

Запорізький національний технічний університет



# Радіоелектроніка Інформатика Управління

**2(21)·2009**

Науковий журнал

Виходить двічі на рік

Видається з березня 1999 року

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  
2009

Постановами ВАК України №1-05/9 від 08.09.1999 р. та №3-05/7 від 30.06.2004 р. журнал «Радіоелектроніка, інформатика, управління» (скорочена назва – РІУ), який видається з 1999 року, включений 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».

Журнал є донором журналу «Telecommunications and Radio Engineering», який видається в США. Інтернет-сторінка журналу: <http://journal.zntu.edu.ua/ric/index.php?page=index>.

Статті, що публікуються в журналі, реферуються в базах даних та РЖ ВІНІТІ (Росія) і «Джерело» (Україна). Журнал РІУ міститься у міжнародній базі наукових видань Index Copernicus (<http://journals.indexcopernicus.com/index.php>), електронна копія журналу розміщена на сайті Національної бібліотеки України імені В. І. Вернадського НАН України у розділі «Наукова періодика України» за адресою: <http://nbuv.gov.ua/portal/>.

Журнал розповсюджується за Каталогом періодичних видань України (передплатний індекс – 22914).

## РЕДАКЦІЙНА КОЛЕГІЯ

**Головний редактор** – д-р техн. наук Піза Д. М.

**Заст. головного редактора** – канд. техн. наук Дубровін В. І.

**Члени редколегії :**

д-р фіз.-мат. наук Ахметшин А. М.  
д-р техн. наук Волков О. В.  
д-р фіз.-мат. наук Горбань О. М.  
д-р фіз.-мат. наук Горр Г. В.  
д-р техн. наук Гостев В. І.  
д-р фіз.-мат. наук Дробахин О. О.  
д-р техн. наук Карпуков Л. М.  
д-р фіз.-мат. наук Корніч Г. В.

д-р фіз.-мат. наук Матюшин В. М.  
д-р фіз.-мат. наук Онуфрієнко В. М.  
д-р фіз.-мат. наук Погосов В. В.  
д-р техн. наук Потапенко Є. М.  
д-р техн. наук Толок В. О.  
д-р техн. наук Труфанов І. Д.  
д-р фіз.-мат. наук Чумаченко В. П.

Рекомендовано до видання вченою радою Запорізького національного технічного університету, протокол № 11 від 1.06.2009 р.

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

Журнал зверстаний редакційно-видавничим відділом  
Запорізького національного технічного університету

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

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

# ЗМІСТ

## РАДІОФІЗИКА ..... 5

*С. Г. Алексин, С. Г. Лебедев, О. О. Дробахин*  
ВОССТАНОВЛЕНИЕ ПРОФИЛЯ  
ДИЭЛЕКТРИЧЕСКОЙ ПРОНИЦАЕМОСТИ  
СЛОИСТЫХ СТРУКТУР С ИСПОЛЬЗОВАНИЕМ  
МЕТОДА ГЕЛЬФАНДА – ЛЕВИТАНА –  
МАРЧЕНКО..... 5

*О. О. Дробахин, П. И. Заболотный, Е. Н. Привалов*  
РЕЗОНАНСНЫЕ СВОЙСТВА МИКРОВОЛНОВЫХ  
РЕЗОНАТОРОВ В ВИДЕ СФЕРИЧЕСКОГО  
СЕКТОРА ..... 11

*І. В. Зависляк, Г. П. Головач, М. О. Попов*  
МОДИФІКОВАНА ЕЛІПТИЧНА СИСТЕМА  
КООРДИНАТ ДЛЯ ЗАДАЧ МАТЕМАТИЧНОЇ  
ФІЗИКИ З СИМЕТРІЄЮ ЕЛІПТИЧНОГО  
ЦИЛІНДРА ..... 16

*V. V. Pogosov, E. V. Vasyutin, A. V. Babich*  
FEATURE OF MAGIC METAL NANOCCLUSERS  
IN MOLECULAR TRANSISTOR ..... 20

*Я. В. Чумаченко, В. П. Чумаченко*  
К ОБОСНОВАНИЮ ЧИСЛЕННОГО РЕШЕНИЯ  
ОДНОЙ ЗАДАЧИ РАССЕЯНИЯ ВОЛН  
ДЛЯ НАГРУЖЕННОГО ИЗЛОМА  
ПРЯМОУГОЛЬНОГО ВОЛНОВОДА ..... 32

## РАДІОЕЛЕКТРОНІКА ТА ТЕЛЕКОМУНІКАЦІЇ ..... 35

*А. Н. Буханько, В. М. Безрук, Е. В. Дуравкин*  
ПРИМЕНЕНИЕ E-СЕТЕЙ  
ДЛЯ МОДЕЛИРОВАНИЯ АЛГОРИТМОВ  
УПРАВЛЕНИЯ УЧАСТКОМ  
ТЕЛЕКОММУНИКАЦИОННОЙ СЕТИ..... 35

*О. М. Ибрагимов*  
ИСПОЛЬЗОВАНИЕ ЧАСТОТНОГО  
РАДИОВЫСОТОМЕРА ДЛЯ ОПРЕДЕЛЕНИЯ  
ШЕРОХОВАТОСТИ ЗЕМНОЙ ПОВЕРХНОСТИ  
КОРРЕЛЯЦИОННЫМ МЕТОДОМ..... 41

*С. В. Морщавка, Д. М. Пица, Е. Л. Белоусов*  
ИНФОРМАЦИОННАЯ ТЕХНОЛОГИЯ  
ОБРАБОТКИ ПРОПАШНЫХ КУЛЬТУР.  
РАДИОМЕТРИЧЕСКИЕ АСПЕКТЫ ..... 45

## МАТЕМАТИЧНЕ ТА КОМП'ЮТЕРНЕ МОДЕЛЮВАННЯ ..... 53

*О. О. Архипова, В. М. Журавльов*  
ЧАСТОТНИЙ АНАЛІЗ ВИКОРИСТАННЯ БУКВ  
УКРАЇНСЬКОЇ МОВИ ..... 53

*И. В. Гребенник, А. Ю. Хабаров, Ю. А. Гриценко,  
Д. В. Иванов*  
ОПТИМИЗАЦИЯ ПУТИ СОЕДИНЕНИЯ ДВУХ  
ТОЧЕК ПРОВОДНОЙ СЕТИ ЭЛЕКТРОСВЯЗИ  
ПРИ ИНТЕРВАЛЬНОЙ НЕОПРЕДЕЛЕННОСТИ  
ДАННЫХ ..... 57

*В. Н. Журавлёв, В. С. Кабак*  
ИНФОРМАЦИОННАЯ МОДЕЛЬ ИМПУЛЬСНОГО  
РЕЧЕСЛУХОВОГО ПРОЦЕССА ..... 62

*Т. В. Кіпріч, В. М. Харитонов, В. І. Дубровін,  
А. В. Притула*  
ПОБУДОВА МЕТОДІВ ТА МОДЕЛЕЙ ДЛЯ  
ПОПЕРЕДЖЕННЯ ПОМПАЖНИХ ЯВИЩ  
У ГАЗОТУРБІННОМУ ДВИГУНІ..... 66

*М. Г. Коляда*  
ИМИТАЦИОННОЕ МОДЕЛИРОВАНИЕ  
КАНАЛОВ СЕТЕЙ С ОЧЕРЕДЯМИ ПАКЕТОВ  
ЗАЩИЩЕННОЙ ИНФОРМАЦИИ..... 70

*А. А. Лисняк, С. И. Гоменюк*  
ПРИМЕНЕНИЕ R-ФУНКЦИЙ  
ДЛЯ ГЕОМЕТРИЧЕСКОГО МОДЕЛИРОВАНИЯ  
ОБЪЕКТОВ СЛОЖНОЙ ФОРМЫ..... 76

*М. Ю. Терновой*  
ПІДХІД ДО ПРЕДСТАВЛЕННЯ ІНФОРМАЦІЇ  
В ТЕРМІНАХ ПРЕДМЕТНОЇ ОБЛАСТІ ..... 81

*А. С. Шкиль, Е. Е. Сыревич, Д. Е. Кучеренко,  
Самер Альмадхоун*  
МЕТОД ОБРАТНОГО ПРОСЛЕЖИВАНИЯ  
ДЛЯ ПОИСКА ОШИБОК ПРОЕКТИРОВАНИЯ  
В HDL-КОДЕ ..... 86

---

## НЕЙРОІНФОРМАТИКА ТА ІНТЕЛЕКТУАЛЬНІ СИСТЕМИ ..... 91

*Е. В. Бодянский, В. В. Волкова, К. В. Коваль*  
АВТОМАТИЧЕСКАЯ КЛАСТЕРИЗАЦИЯ  
ТЕКСТОВЫХ ДОКУМЕНТОВ  
НА ОСНОВЕ ГЕНЕТИЧЕСКОГО АЛГОРИТМА  
С ИСКУССТВЕННЫМ ОТБОРОМ .....91

*С. Г. Удовенко, А. А. Шамраев, Г. Дибё*  
ЦИФРОВОЕ НЕЙРО-НЕЧЕТКОЕ УПРАВЛЕНИЕ  
БЛОКАМИ ЛИНЕЙНОГО СИЛЬНОТОЧНОГО  
УСКОРИТЕЛЯ ЭЛЕКТРОНОВ ..... 101

*О. О. Олійник, С. О. Субботін*  
ОПТИМІЗАЦІЯ НА ОСНОВІ КОЛЕКТИВНОГО  
ІНТЕЛЕКТУ РОЮ ЧАСТОК З КЕРУВАННЯМ  
ЗМІНОЮ ЇХНЬОЇ ШВИДКОСТІ.....96

## ПРОГРЕСИВНІ ІНФОРМАЦІЙНІ ТЕХНОЛОГІЇ..... 107

*В. П. Авраменко*  
СИСТЕМА ПОДДЕРЖКИ ПРИНЯТИЯ РЕШЕНИЙ  
ПО ЛИКВИДАЦИИ НЕШТАТНЫХ  
СИТУАЦИЙ..... 107

*І. О. Романенко, В. Ф. Столбов, В. В. Калачова*  
ШЛЯХИ ОРГАНІЗАЦІЇ КОНТРОЛЮ ЗНАНЬ  
В СИСТЕМАХ ДИСТАНЦІЙНОГО  
НАВЧАННЯ..... 127

*А. Е. Архипов*  
ЭКСПЕРТНО-АНАЛИТИЧЕСКОЕ ОЦЕНИВАНИЕ  
ИНФОРМАЦИОННЫХ РИСКОВ И УРОВНЯ  
ЭФФЕКТИВНОСТИ СИСТЕМЫ ЗАЩИТЫ  
ИНФОРМАЦИИ ..... 111

*С. Ю. Скруцкий, К. Н. Касьян*  
ПРОГРАММНЫЕ МЕТОДЫ КОДИРОВАНИЯ  
ВИДЕОПОТОКОВ, ПРИМЕНЯЕМЫЕ  
В КОМПЬЮТЕРНОЙ ГРАФИКЕ ..... 130

*В. М. Крищук, О. Ю. Малий, О. Ю. Воронай*  
УНІВЕРСАЛЬНА АЛГОРИТМІЧНА МОВА  
ПРОГРАМУВАННЯ МІКРОКОНТРОЛЕРІВ..... 115

*В. В. Литвин*  
МУЛЬТИАГЕНТНІ СИСТЕМИ ПІДТРИМКИ  
ПРИЙНЯТТЯ РІШЕНЬ, ЩО БАЗУЮТЬСЯ  
НА ПРЕЦЕДЕНТАХ ТА ВИКОРИСТОВУЮТЬ  
АДАПТИВНІ ОНТОЛОГІЇ ..... 120

