) = J(\bar{x}(t_f, t_f, x(L_{t_f})))$.

Таким образом, в новых переменных эквивалентная (5) задача имеет вид

$$\begin{aligned} J &= J(\bar{x}(t_f, t_f, x(L_{t_f}))) \rightarrow J^*, \\ \frac{d\bar{x}(t_f, t, x(L_t))}{dt} &= W(t_f, t, x(L_t))B(t)u(t), \\ t \in [t_0, t_f], \quad \bar{x}(t_0, t_0, x(L_{t_0})) &= x(t_0) = x^0. \end{aligned} \quad (7)$$

Для получения управления определим вначале так называемую критериальную функцию времени t .

$$\bar{J}(t) = J(\bar{x}(t_f, t, x(L_t))).$$

Продифференцируем критериальную функцию по времени в силу уравнений (7). Тогда

$$\begin{aligned} \frac{d\bar{J}(t)}{dt} &= \frac{\partial J(\bar{x})}{\partial \bar{x}} \cdot \frac{d\bar{x}(t_f, t, x(L_t))}{dt} = \\ &= \frac{\partial J(\bar{x})}{\partial \bar{x}} W(t_f, t, x(L_t))B(t)u(t), \end{aligned} \quad (8)$$

где $\frac{\partial J(\bar{x})}{\partial \bar{x}}$ – градиент целевой функции (вектор-строка) критерия с подставленными вместо аргумента $x(t)$ ПКС $\bar{x}(t_f, t, x(t))$. Зададим желаемое поведение критериальной функции (КФ) во времени, но не в явном

виде, а посредством дифференциального уравнения с желаемым решением

$$\frac{d\bar{J}(t)}{dt} = f_J(\bar{J}(t)). \quad (9)$$

Например, задав $f_J(\bar{J}(t)) = \frac{J^* - \bar{J}(t)}{T_u}$, получим экспоненциальный переход КФ от некоторого значения, определяемого начальными условиями, к желаемому значению J^* с постоянной времени T_u . При переменном во времени $T_u = t_f - t$ критериальная функция зависит от времени. Приравняв правые части (8) и (9), получаем

$$\left. \frac{\partial J(\bar{x})}{\partial \bar{x}} \right|_{\bar{x} = \bar{x}(t_f, t, x(L_t))} W(t_f, t, x(L_t))B(t)u(t) = f_J(t).$$

Вводя обозначение

$$G = \left. \frac{\partial J(\bar{x})}{\partial \bar{x}} \right|_{\bar{x} = \bar{x}(t_f, t, x(L_t))} W(t_f, t, x(L_t))B(t)$$

и используя операцию псевдообращения для матрицы G , можно получить выражение для управляющего воздействия

$$u(t) = \frac{G^T}{GG^T} f_J(\bar{J}(t)). \quad (10)$$

Следует отметить, что алгоритм применим и в случае, когда матрица коэффициентов при управлении B зависит от вектора состояния системы, т. е. $B \equiv B(t, x(t))$. Алгоритм (10) реализован в виде расширенной версии специализированного пакета для расчета терминального управления нелинейными системами методом конечного состояния и проверен на нескольких моделях систем с запаздываниями. Одной из таких моделей являлась модель химического реактора с рециркуляцией [4], представленной системой дифференциальных уравнений третьего порядка с одним постоянным запаздыванием. Результаты моделирования подтвердили работоспособность метода как при управлении каждой из координат, так и при управлении тремя координатами с использованием квадратичного критерия вида

$$\begin{aligned} J(x(t_f)) &= F_1(x_1(t_f) - x_1^*) + F_2(x_2(t_f) - x_2^*) + \\ &+ F_3(x_3(t_f) - x_3^*) \rightarrow J^*, \end{aligned}$$

где x_1^*, x_2^*, x_3^* – желаемые конечные значения координат, J^* – желаемое достижимое значение терминально-

го критерия (в идеале $J^* = 0$). Соответствующие результаты и их сопоставление с результатами применения линеаризационного алгоритма будут опубликованы в отдельной работе.