## ТЕОРІЯ І МЕТОДИ АВТОМАТИЧНОГО УПРАВЛІННЯ ..... 135

*Е. М. Потапенко, А. Е. Казурова*  
КОМПЕНСАЦИЯ ВОЗДЕЙСТВИЯ  
НЕЛИНЕЙНОГО ТРЕНИЯ  
НА СИСТЕМЫ УПРАВЛЕНИЯ..... 135

## УПРАВЛІННЯ У ТЕХНІЧНИХ СИСТЕМАХ..... 141

*В. И. Гостев*  
ПРОЕКТИРОВАНИЕ НЕЧЕТКОГО РЕГУЛЯТОРА  
С ВЫХОДНЫМИ СЖАТЫМИ, ЗАВИСЯЩИМИ  
ОТ ОДНОГО ПАРАМЕТРА, S-ОБРАЗНОЙ  
И Z-ОБРАЗНОЙ ФУНКЦИЯМИ  
ПРИНАДЛЕЖНОСТИ..... 141

*И. А. Орловский, Ю. С. Бут*  
УПРАВЛЕНИЕ ВЗАИМОСВЯЗАННЫМ  
АСИНХРОННЫМ ЭЛЕКТРОПРИВОДОМ  
МОСТОВОГО КРАНА ДЛЯ ПЕРЕМЕЩЕНИЯ  
МОСТА БЕЗ ВЗАИМОДЕЙСТВИЯ РЕБОРД  
КОЛЕС С РЕЛЬСАМИ ..... 144

АВТОРЫ НОМЕРА ..... 152

---

# РАДІОФІЗИКА

# РАДИОФИЗИКА

# RADIOPHYSICS

---

УДК 621.391

С. Г. Алексин, С. Г. Лебедев, О. О. Дробахин

## ВОССТАНОВЛЕНИЕ ПРОФИЛЯ ДИЭЛЕКТРИЧЕСКОЙ ПРОНИЦАЕМОСТИ СЛОИСТЫХ СТРУКТУР С ИСПОЛЬЗОВАНИЕМ МЕТОДА ГЕЛЬФАНДА – ЛЕВИТАНА – МАРЧЕНКО

---

*Рассматриваются такие прикладные вопросы, связанные с использованием метода Гельфанда – Левитана – Марченко, как ускорение итерационной процедуры, способ получения импульсной характеристики, учет неплоского характера зондирующей волны, влияние потерь энергии в образце на результат восстановления профиля.*

### ВВЕДЕНИЕ

Восстановление одномерного профиля диэлектрической проницаемости (ПДП)  $\epsilon(x)$  является одной из общих задач радиофизики. Задача восстановления  $\epsilon(x)$  относится к обратным задачам электродинамики. Для ее решения был создан метод Гельфанда – Левитана – Марченко (ГЛМ) [1]. Особенностью этого метода является то, что исходными данными для восстановления ПДП для него является не частотная зависимость коэффициента отражения (ЧЗКО), а импульсная характеристика отражения (ИХО) исследуемой структуры, которая связана с ЧЗКО парой преобразований Фурье. Вывод метода основывается на использовании плосковолнового приближения. Метод ГЛМ позволяет получить точное решение задачи в случае, если точно известна ИХО структуры.

Импульсная характеристика предполагает наличие информации о частотной характеристике, начиная с нулевой частоты. Обеспечение локальности измерений при решении задачи неразрушающего контроля изделий из диэлектриков требует использования излучения микроволнового диапазона, что принципиально не позволяет непосредственно получить ИХО

путем преобразования Фурье данных, измеренных в частотной области, вследствие наличия частоты отсечки в волноводных трактах. Кроме того, использование рупорных излучателей не обеспечивает получение строго плоской волны. Волна имеет свойства локальной квазисферической волны с фазовым центром излучения, размещенным внутри рупора. Конечность полосы частот измерений приводит к тому, что обратное преобразование Фурье позволяет получить лишь некоторое приближение к ИХО. Для синтеза более точной ИХО в [2, 3] предложено использовать параметрические методы спектрального анализа, однако, рассмотрен случай, когда имеется в наличии ЧЗКО, начиная с частот, близких к нулевой. В некоторых ситуациях легче ИХО не измерять непосредственно, а синтезировать с использованием методов спектрального анализа на основе данных измерений ЧЗКО на некоторой сетке частот.

### ПОСТАНОВКА ЗАДАЧИ

Целью данной статьи является исследование путем численного эксперимента влияния на точность определения ПДП методом ГЛМ отклонения начальной частоты полосы измерений от нулевой, наличия расходимости излученной волны, влияния потерь в диэлектрике. Численная реализация метода ГЛМ предполагает использования более эффективного алгоритма, основанного на блочной структуре основной матрицы метода. В качестве средства параметрического

спектрального анализа целесообразно применить метод пучка матриц и рассмотреть два варианта: использование непосредственных спектральных оценок и проведение экстраполяции ЧЗКО до нулевой частоты с последующим применением дискретного преобразования Фурье (ДПФ).

## МЕТОД ГЕЛЬФАНДА – ЛЕВИТАНА – МАРЧЕНКО

### 1 «Классический» вариант итерационной процедуры

Перед началом восстановления профиля ИХО подвергается эквидистантной дискретизации, что требует решения вспомогательной оптимизационной задачи о выборе количества отсчетов дискретизированной ИХО. Затем, в итерационной процедуре последовательно обрабатываются отсчеты ИХО и выполняется последовательное, слой за слоем, восстановление ПДП в глубину. В «классическом» варианте [1, 3] итерационной процедуры на  $k$ -й итерации ( $k = 0, 1, 2, \dots$ ) необходимо решить матричное уравнение вида:

$$\begin{pmatrix} 1 & 0 & \dots & 0 & 0 \\ 0 & 1 & \dots & 0 & 0 \\ \vdots & \vdots & \ddots & \vdots & \vdots \\ 0 & 0 & \dots & 1 & 0 \\ 0 & 0 & \dots & 0 & 1 \end{pmatrix} + \begin{pmatrix} 0 & 0 & \dots & 0 & R_0 \\ 0 & 0 & \dots & R_0 & R_1 \\ \vdots & \vdots & \ddots & \vdots & \vdots \\ 0 & R_0 & \dots & R_{k-2} & R_{k-1} \\ R_0 & R_1 & \dots & R_{k-1} & R_k \end{pmatrix} \times \\ \times \begin{pmatrix} K_{-k}^k \\ K_{2-k}^k \\ \vdots \\ K_{k-2}^k \\ K_k^k \end{pmatrix} = - \begin{pmatrix} R_0 \\ R_1 \\ \vdots \\ R_{k-1} \\ R_k \end{pmatrix}, \quad (1)$$

в котором матрица коэффициентов и вектор свободных членов формируются из отсчетов ИХО. Коэффициент  $K_k^k$  (верхний индекс обозначает номер итерации, нижний – номер неизвестного коэффициента), найденный из данного матричного уравнения, позволяет определить коэффициент отражения (КО) границы  $k$ -го и  $(k+1)$ -го слоев структуры по формуле

$$r_{k,k+1} = \frac{1}{(1 + K_k^k) \prod_{i=0}^{k-1} (1 - r_{i,i+1}^2)} - 1, \quad (2)$$

в которую входят КО от всех границ, что предшествуют данной, найденные на предыдущих итерациях.

Относительная диэлектрическая проницаемость (ДП)  $(k+1)$ -го слоя структуры выражается через ДП  $k$ -го слоя и КО границы этих слоев как

$$\epsilon_{k+1} = \epsilon_k \left( \frac{1 - r_{k,k+1}}{1 + r_{k,k+1}} \right)^2. \quad (3)$$

На нулевой итерации в данную формулу входит ДП пространства перед структурой  $\epsilon_0$ .

Толщина  $(k+1)$ -го слоя вычисляется по формуле:

$$d_{k+1} = \frac{c\Delta t}{2\sqrt{\epsilon_{k+1}}}, \quad (4)$$

где  $c$  – скорость света,  $\Delta t$  – шаг дискретизации ИХО.

Вследствие того, что на  $k$ -й итерации необходимо решить систему линейных алгебраических уравнений (СЛАУ)  $(k+1)$ -го порядка, итерации становятся все более медленными. При достаточно густой дискретизации ИХО необходимое количество итераций может достигать нескольких сотен и поэтому полное восстановление ПДП занимает много машинного времени. Поскольку СЛАУ решается полностью для нахождения только одного из неизвестных коэффициентов, очевидна неэффективность использования компьютерного ресурса.

### 2 Ускоренный вариант итерационной процедуры

Особенностью матрицы коэффициентов в матричном уравнении (1) является то, что данная матрица на  $k$ -й итерации включает в себя в качестве блока матрицу коэффициентов, соответствующую итерации  $k-2$ . Это позволяет использовать метод окаймления [4] для ускорения решения СЛАУ с разбиением матрицы коэффициентов на 9 блоков.

Запишем уравнение (1) в матричном виде:

$$\mathbf{A}^k \mathbf{K}^k = -\mathbf{R}^k, \quad (5)$$

где верхним индексом обозначен номер итерации, и используем блочное представление матрицы коэффициентов  $\mathbf{A}^k$ , вектора-столбца неизвестных  $\mathbf{K}^k$  и вектора-столбца свободных членов  $\mathbf{R}^k$ :

$$\mathbf{A}^k = \begin{pmatrix} 1 & (\mathbf{0}^k)^T & R_0 \\ \mathbf{0}^k & \mathbf{A}^{k-2} & \mathbf{r}^k \\ R_0 & (\mathbf{r}^k)^T & 1 + R_k \end{pmatrix}, \quad \mathbf{K}^k = \begin{pmatrix} K_{-k}^k \\ \mathbf{k}^k \\ K_k^k \end{pmatrix}, \\ \mathbf{R}^k = \begin{pmatrix} R_0 \\ \mathbf{r}^k \\ R_k \end{pmatrix},$$

где  $\mathbf{0}^k$  – нулевой вектор-столбец высотой  $k-1$ ;  $\mathbf{A}^{k-2}$  – матрица коэффициентов на итерации  $k-2$ ,  $\mathbf{r}^k$  та  $\mathbf{k}^k$  – введенные векторы-столбцы.

Решение матричного уравнения (5) находится как:

$$\mathbf{K}^k = -\mathbf{B}^k \mathbf{R}^k, \quad (6)$$

где  $\mathbf{B}^k$  – матрица, обратная матрице  $\mathbf{A}^k$ , которая также имеет рекуррентное блочное представление. Нахождение из (6) интересующего нас коэффициента  $K_k^k$  и его подстановка в (2) приводит к следующему выражению для КО границы:

$$r_{k,k+1} = \frac{\delta^k}{\prod_{i=1}^k (1 - r_{i,i+1}^2)} - 1, \quad (7)$$

где был введен параметр  $\delta$  на  $k$ -й итерации, который дается выражением:

$$\delta^k = 1 + R_k - R_0^2 - (\mathbf{r}^k)^T \mathbf{B}^{k-2} \mathbf{r}^k. \quad (8)$$

Рекуррентное выражение для матрицы  $\mathbf{B}^{k-2}$  имеет вид:

$$\mathbf{B}^{k-2} = \begin{pmatrix} 1 & (\mathbf{0}^{k-2})^T & 0 \\ \mathbf{0}^{k-2} & \mathbf{B}^{k-4} & \mathbf{0}^{k-2} \\ 0 & (\mathbf{0}^{k-2})^T & 0 \end{pmatrix} + \frac{1}{\delta^{k-2}} \begin{pmatrix} R_0^2 & R_0 (\mathbf{r}^{k-2})^T \mathbf{B}^{k-4} & -R_0 \\ R_0 \mathbf{B}^{k-4} \mathbf{r}^{k-2} & \mathbf{B}^{k-4} \mathbf{r}^{k-2} (\mathbf{r}^{k-2})^T \mathbf{B}^{k-4} & -\mathbf{B}^{k-4} \mathbf{r}^{k-2} \\ -R_0 & -(\mathbf{r}^{k-2})^T \mathbf{B}^{k-4} & 1 \end{pmatrix}, \quad (9)$$

где  $\mathbf{0}^{k-2}$  – нулевой вектор-столбец высотой  $k-3$ .

Таким образом, на итерациях отпадает необходимость полного решения матричного уравнения (1). Ускоренный метод ГЛМ состоит из двух параллельных итерационных процедур, первая из которых включается на четных итерациях, а вторая – на нечетных. На  $k$ -й итерации необходимо вычислить параметр  $\delta^k$  по формуле (8), коэффициент отражения границы слоев  $r_{k,k+1}$  по формуле (7), найти диэлектрическую проницаемость  $\epsilon_{k+1}$  нового слоя по (3) и его толщину  $d_{k+1}$  по (4). Также на  $k$ -й итерации нужно сформировать матрицу  $\mathbf{B}^k$  по (9), которая будет использоваться на итерации  $k+2$  для вычисления параметра  $\delta^{k+2}$ .

### 3 Модификация метода ГЛМ для учета расходимости зондирующего луча

Несмотря на то, что метод ГЛМ был разработан в предположении плосковолновой модели зондирующего сигнала [1], несложная модификация данного метода делает его пригодным для восстановления и в случае, если рефлектометрические данные были получены при зондировании расходящейся волной. Данная модификация исходит из приближения пара-

ксиальной оптики, согласно которому импульсы ИХО приобретают ослабление, определяемое путем, пройденным зондирующим сигналом внутри структуры [5]. Корректировка данного ослабления и является целью предлагаемой модификации.

Предполагается, что в ИХО, являющейся исходными данными для восстановления профиля, первый импульс имеет амплитуду, равную КО от первой границы  $r_{01}$ , независимо от того, как далеко от образца находится антенна. Данное предположение основано на использовании калибровки амплитуды первого импульса на амплитуду импульса отражения от металлического листа, расположенного на том же расстоянии от антенны, что и исследуемая структура. Амплитуда каждого последующего импульса корректируется перед тем, как будет использована на соответствующей итерации метода ГЛМ. Так, для импульса  $R_k$  ( $k = 1, 2, 3, \dots$ ) корректирующее преобразование имеет вид:

$$R_k \rightarrow R_k \cdot \frac{z + \sum_{i=1}^k \frac{d_i}{\sqrt{\epsilon_i}}}{z},$$

где  $z$  – расстояние от передней грани структуры до фазового центра антенны. Данное корректирующее преобразование ИХО является тем более точным, чем меньшим контрастом обладает структура.

## СПОСОБЫ ПОЛУЧЕНИЯ ИМПУЛЬСНОЙ ХАРАКТЕРИСТИКИ

### 1 Параметрический спектральный анализ

Для синтезирования ИХО на основании данных многочастотной рефлектометрии используются различные методы параметрического спектрального анализа, такие как метод Прони [2], метод пучка матриц [3], метод, основанный на дробно-рациональной аппроксимации в спектрально-сопряженной области [6] и др. Поскольку все эти методы в общем случае дают комплексное время прихода импульсов и комплексные амплитуды, что неприемлемо для метода ГЛМ, перед восстановлением необходимо отбросить их мнимые части. Также перед восстановлением необходимо выполнить эквидистантную дискретизацию ИХО. Основным недостатком такого подхода является наличие ошибок в оценке фазы (поляриности) отражения от границ раздела, что влечет ошибку в определении ПДП.

### 2 ДПФ экстраполированных данных

Альтернативой параметрическому спектральному анализу является следующая методика, основанная на

ДПФ. Для ЧЗКО строится модель линейного предсказания, на основании которой частотная характеристика экстраполируется к нулевой частоте. Это позволяет, после преобразования Фурье во временную область, получить полярность пиков временного сигнала, соответствующую верной полярности отражений. Прежде чем характеристика будет переведена во временную область с помощью ДПФ, она должна быть обработана весовым окном и дополнена большим количеством нулевых отсчетов. Уменьшение уровня боковых лепестков импульсов во временном сигнале за счет весовой обработки ЧЗКО снимает проблему критичности выбора условного начала ИХО при восстановлении, однако загроубляет мелкие детали ПДП. В данной методике весовое окно и ДПФ должны быть использованы в своих энергосохраняющих, а не амплитудосохраняющих формах. В отличие от первого метода получения ИХО, второй метод сразу дает эквидистантную ИХО с действительным временем.

В случае, если в структуре имеются небольшие потери энергии (тангенс угла диэлектрических потерь  $\text{tg}\delta \neq 0$ ), спектральный анализ будет ассоциировать их с мнимой частью времени прихода импульсов, которую следует просто отбросить, после чего метод ГЛМ восстановит профиль действительной части  $\epsilon(x)$ . Однако при применении ДПФ такой простой способ отбрасывания потерь уже не может быть использован.

При наличии шума в данных оценки мнимых частей комплексных времен, полученных с помощью метода пучка матриц [3], имеют значительные погрешности. При наличии потерь в материале и наличии расходимости эти части имеют отрицательный знак, что, при предсказании назад к нулевой частоте данных реального эксперимента, влечет рост амплитуды отсчетов ЧЗКО вплоть до значений, превышающих единицу, что является нефизичным. Для получения верной полярности отражения при использовании экстраполированных данных достаточно для каждого значения комплексного времени  $t_k = t'_k + jt''_k$ , после обнуления  $t''_k$ , рассчитать значения  $z_k = \exp(jt'_k \Delta f)$ . Коэффициенты линейного предсказания  $p_k$  обычно вычисляются как коэффициенты полинома, корнями которого являются  $z_k$ :

$$\prod_{k=1}^M (z - z_k) = z^M + \sum_{k=1}^M p_{M-k} z^{M-k}.$$

В таком случае каждый следующий отсчет ЧЗКО  $X_n$  при предсказании назад получается как линейная комбинация  $M$  последующих:

$$X_n = - \sum_{k=1}^M \frac{p_k}{p_0} X_{n+k}.$$

## ЧИСЛЕННОЕ МОДЕЛИРОВАНИЕ

### 1 Восстановление профиля $\epsilon(x)$ при наличии потерь в материале слоев

Для численного эксперимента была выбрана структура, моделирующая два листа оргстекла с воздушным зазором. Полагалось, что оргстекло имеет  $\epsilon = 2,6$ , толщины листов оргстекла и ширина воздушного зазора равны 40 мм. ЧЗКО была синтезирована в диапазоне 10–25 ГГц с шагом сетки частот 0,1 ГГц. Для перехода во временную область были использованы две методики: параметрический спектральный анализ с использованием метода пучка матриц и экстраполяция к нулю с последующим ДПФ. Порядок спектрального анализа  $M$  в первом случае и порядок линейного предсказания во втором были выбраны равными 15, а ширина окна в методе пучка матриц  $L$  была равна 76 отсчетам. Восстановленные ПДП изображены на рис. 1, *а* и *б*. На рис. 1, *в* помещены соответствующие временные сигналы, а на рис. 1, *г* – частотная характеристика. Отметим, что импульсная характеристика, изображенная на рис. 1, *в* сплошной линией, была отнормирована так, чтобы амплитуды импульсов были равны амплитудам спектральных компонент (пунктирная линия). Однако, при восстановлении по ней профиля с помощью метода ГЛМ, необходимо использовать нормировку, сохраняющую энергию. При рассмотрении случая измерений для начальной частоты, превышающей 10 ГГц, наблюдались ошибки спектрального анализа, которые сопровождались ошибками в восстановленном ПДП.

Выигрыш во времени восстановления, который дает использование ускоренного варианта итерационной процедуры вместо «классического», при восстановлении профиля, изображенного на рис. 1, *б*, составил более двух раз.

Моделирование было повторено, но на этот раз при синтезировании ЧЗКО полагалось, что материал имеет  $\text{tg}\delta = 0,08$ . ПДП, восстановленный по результатам спектрального анализа, визуально не отличался от профиля, изображенного на рис. 1, *а*, в то время как профиль, полученный с использованием экстраполяции и ДПФ, оказался сильно искаженным (рис. 2, *а* и *б*). Попытка устранения потерь путем закладывания чисто мнимых показателей экспонент в модель линейного предсказания только ухудшила результат восстановления профиля (рис. 2, *в* и *г*).

### 2 Восстановление профиля $\epsilon(x)$ при наличии расходимости луча

Для рассмотренной выше структуры (без потерь) и частотного диапазона была синтезирована ЧЗКО с параметром  $z = 100$  мм, задающим расстояние от передней границы первого оргстекла до фазового центра волны. ЧЗКО была подвергнута спектральному



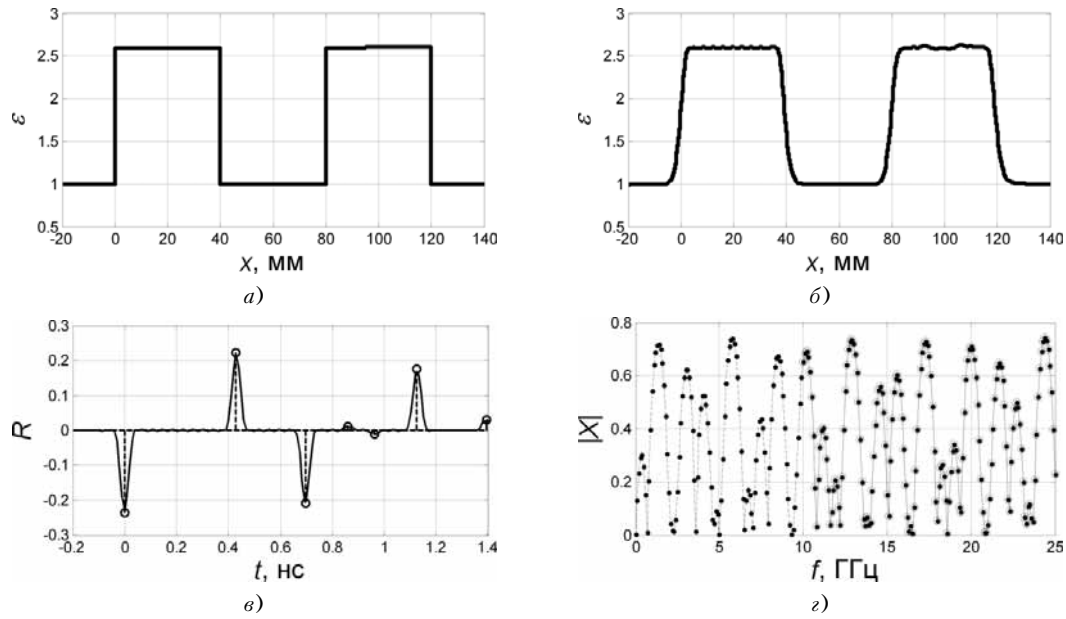


Рисунок 1 – Результат восстановления ПДП

и соответствующий сигнал при использовании двух методик для перехода во временную область:

а и пунктир на в – метод пучка матриц; б и сплошная линия на в – экстраполяция ЧЗКО к нулевой частоте с последующим ДПФ; г – исходная частотная характеристика (сплошная) и ее экстраполированная часть (пунктирная)

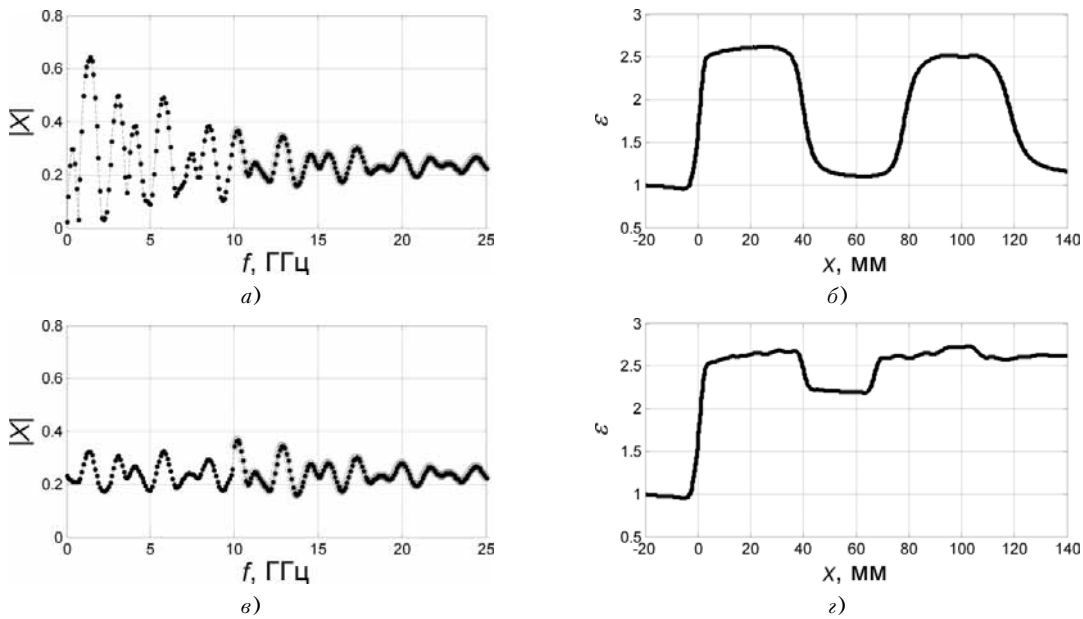


Рисунок 2 – Восстановление ПДП структуры

с потерями энергии при использовании экстраполяции ЧЗКО к нулевой частоте с последующим ДПФ:

а и б – ЧЗКО и профиль без обнуления  $t''$ ; в и г – ЧЗКО и профиль при обнулении  $t''$

анализу методом пучка матриц ( $M = 15$ ,  $L = 76$ ) и передана в программу, реализующую метод ГЛМ. На рис. 3 изображены восстановленные профили  $\epsilon$ , причем в первом случае использовался стандартный вариант метода ГЛМ, а во втором – с коррекцией расходимости. Из рисунка видно преимущество модифи-

цированного варианта итерационной процедуры. Однако, следует отметить, что модифицированный вариант метода не может корректно обрабатывать импульсы переотражений, вследствие чего в третьем слое профиля на рис. 3, б появились две небольшие ступеньки.

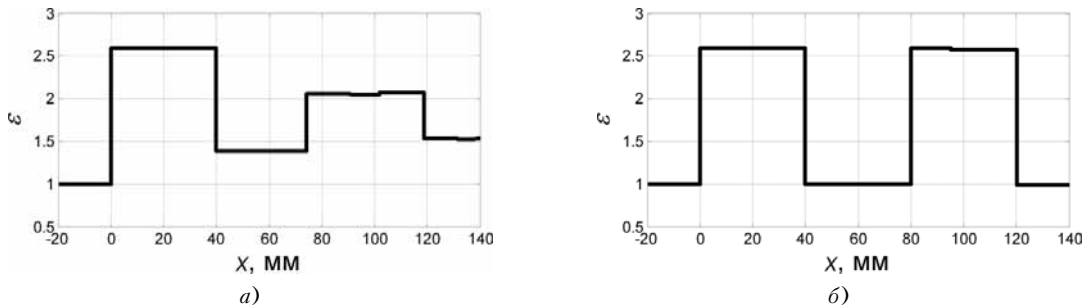


Рисунок 3 – Результат восстановления ПДП в случае расходящегося зондирующего сигнала:  
 а – при использовании стандартного метода ГЛМ; б – при использовании метода ГЛМ с коррекцией расходимости

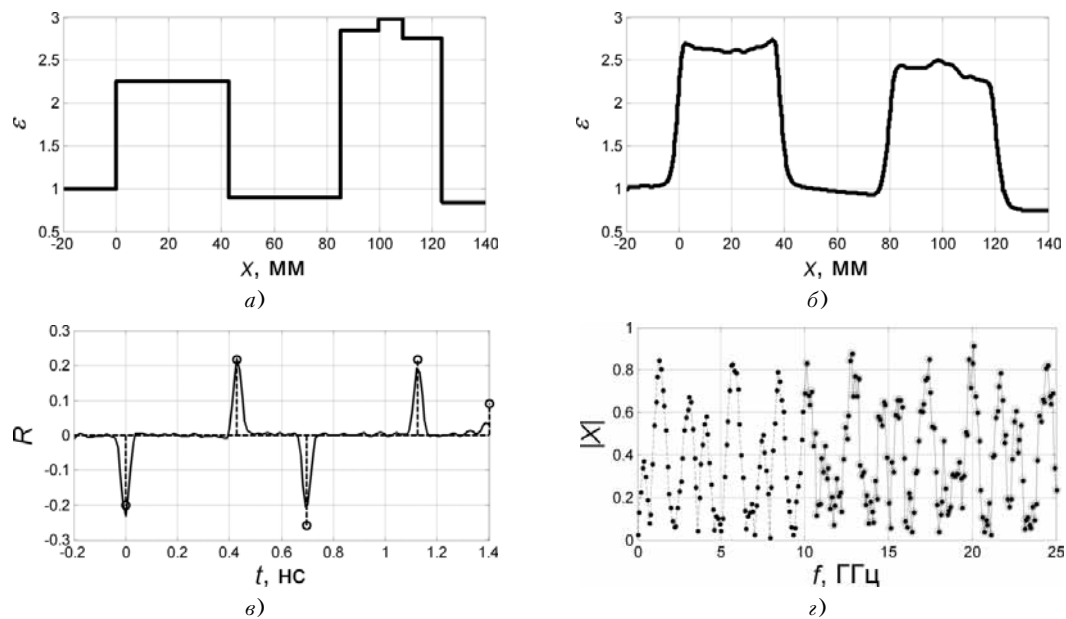


Рисунок 4 – Восстановление ПДП при наличии шума в частотных данных:  
 а и пунктир на в – метод пучка матриц; б и сплошная линия на в – экстраполяция ЧЗКО к нулевой частоте с последующим ДПФ; г – исходная частотная характеристика (сплошная) и ее экстраполированная часть (пунктирная)

### 3 Восстановление профиля $\varepsilon(x)$ при наличии шума

Для рассмотренной выше структуры (без потерь) в предположении плоской зондирующей волны была синтезирована ЧЗКО, в которую был введен аддитивный комплексный шум, имеющий нормальный закон распределения действительной и мнимой части. Действительная и мнимая части шума имели дисперсию  $\sigma^2 = 0,01$  и нулевое матожидание. ПДП, как и раньше, восстанавливался с использованием двух способов перехода во временную область, но на этот раз порядок спектрального анализа и линейного предсказания пришлось уменьшить до 6, поскольку ложные спектральные компоненты, обусловленные наличием шума, влекли получение ложных скачков в ПДП. Результаты восстановления и линейного предсказания изображены на рис. 4. Отметим, что в обо-

их случаях была использована одна и та же реализация шума.

Из рис. 4, а видно, что вследствие ошибок определения величины импульсов воздушный зазор и задний воздушный слой приобрели диэлектрическую проницаемость меньшую единицы, что нефизично. Кроме того, два импульса эха, которые должны были появиться между третьим и четвертым импульсами синтезированной ИХО, не были найдены спектральным анализом (сравните рис. 4, в и рис. 1, в), вследствие чего второй лист оргстекла приобрел ложный скачок диэлектрической проницаемости. Применение экстраполяции позволило несколько более точно определить значение  $\varepsilon$  первого и второго слоев по сравнению с непосредственным применением метода пучка матриц. Таким образом, проведение экстраполяции при обработке данных реального эксперимента

может в отдельных случаях иметь преимущества по сравнению с традиционной методикой.

## ВЫВОДЫ

1. Ускоренный вариант итерационной процедуры ГЛМ, основанный на использовании метода окаймления для решения СЛАУ, позволяет ускорить процесс восстановления профиля. Использование ускоренного варианта вместо обычного является тем более выгодным, чем большее количество итераций метода требуется для полного восстановления.

2. Предложена модификация метода ГЛМ, позволяющая скорректировать искажения в восстановленном профиле, вызванные расходимостью зондирующего луча. Выполняемая коррекция является тем более точной, чем менее контрастным является ПДП.

3. Для структур с потерями энергии показано преимущество спектрального анализа как способа получения ИХО перед способом получения, основанным на экстраполяции ЧЗКО до нулевой частоты с использованием линейного предсказания и последующим выполнением ДПФ.

4. При наличии шума метод получения ИХО на основе экстраполяции данных измерений в частотной области может давать более точные значения диэлектрической проницаемости по сравнению со случаем применения параметрического спектрального анализа.

## ПЕРЕЧЕНЬ ССЫЛОК

1. Аки К. Количественная сейсмология. Теория и методы. Т. 2 / К. Аки, П. Ричардс ; пер. с англ. А. В. Калинина. – М. : Мир, 1983. – 360 с. : ил.

УДК 621.372.22

О. О. Дробахин, П. И. Заболотный, Е. Н. Привалов

# РЕЗОНАНСНЫЕ СВОЙСТВА МИКРОВОЛНОВЫХ РЕЗОНАТОРОВ В ВИДЕ СФЕРИЧЕСКОГО СЕКТОРА

*Исследовано влияние на резонансные свойства и структуру электромагнитного поля геометрии резонаторов, выполненных в виде сферического сектора и усеченного сферического сектора. Показано, что различие между резонансными частотами колебаний типа  $H_{011}$ ,  $E_{111}$  и  $E_{112}$  в большей мере определяется углом при вершине сферического сектора. Выявлено, что в усеченных резонато-*

2. Андреев М. В. Модернизация метода Гельфанда – Левитана для решения одномерной обратной задачи с помощью метода Прони / М. В. Андреев, О. О. Дробахин, А. Г. Новомлинов, В. Г. Короткая // Системні технології. Системи і процеси обробки інформації та управління. – 2003. – Вып. 5 (28). – С. 59–63.
3. Андреев М. В. Модификация метода Гельфанда – Левитана для решения одномерной обратной задачи с помощью метода пучка матриц / М. В. Андреев, О. О. Дробахин, А. Г. Новомлинов, В. Г. Короткая, А. В. Сазонов // Радіоелектроніка, інформатика, управління. – 2002. – № 2. – С. 9–13.
4. Фортсайт Дж. Численное решение систем линейных алгебраических уравнений / Дж. Фортсайт, К. Молер ; пер. с англ. под ред. Г. И. Марчука. – М. : Мир, 1969. – 168 с.
5. Дробахин О. О. Определение зависимости амплитуды отраженного импульса в методе синтеза радиопульса / О. О. Дробахин // Дефектоскопия. – 1994. – № 4. – С. 48–55.
6. Andreev M. V. Rational Representation in Spectrum Conjugate Domain for Parameters Determination of Reflecting Structures / M. V. Andreev, V. F. Borulko, O. O. Drobakhin, D. Yu. Saltykov // 10th Int. Conf. Mathematical Methods in Electromagnetic Theory, September 14–17, 2004 : Proc. – Dnepropetrovsk. – 2004. – P. 449–451.