ЗАКЛЮЧЕНИЕ

Рассмотрена задача терминального управления нелинейными системами с запаздываниями вида (1). Опираясь на методику, изложенную в [5, 6] и результаты, полученные в работе [7], предложено обобщение метода конечного состояния на нелинейные системы с запаздываниями. Получено выражение (10) для определения управляющего воздействия, приводящего нелинейную систему с запаздываниями в заданное конечное состояние за заданное время.

ПЕРЕЧЕНЬ ССЫЛОК

1. Narendra K. S. Annaswamy A. M. Stable adaptive systems / Englewood Cliffs, Nj: Prentice Hall, 1989. – 494 с.
2. Гурецкий Х. Анализ и синтез систем управления с запаздыванием. М.: Машиностроение, 1973. – 328 с.
3. Терновая Г. Н. Робастно-адаптивное управление нелинейным объектом с запаздыванием по состоянию с эталонной моделью // Труды VI Международной конференции SICPRO'07. М.: ИПУ им. В. А. Трапезникова РАН, 2007. – С. 587–594.
4. Wu Wei. Lyapunov-based design procedures for a state-delay chemical process // 14th World Congress of IFAC, 1999. – N-7A-05.
5. Шушляпин Е. А., Шушляпина А. Е. Управление нелинейными дифференциальными системами на основе

- идентифицированных моделей конечного состояния // Труды III Международной конференции SICPRO'04. М.: ИПУ им. В. А. Трапезникова РАН, 2004. – С. 607–636.
6. Шушляпин Е. А., Подольская О. Г. Управление нелинейными непрерывно-дискретными системами методом конечного состояния // Труды IV Международной конференции SICPRO'05. М.: ИПУ им. В. А. Трапезникова РАН, 2005. – С. 1495–1513.
 7. Шушляпин Е. А., Виноградов Е. В. Модель конечного состояния для представления нелинейных систем с переменными запаздываниями // Труды V Международной конференции SICPRO'06. М.: ИПУ им. В. А. Трапезникова РАН, 2005. – С. 1376–1386.
 8. Алексеев В. М. Об одной оценке возмущений обыкновенных дифференциальных уравнений // Вестн. Москов. ун-та. Сер. 1. Математика, механика. – 1961. – № 2. – С. 28 – 36.
 9. Shushlyapin E. A. An Approach to the Design of Nonlinear Control System // Journal of Mathematics Science. New York:Plenum, 2001. – V. 103, № 1. – P.34–37.
 10. Shushlyapin E. A. Alternative Form of Optimal Control for the Linear-Quadratic Problem // Journal of Mathematics Science. New York:Plenum, 2001. – V. 103, № 2. – P. 154–157.
 11. Барабанов А. Т., Агранович Г. А. Линейные модели в оптимизации непрерывно-дискретных динамических систем // Динамические системы. – Киев: Вища школа, 1983. Выпуск 2. – С. 17–24.
 12. Методы исследования нелинейных систем автоматического управления / Под ред. Нелепина Р. А. – М.: Наука, 1975. – 448 с.

Надійшла 11.01.07

Для нелінійних динамічних систем з кількома змінними у часі запізнюваннями вирішена задача термінального керування методом кінцевого стану.

A problem of terminal control was solved with the method of the terminal state for nonlinear dynamic systems with many time-variable delays.

УДК 62-50

И. В. Щербань, В. А. Бертенев

НЕПАРАМЕТРИЧЕСКАЯ СТОХАСТИЧЕСКИХ ИДЕНТИФИКАЦИЯ НЕЛИНЕЙНЫХ СИСТЕМ НА ОСНОВЕ КВАДРАТИЧНОГО КРИТЕРИЯ

В статье рассматривается теоретическое решение задачи непараметрической идентификации стохастических систем. Найдено эффективное с вычислительной точки зрения приближенное решение, обеспечивающее возможность идентификации в реальном масштабе времени. Приведен пример, иллюстрирующий возможность практического применения предложенного подхода. Полученные результаты позволяют сделать вывод о высокой точности рассмотренной процедуры непараметрической идентификации и возможности ее практического использования в реальных системах.

ВВЕДЕНИЕ

До настоящего времени задача идентификации нелинейных динамических систем удовлетворительно решалась для различных случаев параметрической идентификации [1, 2]. В статье рассматривается проблема идентификации в целом всей динамической системы, описываемой нелинейным векторным стохастическим уравнением

$$\dot{x} = f(x, t) + W_t, \quad (1.1)$$

где x – вектор состояния; $f(x, t)$ – неизвестная (идентифицируемая) вектор-функция; W_t – белый гауссовский вектор-шум (БГШ) с нулевым средним и известной матрицей интенсивности $D_W(t)$ с измерением на заданном конечном интервале времени $T = [t_0, t_k]$ зашумленным измерителем

$$z = h(x, t) + V_t, \quad (1.2)$$

где z – вектор выходных сигналов измерителя; $h(x, t)$ – нелинейная вектор функция наблюдения; V_t – БГШ с нулевым средним и матрицей интенсивности $D_V(t)$. Указанная проблема в настоящее время остается нерешенной.

В [3] решена задача, но для локальной непараметрической идентификации. Это ограничивает возможность практического использования такого подхода при необходимости идентификации глобальных в смысле времени критериев оценивания на всем промежутке функционирования стохастической системы $T \in [t_0, t_k]$.

1 ФОРМУЛИРОВКА ИСХОДНЫХ ПОЛОЖЕНИЙ

Для решения данной проблемы выберем в качестве критерия, обеспечивающего максимальную точность идентификации, классический метод наименьших квадратов (МНК), одновременно потребовав, в силу принципа Ферма, минимума квадратичной формы функции f на временном интервале T (т. е. минимальной энергии системы (1.1)). Минимизируемый критерий J в этом случае принимает вид

$$J = \int_T \{ [z - h(x, t)]^T [z - h(x, t)] + f^T(x, t) f(x, t) \} dt. \quad (2.1)$$

Так как идентификация (формирование функции f) должна быть осуществлена на основании совокупности наблюдений z , сделанных на интервале T , то естественно рассматривать синтез вектора f в функции не самого стохастического вектора x , а вектора его оценки \hat{x} , полученной на основании одного из существующих методов фильтрации [1, 2, 3]. В качестве такого фильтра используем далее наиболее распространенный нелинейный гауссовский фильтр (как будет показано ниже, вид фильтра не оказывает принципиального влияния на существо рассматриваемого метода), уравнения которого для системы «объект-наблюдатель» (1.2) имеют вид

$$\begin{aligned} \dot{\hat{x}} &= f(\hat{x}, t) + K(\hat{x}, t)[z - h(\hat{x}, t)], \quad K = R \left(\frac{\partial h}{\partial x} \right) D_V^{-1}, \\ \dot{R} &= R \left(\frac{\partial f}{\partial x} \right)^T + \frac{\partial f}{\partial x} R + D_W - K D_V K^T, \end{aligned}$$

$$\hat{x}_0 = M(x_0), \quad R_0 = M[(x_0 - \hat{x}_0)(x_0 - \hat{x}_0)^T]. \quad (2.2)$$

Так как для последующего решения задачи в уравнениях (2.2) основной интерес представляют составляющие, содержащие вектор f , то запишем данные уравнения в более компактной форме

$$\begin{aligned} \dot{\hat{x}} &= f(\hat{x}, t) + G_x(\hat{x}, R, t), \\ \dot{R} &= R \left(\frac{\partial f}{\partial x} \right)^T + \frac{\partial f}{\partial x} R + G_R(\hat{x}, R, t), \end{aligned} \quad (2.3)$$

дополнительно отметив, что к форме (2.3) сводятся уравнения практически всех известных на сегодняшний день фильтров [1, 2, 3].

Таким образом, поставленная задача может быть сформулирована как задача поиска вектор-функции f , доставляющей минимум критерию J и определенной на множестве решений системы (2.3).