Надійшла 21.04.2009

*Розглядаються такі прикладні питання, пов'язані з використанням методу Гельфанда – Левітана – Марченка, як пришвидшення ітераційної процедури, спосіб отримання імпульсної характеристики, урахування неплоского характеру хвилі, що зондує, та вплив втрат енергії в зразку, що досліджується, на результат відновлення профілю.*

*Considered are applied problems related to the usage of Gelfand – Levitan – Marchenko's method as acceleration of iterative procedure, methods to obtain impulse characteristics, consideration of non-flat character of probing wave, influence of energy losses in a sample as to the result of the building-up of the profile.*

*рах для E типов колебаний с уменьшением объема резонатора его резонансная частота уменьшается.*

## ВВЕДЕНИЕ

Аксиально-симметричные микроволновые резонаторы широко применяются в современных электронных

приборах, например, волномерах, датчиках, СВЧ-фильтрах и т. п. [1]. Наиболее широкое распространение получили устройства на основе микроволновых цилиндрических резонаторов [2] с рабочим типом колебаний  $H_{011}$ . Однако, обеспечение гарантированно одномодового режима работы этих резонаторов сложно осуществимо и требует непростых технических решений, позволяющих предотвратить возникновение высших типов колебаний с резонансными частотами, попадающими в тот же частотный интервал, которому принадлежит и основной рабочий тип колебаний [2]. Особенно сложно устранять вырождение типов колебаний  $H_{011}$  и  $E_{111}$ .

### ПОСТАНОВКА ЗАДАЧИ

В ряде работ [3, 4] было показано, что в аксиально-симметричных резонаторных системах, содержащих участки с конической образующей, вырождение типов колебаний  $H_{011}$  и  $E_{111}$  устраняется. При этом были рассмотрены случаи резонаторов в виде конуса и усеченного конуса, то есть предполагалось наличие плоских оснований. Анализ полей проводился в сферической либо в цилиндрической системах координат (для малых углов наклона конических образующих). Определение собственных резонансных частот сводилось к поиску корней уравнений, содержащих функции Бесселя и присоединенные функции Лежандра.

Представляет интерес исследовать структуру электромагнитного поля и поведение собственных резонансных частот при замене плоских оснований на сферические, то есть перейти к резонаторам в виде сферического сектора и усеченного сферического сектора. Целью таких исследований является оценка влияния кривизны оснований, а, следовательно, и угла при вершине конического сектора, на снятие вырождения типов колебаний  $H_{011}$ ,  $E_{111}$ , трансформацию других типов колебаний с близкими частотами.

Для изучения особенностей поведения резонансных частот были рассмотрены резонаторы в виде:

- сферического сектора со сферической формой торца (рис. 1, а);
- усеченного сферического сектора с двумя сферическими торцами (рис. 1, б).

При оценке влияния изменения геометрических размеров рассматриваемых резонаторов на снятие вырождения ближайших к  $H_{011}$  типов колебаний ограничимся поиском решений для типов колебаний  $E_{111}$  и  $E_{112}$ . При этом будем рассматривать собственные колебания без учета влияния элементов возбуждения и связи.

Резонансные частоты резонаторов будем находить на основе решения уравнений Максвелла в сферической системе координат [3] для поперечно-электриче-

ских  $TE_{nmp}$  ( $r, \varphi, \theta$ ) и поперечно-магнитных  $TM_{nmp}$  ( $r, \varphi, \theta$ ) колебаний. Под  $r, \varphi, \theta$  понимаются координаты в сферической системе координат.

Для получения уравнения, корни которого и определяют резонансные частоты резонатора в виде сферического сектора радиуса  $R$  со сферической формой торца, необходимо воспользоваться граничными условиями для электрического поля  $\vec{E}$  на металлической поверхности  $[\vec{r}_0 \times \vec{E}] = 0$  (явный вид для которого приведен, например, в [3]), где  $\vec{r}_0$  – единичный вектор нормальный к сферической поверхности.

С учетом известных соотношений [5] между сферическими и цилиндрическими функциями Бесселя уравнение для определения резонансных частот типов колебаний  $E_{111}$  и  $E_{112}$  принимает вид

$$\frac{d}{d(kr)} [\sqrt{kr} \cdot J_{\mu+0.5}(kr)]_{r=R} = 0 \quad (1)$$

или

$$J_{\mu-0.5}(kR) - \frac{2n-1}{2kR} J_{\mu+0.5}(kR) = 0, \quad (2)$$

а для колебаний  $H_{011}$

$$J_{\nu+0.5}(kR) = 0, \quad (3)$$

где  $\mu$  и  $\nu$  порядок цилиндрической функции Бесселя  $J_\mu$  и  $J_\nu$  для поперечно-магнитных и поперечно-электрических колебаний, соответственно;  $k = 2 \cdot \pi \cdot (f/c)$  – волновое число;  $f$  – резонансная частота;  $c$  – скорость света в вакууме.

Для определения порядков степеней  $\mu$  и  $\nu$  присоединенных функций Лежандра, входящих в выражения для полей сферического резонатора, использовались граничные условия на конической поверхности сферического конуса при  $\theta = \theta_0$ , сводящиеся к требованию обращения в нуль тангенциальной составляющей электрического поля (в данном случае  $E_\varphi$  и  $E_r$ ). С учетом одинаковой зависимости обеих компонент от угла  $\theta$ , условие равенства нулю обеих компонент сводится к одинаковым уравнениям, а именно: для поперечно-магнитных колебаний  $E_{11p}$

$$P_\mu^1(\cos(\theta)) = 0, \quad (4)$$

а для поперечно-электрических колебаний  $H_{011}$

$$\frac{d}{d\theta} P_\nu^1(\cos(\theta)), \quad (5)$$

где  $P_\mu^1, P_\nu^1$  присоединенные функции Лежандра первой степени порядка  $\mu$  и  $\nu$ , соответственно.

Для удобства вычислений функций Бесселя значения  $\theta$  подбирались таким образом, чтобы они соответствовали целым порядкам функции Бесселя. Все

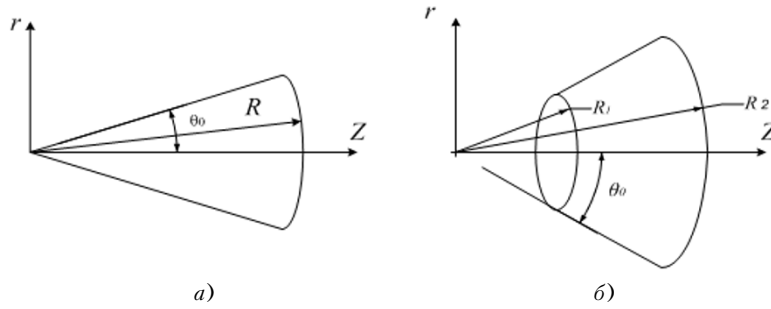


Рисунок 1 – Вид исследованных микроволновых резонаторов:  
а – сферический сектор; б – усеченный сферический сектор

полученные результаты приводились к безразмерному виду. Геометрические размеры резонаторов нормировались к радиусу  $R$  торца основания сферического сектора. Частотные характеристики резонаторов представлялись через безразмерную величину  $k \cdot R$ .

При рассмотрении резонатора в виде усеченного сферического сектора (рис. 1, б) граничные условия на сферических торцевых поверхностях при  $r = R_1$  и  $r = R_2$  и для конической поверхности при  $\theta = \theta_0$  задавались так же как и для микроволнового резонатора в виде сферического сектора со сферической формой торца. Геометрические размеры усеченных резонаторов нормировались к радиусу  $R_2$ .

Соответствующее уравнение для резонансных частот для типов колебаний  $E_{111}$  и  $E_{112}$  принимает вид:

$$\begin{aligned} & \left[ J_{n-1}(k\eta) - \frac{2n-1}{2k\eta} J_n(k\eta) \right] \times \\ & \times \left[ N_{n-1}(kR_2) - \frac{2n-1}{2kR_2} N_n(kR_2) \right] - \\ & - \left[ J_{n-1}(kR_2) - \frac{2n-1}{2kR_2} J_n(kR_2) \right] \times \\ & \times \left[ N_{n-1}(k\eta) - \frac{2n-1}{2k\eta} N_n(k\eta) \right] = 0, \end{aligned} \quad (6)$$

а для типа колебаний  $H_{011}$

$$J_m(k\eta) \cdot N_m(kR_2) - J_m(kR_2) \cdot N_m(k\eta) = 0, \quad (7)$$

где  $\eta = R_1/R_2$ ,  $N_j$  – функция Неймана  $j$ -го порядка.

Исследование влияния геометрических размеров и формы резонаторов на структуру их электромагнитных полей, выполнялось путем численного расчета методом конечных элементов (МКЭ) [6]. Предполагалось, что резонаторы изготовлены из меди, а их внутренний объем заполнен воздухом. Для сокращения объема вычислений и, с учетом азимутальной симметрии, расчеты проводились в областях, ограниченных двумя продольными секущими плоскостями,

расположенными под углом  $10^\circ$  друг к другу. На поверхности секущих плоскостей задавались граничные условия, соответствующие идеальной электрической стенке. Максимальный линейный размер элементов расчетной сетки не превышал  $0,05$  радиуса  $R_2$  торцевой поверхности резонатора.

### ОБСУЖДЕНИЕ РЕЗУЛЬТАТОВ РАСЧЕТОВ

На рис. 2 представлены рассчитанные с помощью уравнений (2)–(3) значения нормированных резонансных частот  $k \cdot R$  для типов колебаний  $H_{011}$ ,  $E_{111}$  и  $E_{112}$  в зависимости от угла  $\theta_0$  сферического сектора.

Как видно из представленных результатов, при увеличении угла  $\theta_0$  сферического сектора, начиная со значения, приближенно равного  $4,3^\circ$ , наблюдается рост разности значений резонансных частот для колебаний типов  $H_{011}$ ,  $E_{111}$  и  $E_{112}$  и, соответственно, снимается вырождение рассматриваемых типов колебаний. Наиболее значительный рост разности значений резонансных частот между типами колебаний  $H_{011}$  и  $E_{111}$ , а также  $H_{011}$  и  $E_{112}$ , наблюдается в диапазоне изменения угла  $\theta_0$  сферического сектора от  $4,3^\circ$  до  $60^\circ$ .

Для оценки влияния формы торца резонатора на его параметры с помощью МКЭ были проведены расчеты резонансной частоты и добротности при различных значениях  $\theta_0$  для резонаторов с плоскими и сферическими торцами. Часть полученных результатов приведена в табл. 1.