2 РЕШЕНИЕ ЗАДАЧИ

Для возможности использования в процессе решения аппарата современной теории оптимального управления необходимо предварительно трансформировать систему (2.3), где R – матрица, в единую векторную систему относительно расширенного вектора $\gamma = \begin{bmatrix} \hat{x} \\ R^{(v)} \end{bmatrix}$,

где (v) – операция преобразования матрицы в вектор, введенная ранее в (2.2). При этом сделаем единственное упрощающее допущение относительно вектора f , считая, что все его компоненты могут быть представлены конечным разложением в ряд по некоторой заданной системе многомерных функций (ортогональных, степенных и т. д.) $\{\psi_1, \dots, \psi_s\}$. Обозначив вектор-строку $[\psi_1, \dots, \psi_s] = \psi$, представим далее аппроксимацию вектора f в виде

$$f(\hat{x}, t) = (E \otimes \psi^T) f = \psi_E f, \quad (3.1)$$

где $f = [f_{11} \dots f_{1s} \quad f_{21} \dots f_{2s} \quad f_{N1} \dots f_{Ns}]^T$ – вектор коэффициентов разложения, $f_i(\hat{x}, t) = \sum_{j=1}^s f_{ij}(t) \psi_j(x)$ – i -й компонент искомого вектора f .

Тогда система уравнений (2.3) может быть трансформирована в следующую:

$$\begin{aligned} \dot{\hat{x}} &= \psi_E f + G_x, \\ \dot{R}^{(v)} &= (E \otimes R) \left(\frac{\partial \psi_E}{\partial \hat{x}} \right)^* f + (R \otimes E) \frac{\partial \psi_E}{\partial \hat{x}} f + G_R^{(v)}. \end{aligned} \quad (3.2)$$

При использовании операций, ранее введенных в [4, 5],

$$\left(\frac{\partial \Psi_E}{\partial \hat{x}} f\right)^{(v)} = \frac{\partial \hat{\Psi}_E}{\partial \hat{x}} f; \left[f^T \left(\frac{\partial \Psi_E}{\partial \hat{x}}\right)^T \right]^{(v)} = \left(\frac{\partial \Psi_E}{\partial \hat{x}}\right)^* f; \quad (ASC)^{(v)} = (C^T \otimes A)S^{(v)}, \quad (3.3)$$

где A, C, S – произвольные прямоугольные матрицы; символ « v » означает преобразование блочной матрицы

$$B = |B_1; B_2; \dots; B_N| \text{ в блочную матрицу } \begin{bmatrix} B_1 \\ B_2 \\ \vdots \\ B_N \end{bmatrix}; \text{ символ}$$

«*», соответственно, матрицы B^T в матрицу

$$\begin{bmatrix} \left[(B_{(1)}^T)^{(\wedge)} \right]^T \\ \left[(B_{(2)}^T)^{(\wedge)} \right]^T \\ \vdots \\ \left[(B_{(N)}^T)^{(\wedge)} \right]^T \end{bmatrix}, \text{ где } B_{(i)}^T - i\text{-й столбец } B^T; (\wedge) -$$

операция, обратная операции (v) ; $[R^{(v)}]^{(\wedge)} = R$; \otimes – символ кронекеровского произведения.

В окончательной (канонической) форме система уравнений расширенного вектора y может быть представлена следующим образом:

$$\begin{aligned} \begin{bmatrix} \dot{\hat{x}} \\ \dot{R}^{(v)} \end{bmatrix} &= \dot{y} = \\ &= \begin{bmatrix} \Psi_E \\ |E \otimes R| \left(\frac{\partial \Psi_E}{\partial \hat{x}}\right) + (R \otimes E) \frac{\partial \hat{\Psi}_E}{\partial \hat{x}} \end{bmatrix} f + \begin{bmatrix} G_x \\ G_R^{(v)} \end{bmatrix}, \quad (3.4) \end{aligned}$$

или

$$\dot{y} = B(y, t)f + B_0(y, t). \quad (3.5)$$

При подобном представлении вектора $f(\hat{x}, t)$ выражение критерия J также преобразуется к следующему виду:

$$J = \int_T \left\{ (z-h)^T (z-h) + f^T \Psi_E^T \Psi_E f \right\} dt \quad (3.6)$$