Из представленных в табл. 1 результатов видно, что нормированная резонансная частота резонатора со сферической формой торца имеет более низкое значение по сравнению с резонансной частотой резонатора с плоским основанием и разница между ними возрастает пропорционально росту  $\theta_0$ . Добротность резонаторов со сферической формой торцевой поверхности, при углах  $\theta_0 > 16^\circ$  имеет более высокое значение по сравнению с добротностью резонатора с плоским торцом. Однако, при значениях угла  $\theta_0 < 16^\circ$

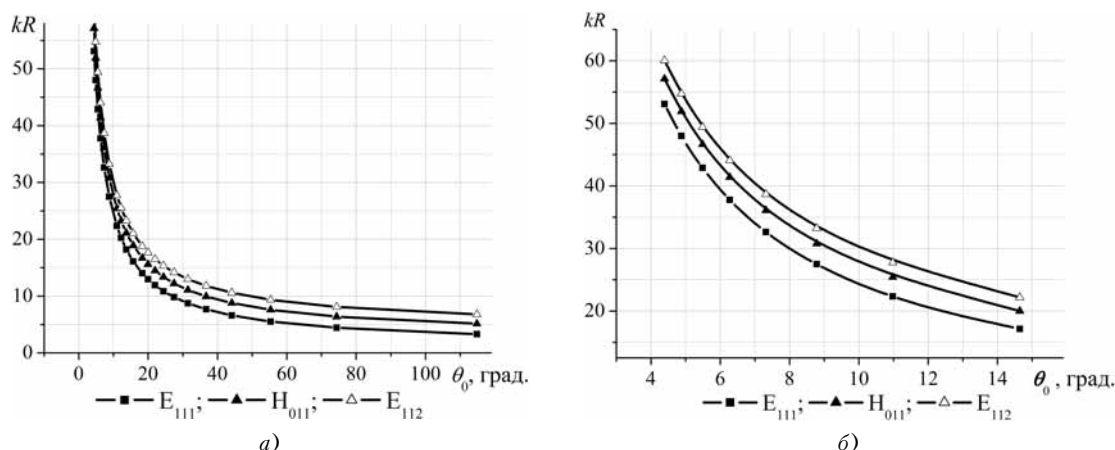


Рисунок 2 – Зависимость нормированных резонансных частот резонатора в виде сферического сектора для типов колебаний  $H_{011}$ ,  $E_{111}$  и  $E_{112}$  от угла  $\theta_0$  сферического сектора:

*a* – во всем исследованном диапазоне изменения  $\theta_0$ ; *б* – при малых значениях  $\theta_0$

Таблица 1 – Зависимость нормированной резонансной частоты  $k \cdot R$  и добротности  $Q$  резонатора в виде сферического сектора от угла  $\theta_0$  сферического сектора и формы торца

$\theta_0$ , град	Сферический торец		Плоский торец	
	$k \cdot R$	$Q$	$k \cdot R$	$Q$
16	19,01	22130	18,51	22320
21	14,95	25350	15,70	22410
26	12,85	27620	13,83	22540
30	11,45	29540	12,69	26350
34	10,47	31300	11,98	26940
38	9,77	32890	11,56	27330
41	9,22	34460	11,29	27570

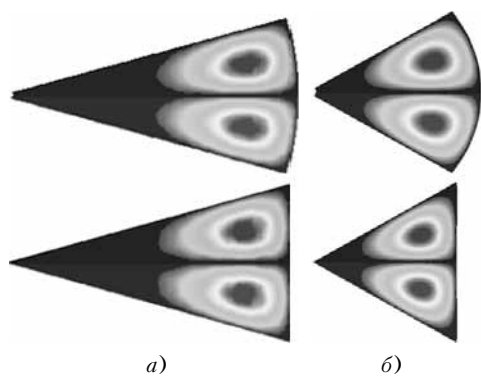


Рисунок 3 – Структура электрического поля в продольном сечении резонаторов с различной формой торцов:

*a* – при  $\theta_0 = 16^\circ$ ; *б* – при  $\theta_0 = 38^\circ$

и одинаковых  $R$ , добротности этих двух видов резонаторов приближенно одинаковы.

Данные в табл. 1 подтверждают представленные на рис. 3 результаты визуализации рассчитанной структуры электрического поля в продольном сечении ре-

зонаторов с плоскими и сферическими торцами для углов  $\theta_0$ , равных 16 и 38 градусам.

Из представленных рисунков видно, что при малых углах сферического сектора структура электрического поля в случае плоского торца визуально идентична структуре поля резонатора со сферическим торцом. По мере увеличения угла  $\theta_0$  сферического сектора наблюдается рост степени отличий электрического поля резонатора, имеющего торец сферической формы, по сравнению с полем резонатора с плоским торцом.

На рис. 4 представлены результаты расчета резонансных частот резонатора в виде усеченного сферического сектора с  $\theta_0 = 11^\circ$  для всех рассматриваемых в данной работе типов колебаний. При изменении значения угла  $\theta_0$  наблюдается смещение частотного интервала, где возникают резонансы, однако характер зависимости резонансных частот от  $R_1/R_2$  и взаимное расположение резонансных кривых для рассматриваемых типов колебаний остается практически неизменным.

Анализ зависимостей резонансных частот типов колебаний  $E_{111}$ ,  $H_{011}$  и  $E_{112}$  резонатора в виде усечен-

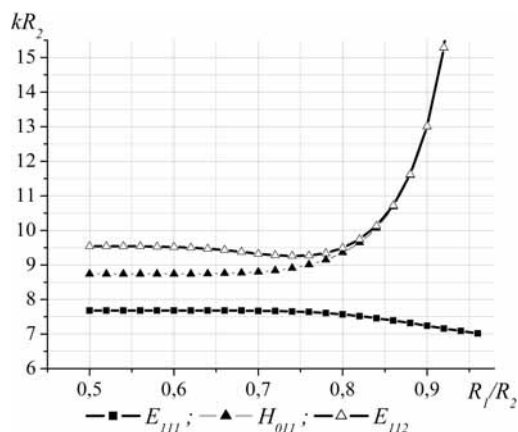


Рисунок 4 – Зависимость резонансных частот типов колебаний  $E_{111}$ ,  $H_{011}$  и  $E_{112}$  резонатора в виде усеченного сферического сектора от отношения радиусов сферических торцевых поверхностей при  $\theta_0 = 11^\circ$

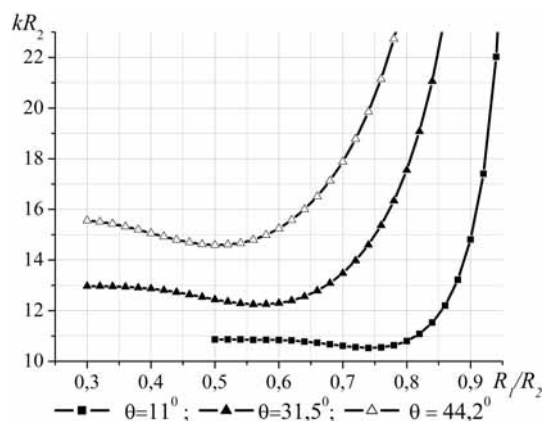


Рисунок 5 – Зависимость резонансных частот типа колебаний  $E_{112}$  резонатора в виде усеченного сферического сектора от отношения радиусов сферических торцевых поверхностей для ряда углов  $\theta_0$

ного сферического сектора от отношения радиусов сферических торцевых поверхностей, одна из которых показана на рис. 4, позволил выявить аномальный характер поведения резонансных частот для  $E$  типов колебаний. Так для колебания  $E_{111}$  это проявляется в том, что с уменьшением объема резонатора его резонансная частота уменьшается. Для типа колебаний  $E_{112}$ , помимо наличия падающего участка кривой, имеется еще и пологий минимум. Это хорошо видно при рассмотрении представленной на рис. 5 зависимости резонансных частот резонатора для колебаний типа  $E_{112}$  от отношения радиусов сферических торцевых поверхностей при углах  $\theta_0$ , равных  $11^\circ$ ,  $31,5^\circ$  и  $44,2^\circ$ .

Возможным объяснением аномального характера поведения резонансных частот для  $E$  типов колебаний может быть то, что при больших значениях отношения  $R_1/R_2$  структура электрического поля стремится к структуре колебания  $E_{110}$  (наблюдаемой в цилиндрическом резонаторе), для которой характерна возможность существования колебаний при расстояниях между торцевыми поверхностями намного меньших половины длины электромагнитной волны. Именно для типа колебаний  $E_{110}$  характерно отсутствие зависимости резонансной частоты от длины резонатора и ее уменьшение с увеличением радиуса резонатора.

Таким образом, для усеченного сферического конуса уменьшение резонансной частоты с уменьшением резонансного объема можно трактовать, как и уменьшение резонансной частоты колебаний типа  $E_{110}$  в цилиндрическом резонаторе при увеличении его радиуса, так как в усеченном коническом резонаторе при сближении торцевых поверхностей средний радиус резонансного объема увеличивается.

## ВЫВОДЫ

Проведенный анализ резонансных явлений для типов колебаний  $H_{011}$ ,  $E_{111}$  и  $E_{112}$  в резонаторах в виде сферического сектора показал:

- возможность снятия вырождения между типами колебаний  $H_{011}$ ,  $E_{111}$  и  $E_{112}$  в широком диапазоне изменения геометрических размеров резонаторов, причем разница резонансных частот в большей мере определяется углом при вершине сферического сектора;
- при малых углах сферического сектора структура электрического поля в случае плоского торца визуально идентична структуре поля резонатора со сферическим торцом;
- по мере увеличения угла  $\theta$  сферического сектора наблюдается рост степени отличия электрического поля резонатора, имеющего торец сферической формы, по сравнению с полем резонатора с плоским торцом;
- аномальный характер поведения резонансных частот для  $E$  типов колебаний в усеченных резонаторах, проявляющийся в том, что с уменьшением объема резонатора его резонансная частота уменьшается;
- для колебаний типа  $E_{112}$  в резонаторах в виде усеченного сферического конуса имеется пологий экстремум, содержащий участок с аномальным поведением резонансных частот.

## ПЕРЕЧЕНЬ ССЫЛОК

1. Makimoto M. Microwave resonators for wireless communication. Theory, design and application // M. Makimoto, S. Yamashita. – Berlin : Springer – Verlag, 2001. – 162 p.
2. Дробахин О. О. Датчики перемещений на основе круглых цилиндрических СВЧ-резонаторов / О. О. Дробахин, П. И. Заболотный, Е. Н. Привалов // Техника и приборы СВЧ. – 2008. – № 2. – С. 24–29.
3. Harrington R. F. Time-harmonic electromagnetic fields. – New York : Wiley, 2001. – 480 p.

4. Van't Hof J. P. Eigenfrequencies of a Truncated Conical Resonator via the Classical and Wentzel-Kramers-Brillouin Methods / J. P. Van't Hof, D. D. Stancil // Transactions on Microwave Theory and Techniques. – Vol. MTT-56. – 2008. – No. 8. – P. 1909–1916.
5. Справочник по специальным функциям / под. ред. М. Абрамовица и И. Стиган. – М.: Наука, 1979. – 830 с.
6. Григорьев А. Д. Электродинамика и микроволновая техника / А. Д. Григорьев. – СПб.: Лань, 2007. – 704 с.

Надійшла 21.04.2009

Досліджено вплив на резонансні властивості й структуру електромагнітного поля геометрії резонаторів, виконаних у вигляді сферичного сектора й усіченого сферичного сектора. Показано, що розходження

між резонансними частотами коливань типу  $H_{011}$ ,  $E_{111}$  і  $E_{112}$  у більшій мері визначається кутом при вершині сферичного сектора. Виявлено, що в усічених резонаторах для  $E$  типів коливань зі зменшенням об'єму резонатора його резонансна частота зменшується.