и, следовательно, поставленная задача идентификации далее рассматривается уже как задача поиска вектора f , минимизирующего критерий (3.6) на решениях системы (3.5), и решается с использованием существующих методов теории оптимального управления. Для последующего синтеза вектора f сформируем гамильтониан H для системы (3.5) в виде

$$H = (z-h)^T (z-h) + f^T \Psi_E^T \Psi_E f + \lambda^T (Bf + B_0), \quad (3.7)$$

где λ – вектор сопряженных переменных. Из необходимых условий наличия экстремума последнего вытекает уравнение для оптимального вектора f

$$2f^T \Psi_E^T \Psi_E + \lambda^T B = 0. \quad (3.8)$$

Из последнего уравнения определяется искомым вектор

$$f_{\text{opt}} = -\frac{1}{2} (\Psi_E^T \Psi_E)^{-1} B^T \lambda. \quad (3.9)$$

Подстановка выражения (3.9) в систему канонических сопряженных уравнений

$$\begin{aligned} \dot{y} &= -\frac{1}{2} B (\Psi_E^T \Psi_E)^{-1} B^T \lambda + B_0; \\ \dot{\lambda} &= 2 \left| \frac{\partial H}{\partial \hat{x}} \right| \mathbf{0} \left| (z-h) - \frac{1}{2} \lambda^T B (\Psi_E^T \Psi_E)^{-1} \left| \frac{\partial \Psi_E}{\partial \hat{x}} \right| \mathbf{0} \right|^T \times \\ &\times \Psi_E^T B^T \lambda - \left[\left(\frac{\partial B_0}{\partial y} \right)^T - \frac{1}{2} \lambda^T B (\Psi_E^T \Psi_E)^{-1} \left(\frac{\partial B}{\partial y} \right)^T \right] \lambda; \\ y(t_0) &= y_0; \lambda(t_k) = 0 \quad (3.10) \end{aligned}$$

после интегрирования последней и формирования по результатам решения вектора f_{opt} (3.9) завершает, по существу, теоретическое решение поставленной задачи.

3 АНАЛИЗ ПУТЕЙ ВЫЧИСЛИТЕЛЬНОЙ РЕАЛИЗАЦИИ ПРОЦЕДУРЫ ИДЕНТИФИКАЦИИ

Интегрирование системы (3.10), представляющее собой решение двухточечной краевой задачи (ДТКЗ), может быть осуществлено различными известными методами: пристрелки, квазилинеаризации, инвариантного погружения и т. д. [2, 6]. С точки зрения вычислительных затрат, как показано в многочисленных исследованиях [2, 6], одним из наиболее эффективных является приближенный метод инвариантного погружения [6], позволяющий получить приближенное решение ДТКЗ в реальном масштабе времени уже для существующих классов вычислителей. Его применение

к системе (3.10) приводит к следующему уравнению приближенной оценки \hat{y} вектора y :

$$\begin{aligned} \dot{\hat{y}} &= 2D \left[\frac{\partial h}{\partial x} \vdots 0 \right]^T (z-h) + B_0, \quad \hat{y}(t_0) = y_0, \\ \dot{D} &= 2 \frac{\partial B_0}{\partial \hat{y}} D + D \left(\frac{\partial B_0}{\partial \hat{y}} \right)^T + \frac{1}{2} B (\Psi_E^T \Psi_E)^{-1} B^T + \\ &+ 4D \left(\left[\frac{\partial^2 h}{\partial \hat{x}^2} \vdots 0 \right]^T (z-h) + \left[\frac{\partial h}{\partial \hat{x}} \vdots 0 \right]^T \frac{\partial h}{\partial \hat{x}} \right), \end{aligned} \quad (4.1)$$

интегрирование которого может быть обеспечено современными вычислительными машинами в реальном масштабе времени.

Следует при этом отметить, что подобный подход позволяет, в отличие от остальных известных, сформировать в явной форме приближенную оценку \hat{f}_{opt} самого вектора f_{opt} в текущем времени. Действительно, сравнивая правые части уравнений для \hat{y} (4.1) и y (3.5) (в предположении соответствующего их совпадения, удовлетворяющего требованиям практической точности идентификации), имеем

$$B(\hat{y}, t) \hat{f}_{opt} = 2D \left[\frac{\partial h}{\partial x} \vdots 0 \right]^T (z-h), \quad (4.2)$$

откуда

$$\hat{f}_{opt} = 2B^{-1} D \left[\frac{\partial h}{\partial x} \vdots 0 \right]^T (z-h). \quad (4.3)$$

Формирование подобного явного выражения \hat{f}_{opt} может оказаться необходимым, например, при последующем решении задачи синтеза управления объектом (1.1).

Для иллюстрации возможности практического использования предложенного подхода рассмотрим следующий пример.

4 ПРИМЕР

Точность идентификации оценивалась для линейного объекта

$$\dot{x} = -x^3 + a \sin t, \quad x_0 = 0, \quad a = 10^{-2}, \quad (5.1)$$

наблюдаемого квадратичным измерителем

$$z = x^2 + V \quad (5.2)$$

на временном интервале $[0,600]$ с, где V – БГШ с БГШ с нулевым средним и $D_V = 10^{-1}$.

В качестве функций разложения идентифицируемой правой части в ряд была выбрана система степенных

функций до 4-го члена разложения $\psi = |1 \ x \ x^2 \ x^3|^T$, что определило следующий вид функций B и B_0 :

$$B = \begin{bmatrix} \psi^T \\ 2R|0 \ 1 \ 2x \ 3x^2| \end{bmatrix}; \quad B_0 = \begin{bmatrix} 20 \ R\hat{x} \\ -40 \ R^2\hat{x}^2 \end{bmatrix}. \quad (5.3)$$

Использованная в процессе идентификации система уравнений, построенная аналогично уравнениям (4.1)

$$\begin{aligned} \dot{\hat{y}} &= 4D \cdot \begin{bmatrix} \hat{x} \\ 0 \end{bmatrix} \cdot (z-\hat{x})^2 + 20 \cdot R\hat{x} \cdot \begin{bmatrix} 1 \\ -2R\hat{x} \end{bmatrix}; \\ \dot{D} &= 40 \begin{bmatrix} R & \hat{x} \\ -4R^2\hat{x} & -4R\hat{x}^2 \end{bmatrix} D + 20D \begin{bmatrix} R & -4R^2\hat{x} \\ \hat{x} & -4R\hat{x}^2 \end{bmatrix} + \\ &+ \frac{1}{2} B (\Psi^T \Psi)^{-1} B^T + 8D \begin{bmatrix} z + \hat{x}^2 \\ 0 \end{bmatrix}, \end{aligned} \quad (5.4)$$

интегрировалась методом Рунге-Кутты 4-го порядка. По окончании моделирования оценка точности идентификации производилась путем вычисления ошибки между реальной координатой объекта x и ее оценкой \hat{x} с последующим усреднением ошибки на интервале $[500,600]$ с. В результате численного моделирования было установлено, что средняя ошибка оценки не превышает 4 %.

ЗАКЛЮЧЕНИЕ

Полученные результаты позволяют сделать вывод о высокой точности рассмотренной процедуры непараметрической идентификации и возможности ее практического использования в реальных системах управления, связи и наблюдения – радионавигационных, телеметрических, астрофизических и т. д.

ПЕРЕЧЕНЬ ССЫЛОК

1. Казаков И. Е. Статистическая теория систем управления в пространстве состояний. – М.: Наука, 1975. – 432 с.
2. Справочник по теории автоматического управления / Под ред. А. А. Красовского. – М.: Наука, 1987. – 712 с.
3. Щербань О. Г. Решение задачи идентификации динамических объектов в условиях априорной неопределенности // Радиоэлектроника, информатика, управление. – 2004. – № 1. – С. 43–47.
4. Тихонов В. И., Харисов В. Н. Статистический анализ и синтез радиотехнических устройств и систем. – М.: Радио и связь, 1991. – 608 с.
5. Чернов А. А., Ястребов В. Д. Метод оценки возмущений в алгоритмах решения навигационных задач // Изв. РАН. Косм. исслед. 1984. – Т. 22, № 3. – С. 361–368.
6. Соколов С. В., Хуторцев В. В. Современные принципы управления и фильтрации в стохастических системах. – М.: Радио и связь, 2001. – 808 с.
7. Соколов С. В., Коляда Ю. И., Мельниченко Ф. В. Синтез оптимального управления наблюдениями нелинейных стохастических процессов // Автом. и выч. техника. – 1997. – № 6. – С. 5–8.