The influence of spherical and truncated spherical sector geometrical structure resonators on resonance frequency and em field structure is investigated. It has been shown that the resonance frequency discrepancies for  $H_{011}$ ,  $E_{111}$  and  $E_{112}$  modes are mainly defined by vertex angle of spherical sector. For  $E$  modes in truncated cavities an effect of resonance frequency diminishing with cavity volume reduction has been detected.

УДК 537.612, 537.635

І. В. Зависяк, Г. П. Головач, М. О. Попов

## МОДИФІКОВАНА ЕЛІПТИЧНА СИСТЕМА КООРДИНАТ ДЛЯ ЗАДАЧ МАТЕМАТИЧНОЇ ФІЗИКИ З СИМЕТРІЄЮ ЕЛІПТИЧНОГО ЦИЛІНДРА

Представлена оригінальна модифікована еліптична система координат, розглянуто її властивості, наведено корисні співвідношення. Запропоновано застосовувати цю систему координат при розгляді задач математичної фізики, які мають симетрію еліптичного циліндра, зокрема, задач про знаходження власних хвиль і коливань в хвилеводах і резонаторах еліптичного перерізу та продемонстровано приклади її використання.

### ВСТУП

Для розв'язання багатьох задач математичної фізики необхідно використовувати еліптичну систему координат. Так, для електродинаміки і техніки як НВЧ, так і оптичного діапазонів значний інтерес становлять еліптичні хвилеводи і резонатори, які займають проміжне положення між своїми аналогами круглого та прямокутного перерізу. В них, на відміну від круглої геометрії, знімається азимутальне виродження власних хвиль і коливань, фіксується поляризація, розширюється робочий частотний діапазон хвилеводів, розріджується спектр резонаторів [1, 2].

В спілхвильовій електродинаміці еліптична система координат застосовувалася в [3], де розглядалися магнітостатичні хвилі в еліптичному циліндричному феритовому хвилеводі. Характерною особливістю розв'язку є те, що при цьому магнітостатичний потенціал (і магнітне поле також) повинні бути представлені у вигляді нескінченних рядів з функцій Матьє (ззовні зразка) чи модифікованих функцій

Матьє (всередині зразка). При цьому розв'язок характеристичного рівняння можливий лише наближеними числовими методами.

Метою даної роботи є запровадження нового ефективного аналітичного методу аналізу електродинамічних задач в системах з симетрією еліптичного циліндра і демонстрація її ефективності на прикладі повного аналітичного розв'язання спектральної задачі для поверхневих магнітостатичних коливань (ПМСК) в феромагнітних системах з геометрією еліптичного циліндра.

### МОДИФІКОВАНА ЕЛІПТИЧНА СИСТЕМА КООРДИНАТ

Зазвичай, при розв'язанні задач математичної фізики в областях, які мають симетрію еліптичного циліндра, використовується стандартна еліптична система координат [3, 4], у якої одна із координатних поверхонь є еліпсом:

$$x = c \cdot ch\xi \cdot \cos \eta, \quad y = c \cdot sh\xi \cdot \sin \eta, \quad z = z, \quad (1)$$

де  $\eta \in [-\pi, \pi]$ ,  $\xi \in [0, \infty)$ ,  $z \in (-\infty, \infty)$ ,  $c$  – напівфокальна відстань:  $c^2 = a^2 - b^2$ , де  $a$  і  $b$  – це відповідно велика і мала півосі еліпса, який відповідає  $\xi = \text{const}$ .

На площині  $xy$  заміна змінних (1) перетворює внутрішню область еліпса з півосями  $a$  і  $b$  на прямо-



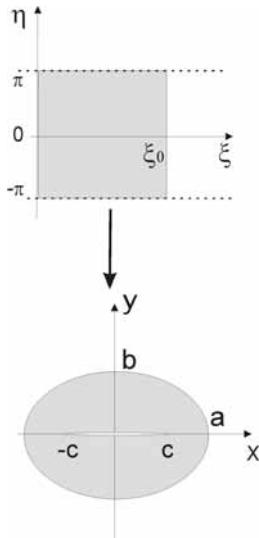


Рисунок 1 – Конформне перетворення області при заміні (1)

кутну область  $\eta \in [-\pi, \pi]$ ,  $\xi \in [0, \xi_0]$ ,  $\xi_0 = \frac{1}{2} \ln \left( \frac{a+b}{a-b} \right)$  або, враховуючи періодичність змінної  $\eta$ , на смугу  $\xi \in [0, \xi_0]$  (рис. 1).

При цьому розріз  $x \in [-c, c]$  переходить у пряму  $\xi = 0$ , а поверхня еліпса – в  $\xi = \xi_0$ . Аналогічно зовнішня частина простору переходить у смугу  $\xi \in [\xi_0, \infty)$ .

Недоліком цієї системи координат є незручний перехід до полярної системи координат (ПСК). У цьому випадку  $c \rightarrow 0$ ,  $\xi \rightarrow \infty$ , відповідно, для певного виразу потрібно знайти подвійну границю при  $\lim_{\substack{c \rightarrow 0 \\ \xi \rightarrow \infty}}(\dots)$

одночасному виконанні умови  $\lim_{\substack{c \rightarrow 0 \\ \xi \rightarrow \infty}} \frac{ce^\xi}{2} = \rho$ ,  $\rho$  – радіус в ПСК. Дійсно, за цих умов  $\lim_{\substack{c \rightarrow 0 \\ \xi \rightarrow \infty}} (c \cdot sh\xi) =$

$= \lim_{\substack{c \rightarrow 0 \\ \xi \rightarrow \infty}} (c \cdot ch\xi) = \lim_{\substack{c \rightarrow 0 \\ \xi \rightarrow \infty}} \left( \frac{ce^\xi}{2} \right) \rightarrow \rho$ , і формула (1) переходить у відомий вираз для ПСК.

Якщо тепер у (1) одразу зробити заміну  $\frac{ce^\xi}{2} = \rho$ , то отримуємо:

$$\begin{aligned} x &= \left( \rho + \frac{c^2}{4\rho} \right) \cdot \cos\varphi, \quad y = \left( \rho - \frac{c^2}{4\rho} \right) \cdot \sin\varphi, \\ z &= z, \quad \varphi \in [-\pi, \pi], \\ \rho &\in \left[ \frac{c}{2}, \infty \right), \quad z \in (-\infty, \infty). \end{aligned} \quad (2)$$

При цьому перехід до полярної системи координат здійснюється простою підстановкою  $c = 0$ . Перетво-

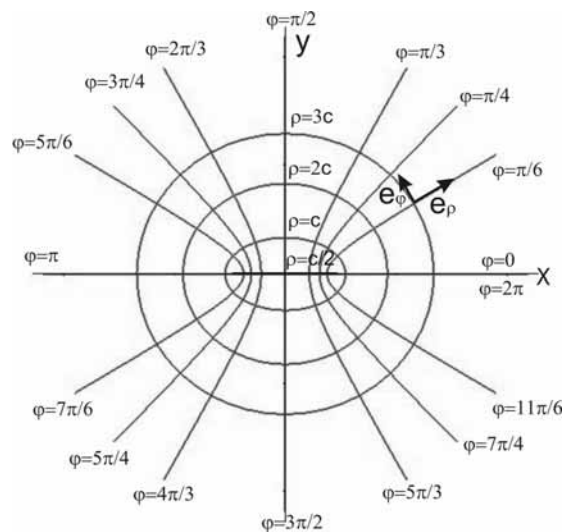


Рисунок 2 – Координатні поверхні та орти МЕСК

рення (2) будемо вважати означенням модифікованої еліптичної системи координат (МЕСК) і надалі використовуватимемо як заміну змінних для розв'язку поставленої задачі. МЕСК у багатьох відношеннях схожа з ПСК, що робить отримані результати більш наочними.

Розглянемо тепер детально властивості МЕСК. Як вже було сказано, зв'язок з декартовою системою координат визначається рівняннями (2), зворотне перетворення задається формулами:

$$\begin{aligned} \sin^2\varphi &= \frac{1}{2c^2} \times \\ &\times \left( c^2 - x^2 - y^2 + \sqrt{x^2(x^2 + 2y^2 - 2c^2) + (y^2 + c^2)^2} \right), \\ \rho &= \frac{1}{2} \left( \frac{x}{\cos\varphi} + \frac{y}{\sin\varphi} \right). \end{aligned} \quad (3)$$

Конкретний корінь першого рівняння системи (3) обирається, виходячи зі знаків змінних  $x$  та  $y$ .

Координатні поверхні являють собою набір конфокальних еліпсів ( $\rho = R = \text{const}$ ), які описуються рівнянням  $\frac{x^2}{\left(R + \frac{c^2}{4R}\right)^2} + \frac{y^2}{\left(R - \frac{c^2}{4R}\right)^2} = 1$ , та гіпербол ( $\varphi = \Phi = \text{const}$ ):  $\frac{x^2}{c^2 \cos^2\Phi} - \frac{y^2}{c^2 \sin^2\Phi} = 1$ , зображених на рис. 2.

Еліпсу з півосями  $a$  і  $b$  у даній системі координат відповідає значення  $\rho = \frac{a+b}{2}$ .

Розрахуємо масштабні множники Ламе для МЕСК:

$$h_\rho = \sqrt{\left(\frac{\partial x}{\partial \rho}\right)^2 + \left(\frac{\partial y}{\partial \rho}\right)^2} = \sqrt{\left(1 - \frac{c^2}{4\rho^2}\right)^2 + \frac{c^2}{\rho^2} \sin^2 \varphi},$$

$$h_\varphi = \sqrt{\left(\frac{\partial x}{\partial \varphi}\right)^2 + \left(\frac{\partial y}{\partial \varphi}\right)^2} =$$

$$= \sqrt{\rho^2 \left[ \left(1 - \frac{c^2}{4\rho^2}\right)^2 \cos^2 \varphi + \left(1 + \frac{c^2}{4\rho^2}\right)^2 \sin^2 \varphi \right]} = \rho h_\rho,$$

$$h_z = 1. \quad (4)$$

Бачимо, що при  $c = 0$ , формули (4) переходять у вирази для циліндричної системи координат ( $h_\rho = 1$ ,  $h_\varphi = \rho$ ,  $h_z = 1$ ), як і має бути.

Якобіан МЕСК має вигляд:

$$\left| \frac{\partial(x, y, z)}{\partial(\rho, \varphi, z)} \right| = \rho \left( 1 + \frac{c^4}{16\rho^4} - \frac{c^2}{2\rho^2} \cos(2\varphi) \right).$$

Якобіан дорівнює нулю лише в двох точках:  $\rho = c/2$ ,  $\varphi = 0$  та  $\rho = c/2$ ,  $\varphi = \pi$ .