8. Первачев С. В., Перов А. И. Адаптивная фильтрация сообщений. – М.: Радио и связь. 1991. – 220 с.

тої процедури непараметричної ідентифікації і можливості її практичного застосування у реальних системах.

Надійшла 14.06.06

У статті розглядається теоретичне розв'язання задачі непараметричної ідентифікації стохастичних систем. Знайдено ефективно з обчислювальної точки зору приближене розв'язання, що забезпечує можливість ідентифікації у реальному масштабі часу. Надано приклад, що ілюструє можливість практичного впровадження запропонованого підходу. Отримані результати дозволяють зробити висновок щодо високої точності розгляну-

In the article a principal theoretical solution of the problem of the nonparametric identification of the stochastic systems is considered. An efficient (from the computing point of view) approximated solution providing the identification feasibility in the real time scale has been found. An example illustrating the feasibility to apply practically the approach suggested has been given. The results obtained allow to make a conclusion on a high accuracy of the nonparametric identification procedure considered and on the feasibility of its practical application in the real systems.

АВТОРЫ НОМЕРА

Алексеев Сергей Викторович

кандидат технических наук, старший научный сотрудник Харьковского университета Воздушных Сил

Алипов Николай Васильевич

доктор технических наук, профессор Харьковского национального университета радиоэлектроники

Артюшенко Богдан Андреевич

аспирант Запорожского национального технического университета

Афонин Юрий Сергеевич

аспирант Запорожского национального технического университета

Бертенев Владимир Александрович

адъютант Ростовского военного института ракетных войск

Ващенко Вадим Владимирович

заведующий лабораторией Запорожского национального технического университета

Вершина Александр Иванович

кандидат технических наук, доцент Запорожского национального технического университета

Ветрова Наталия Моисеевна

кандидат экономических наук, заведующая кафедрой Национальной академии природоохранного и курортного строительства

Виноградов Евгений Викторович

аспирант Севастопольского национального технического университета

Войтович Сергей Антонович

кандидат технических наук, начальник кафедры Харьковского университета Воздушных Сил

Воропай Алексей Юрьевич

ассистент Запорожского национального технического университета

Гапоненко Николай Прокофьевич

кандидат технических наук, доцент Запорожского национального технического университета

Герасин Сергей Николаевич

кандидат технических наук, доцент Харьковского национального университета радиоэлектроники

Говоров Антон Олегович

студент Запорожского национального технического университета

Горецкий Антон Сергеевич

аспирант Запорожского национального технического университета

Гусятин Максим Владимирович

инженер Харьковского национального университета радиоэлектроники

Демин Максим Константинович

аспирант Восточноукраинского национального университета им. Владимира Даля, г. Луганск

Дубровин Валерий Иванович

кандидат технических наук, доцент Запорожского национального технического университета

Емельяненко Сергей Сергеевич

аспирант Сумского государственного университета

Зинченко Руслан Николаевич

кандидат технических наук, доцент Сумского государственного университета

Ильяшенко Матвей Борисович

аспирант Запорожского национального технического университета

Калачева Вероника Валериевна

кандидат технических наук, старший научный сотрудник Харьковского военного университета Воздушных Сил

Карпуков Леонид Матвеевич

доктор технических наук, заведующий кафедрой Запорожского национального технического университета

Килимник Ирина Михайловна

кандидат технических наук, доцент Запорожского национального технического университета

Коваленко Дарья Аркадьевна

аспирант Запорожского национального технического университета

Козина Галина Леонидовна

кандидат физико-математических наук, доцент Запорожского национального технического университета

Коломеец Геннадий Павлович

кандидат физико-математических наук, профессор Гуманитарного университета «Запорожский институт государственного и муниципального управления»

Кораблев Николай Михайлович

кандидат технических наук, доцент Харьковского национального университета радиоэлектроники

Кративной Александр Викторович

кандидат физико-математических наук, доцент Запорожского национального технического университета

Крищук Владимир Николаевич

кандидат технических наук, заведующий кафедрой Запорожского национального технического университета

Левинзон Давид Иделевич

доктор технических наук, профессор Гуманитарного университета «Запорожский институт государственного и муниципального управления»

Лысенко Татьяна Владимировна

кандидат технических наук, доцент Одесского национального политехнического университета

Манило Татьяна Владимировна

студентка Запорожского национального технического университета

Михайлова Мария Валентиновна

студентка НТУУ «КПИ»

Намлинский Антон Александрович

аспирант Запорожского национального технического университета

Небеснюк Оксана Юрьевна

кандидат технических наук, доцент Запорожской государственной инженерной академии

Никонова Алина Александровна

аспирант Запорожской государственной инженерной академии

Никулищев Геннадий Игоревич

студент Запорожского национального технического университета

Огренич Евгений Викторович

аспирант Запорожского национального технического университета

Орловский Игорь Анатольевич

кандидат технических наук, докторант Запорожского национального технического университета

Романенко Игорь Александрович

кандидат военных наук, заместитель главнокомандующего Воздушных Сил Вооруженных Сил Украины

Рубан Игорь Викторович

кандидат технических наук, начальник кафедры Харьковского университета Воздушных Сил

Рубцов Антон Геннадьевич

аспирант Института вычислительного моделирования Сибирского отделения Российской академии наук, г. Красноярск

Садовский Михаил Георгиевич

доктор физико-математических наук, заведующий отделом Института вычислительного моделирования Сибирского отделения Российской академии наук, г. Красноярск

Семерюк Татьяна Николаевна

ассистент Запорожского национального технического университета

Сенашова Мария Юрьевна

кандидат физико-математических наук, старший научный сотрудник Института вычислительного моделирования Сибирского отделения Российской академии наук, г. Красноярск

Синявский Андрей Александрович

аспирант Запорожского национального технического университета

Слипченко Елена Викторовна

кандидат технических наук, старший научный сотрудник НТУ «ХПИ»

Солдатов Борис Трофимович

доцент Запорожского национального технического университета

Становский Александр Леонидович

доктор технических наук, заведующий кафедрой Одесского национального политехнического университета

Тараненко Юрий Карлович

кандидат технических наук, старший научный сотрудник Днепропетровского университета экономики и права

Фарафонов Алексей Юрьевич

старший преподаватель Запорожского национального технического университета

Хиль Михаил Иванович

кандидат технических наук, заведующий кафедрой Северодонецкого технологического института Восточноукраинского национального университета имени В. Даля

Худенко Надежда Петровна

кандидат технических наук, доцент Одесской национальной академии пищевых технологий

Худяев Александр Андреевич

кандидат технических наук, доцент Украинской инженерно-педагогической академии, г. Харьков

Чумаченко Виталий Павлович

доктор физико-математических наук, заведующий кафедрой Запорожского национального технического университета

Чумаченко Ярослав Виталиевич

ассистент Запорожского национального технического университета

Шило Галина Николаевна

кандидат технических наук, доцент Запорожского национального технического университета

Шушляпин Евгений Андреевич

доктор технических наук, профессор Севастопольского национального технического университета

Щербань Игорь Васильевич

кандидат технических наук, доцент Ростовского военного института ракетных войск

Ющенко Светлана Сергеевна

аспирант Национального авиационного университета, г. Киев

Ярымбаш Дмитрий Сергеевич

аспирант Запорожского национального технического университета

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**

№ 1 / 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 26.06.2007. Формат 60x84/8.
Папір офс. Різогр. друк. Ум. друк. арк. 18,3.
Тираж 300 прим. Зам. № 970.*

69063 м. Запоріжжя, ЗНТУ, друкарня, вул. Жуковського, 64