Скориставшись загальним виглядом оператора Лапласа в довільній криволінійній системі координат [4]

$$\Delta V = \frac{1}{h_1 h_2 h_3} \times$$

$$\times \left[ \frac{\partial}{\partial x_1} \left( \frac{h_2 h_3}{h_1} \frac{\partial V}{\partial x_1} \right) + \frac{\partial}{\partial x_2} \left( \frac{h_3 h_1}{h_2} \frac{\partial V}{\partial x_2} \right) + \frac{\partial}{\partial x_3} \left( \frac{h_1 h_2}{h_3} \frac{\partial V}{\partial x_3} \right) \right],$$

можемо обчислити лапласіан в МЕСК:

$$\Delta V = \frac{1}{h_\rho h_\varphi} \left[ \frac{\partial}{\partial \rho} \left( \rho \frac{\partial V}{\partial \rho} \right) + \frac{1}{\rho} \frac{\partial^2 V}{\partial \varphi^2} \right] + \frac{\partial^2 V}{\partial z^2}.$$

Диференціальні оператори в МЕСК визначаються наступними формулами

$$\text{grad} V = \left( \frac{1}{h_\rho} \frac{\partial V}{\partial \rho}, \frac{1}{\rho h_\varphi} \frac{\partial V}{\partial \varphi}, \frac{\partial V}{\partial z} \right),$$

$$\text{div} \vec{A} = \frac{1}{\rho h_\rho^2} \left( \frac{\partial}{\partial \rho} (\rho h_\rho A_\rho) + \frac{\partial}{\partial \varphi} (h_\rho A_\varphi) \right) + \frac{\partial A_z}{\partial z},$$

$$(\text{rot} \vec{A})_\rho = \frac{1}{\rho h_\rho} \frac{\partial A_z}{\partial \varphi} - \frac{\partial A_\varphi}{\partial z},$$

$$(\text{rot} \vec{A})_\varphi = \frac{\partial A_\rho}{\partial z} - \frac{1}{h_\rho} \frac{\partial A_z}{\partial \rho},$$

$$(\text{rot} \vec{A})_z = \frac{1}{\rho h_\rho^2} \left( \frac{\partial}{\partial \rho} (\rho h_\rho A_\varphi) - \frac{\partial}{\partial \varphi} (h_\rho A_\rho) \right).$$

Для того, щоб розглянути трансформацію області, зробимо допоміжні перетворення:

$$u = \rho \cos \varphi, \quad v = \rho \sin \varphi. \quad (5)$$

Тоді отримуємо:

$$x = u \left( 1 + \frac{c^2}{4(u^2 + v^2)} \right), \quad y = v \left( 1 - \frac{c^2}{4(u^2 + v^2)} \right).$$

Побудуємо відображення:

$$Z = x + iy = u \left( 1 + \frac{c^2}{4(u^2 + v^2)} \right) + iv \left( 1 - \frac{c^2}{4(u^2 + v^2)} \right) =$$

$$= u + iv + \frac{c^2(u - iv)}{4(u^2 + v^2)} = u + iv + \frac{c^2}{4(u + iv)}.$$

Позначивши  $W = u + iv$ , отримуємо, що

$$Z = W + \frac{c^2}{4W}. \quad (6)$$

Вираз (6) є відомим конформним перетворенням, що переводить коло в еліпс [5, 6]. Таким чином, перетворення змінних, що задається формулами (2), складається з двох частин: 1) конформне відображення (6); 2) перехід від декартових змінних ( $u, v$ ) до полярних ( $\rho, \varphi$ ). Відзначимо, що в [7], при розв'язанні задачі спінхвильової електродинаміки з симетрією еліптичного циліндра, використовувалися декартові змінні, згідно з (6). Це призвело до нефізичної розбіжності компонент магнітного поля ПМСК у фокусах еліпса і не дозволило авторам аналітично побудувати розподіл магнітного поля. При використанні МЕСК розбіжності у фокусах еліпса не виникає.

Таким чином, при використанні (2) внутрішня частина еліпса з півосями  $a$  і  $b$  переходить у смугу  $\rho \in \left[ \frac{c}{2}, \frac{a+b}{2} \right]$  (з урахуванням періодичності по  $\pi$ ) через проміжну область у формі кільця з внутрішнім і зовнішнім радіусами  $\frac{c}{2}$  та  $\frac{a+b}{2}$ , відповідно. Зовнішня область – у смугу  $\rho \in \left[ \frac{a+b}{2}, \infty \right)$  (рис. 3).

Наведені вище співвідношення створюють підґрунтя для використання МЕСК в задачах математичної фізики з симетрією еліптичного циліндра, зокрема для аналізу металевих, діелектричних та гіротропних хвильоводів та резонаторів еліптичного перерізу.

### СПЕКТРАЛЬНА ЗАДАЧА ДЛЯ ФЕРОМАГНІТНИХ СТРУКТУР З СИМЕТРІЄЮ ЕЛІПТИЧНОГО ЦИЛІНДРА

Продемонструємо метод розділення змінних на основі застосування модифікованої еліптичної системи

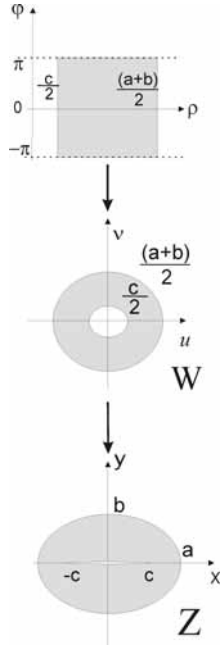


Рисунок 3 – Конформне перетворення областей при заміні (2)

координат при розв'язанні спектральних задач в магнітостатичних резонаторах еліптичної геометрії. Будемо виходити з рівняння Уокера [8]:

$$\operatorname{div}(\mu \operatorname{grad} \Psi) = 0, \quad (7)$$

тут  $\hat{\mu}$  – тензор магнітної проникності, що задається співвідношеннями:

$$\hat{\mu} = \begin{pmatrix} \mu & -i\mu_a & 0 \\ i\mu_a & \mu & 0 \\ 0 & 0 & 1 \end{pmatrix}$$

де  $\mu = \frac{(\omega^2 - \omega_1^2)}{(\omega^2 - \omega_H^2)}$ ,  $\mu_a = \frac{\omega\omega_M}{(\omega^2 - \omega_H^2)}$ ,  $\omega_1^2 = \omega_H(\omega_H + \omega_M)$ ,  $\omega_M = \gamma 4\pi M_0$ ,  $\omega_H = \gamma H_{\text{eff}0}$ ,  $\omega$  – частота,  $M_0$  – намагніченість насичення,  $H_{\text{eff}0}$  – ефективне сталие магнітне поле,  $\gamma$  – гіромагнітне відношення. В формулі (7)  $\Psi$  – магнітостатичний потенціал, що уводиться згідно з  $\vec{h} = \operatorname{grad}\Psi$ ,  $\vec{h}$  – змінна складова магнітного поля. В МЕСК (7) набуває вигляду

$$\mu \frac{1}{h_p h_\phi} \left[ \frac{\partial}{\partial \rho} \left( \rho \frac{\partial \Psi}{\partial \rho} \right) + \frac{1}{\rho} \frac{\partial^2 \Psi}{\partial \phi^2} \right] + \frac{\partial^2 \Psi}{\partial z^2} = 0.$$

Для розділення змінних підставимо  $\Psi(\rho, \phi, z) = R(\rho)\Phi(\phi)Z(z)$ :

$$\mu \frac{1}{h_p h_\phi} \left[ \frac{1}{R} \frac{\partial}{\partial \rho} \left( \rho \frac{\partial R}{\partial \rho} \right) + \frac{1}{\Phi} \frac{1}{\rho} \frac{\partial^2 \Phi}{\partial \phi^2} \right] + \frac{1}{Z} \frac{\partial^2 Z}{\partial z^2} = 0.$$

Відокремимо функцію  $Z(z)$ :  $\frac{1}{Z} \frac{\partial^2 Z}{\partial z^2} = -k^2$ , тоді:

$$\frac{1}{R} \frac{\partial}{\partial \rho} \left( \rho \frac{\partial R}{\partial \rho} \right) + \frac{1}{\Phi} \frac{1}{\rho} \frac{\partial^2 \Phi}{\partial \phi^2} + \frac{h_p h_\phi k^2}{\mu} = 0.$$

Підставивши в явному вигляді коефіцієнти Ламе, отримаємо:

$$\left[ \frac{1}{R} \rho \frac{\partial}{\partial \rho} \left( \rho \frac{\partial R}{\partial \rho} \right) + \frac{k^2}{\mu} \rho^2 \left( 1 - \frac{c^2}{4\rho^2} \right)^2 \right] + \left[ \frac{1}{\Phi} \frac{\partial^2 \Phi}{\partial \phi^2} + \frac{k^2 c^2}{\mu} \sin^2 \phi \right] = 0.$$

Остаточно розділивши змінні, маємо:

$$\frac{1}{\Phi} \frac{\partial^2 \Phi}{\partial \phi^2} + \frac{k^2 c^2}{\mu} \sin^2 \phi = -n^2,$$

$$\frac{1}{R} \rho \frac{\partial}{\partial \rho} \left( \rho \frac{\partial R}{\partial \rho} \right) + \frac{k^2}{\mu} \rho^2 \left( 1 - \frac{c^2}{4\rho^2} \right)^2 = n^2.$$

Звідки після перетворень для першого рівняння отримуємо:

$$\frac{\partial^2 \Phi}{\partial \phi^2} + \left( \left( n^2 + \frac{k^2 c^2}{2\mu} \right) - 2 \frac{k^2 c^2}{4\mu} \cos 2\phi \right) \Phi = 0. \quad (8)$$

Це є рівняння Мат'є [4],  $2\pi$  – періодичними по змінній  $\phi$  розв'язками якого є парні та непарні функції Мат'є першого роду (функції Мат'є з цілим індексом)  $ce_m\left(\phi, \frac{k^2 c^2}{4\mu}\right)$  та  $se_m\left(\phi, \frac{k^2 c^2}{4\mu}\right)$ . Друге рівняння можна переписати в наступному вигляді:

$$\frac{\partial^2 R}{\partial \rho^2} + \frac{1}{\rho} \frac{\partial R}{\partial \rho} + \left( \frac{k^2}{\mu} \left( 1 - \frac{c^2}{4\rho^2} \right)^2 - \frac{n^2}{\rho^2} \right) R = 0. \quad (9)$$

Бачимо, що при  $c = 0$ , рівняння (9) перетворюється у рівняння Бесселя.

Запропонований математичний апарат було застосовано для аналізу власних поверхневих магнітостатичних коливань в електродинамічних структурах еліптичного перерізу: намагнічених вздовж осі отвору у феромагнетик [9], феромагнітному циліндрі [9] та трубці [10, 11], а також еліптичних циліндричних доменах у матеріалах з одновісною анізотропією [12]. Було знайдено в аналітичному вигляді вирази для резонансних частот та власних функцій, проаналізовано поляризацію та просторовий розподіл магнітного поля. Порівняння теоретичних частотно-польових залежностей з експериментальними показало їх гарне узгодження. Також продемонстровано

можливість застосування побудованої теорії для розрахунків спектрів феромагнітних нанодотів.

## ВИСНОВКИ

В роботі була представлена оригінальна модифікована еліптична система координат, розглянуто її основні властивості, знайдено масштабні множники Ламе, побудовано координатні поверхні, записано вирази для диференціальних операторів градієнта, ротора, дивергенції та оператора Лапласа. Запропоновано використовувати модифіковану еліптичну систему координат при розв'язанні задач математичної фізики в системах з геометрією еліптичного циліндра і наведено приклади успішного застосування.

## ПЕРЕЛІК ПОСИЛАНЬ

1. Mahmoud S. F. Electromagnetic waveguides: theory and applications / S. F. Mahmoud. – London : Peter Peregrinus Ltd., 1991. – 228 p.
2. Dyott R. B. Elliptical Fiber Waveguides / R. B. Dyott. – Norwood, MA : Artech House, 1995. – 217 p.
3. De Wames R. E. Magnetostatic surface modes in an axially magnetized elliptical cylinder / R. E. De Wames, T. Wolfram // Appl. Phys. Lett. – 1970. – Vol. 16, № 8. – P. 305–308.
4. Анго А. Математика для электро- и радиоинженеров / А. Анго. – М. : Наука, 1965. – 779 с.
5. Иванов В. И. Конформные отображения и их приложения / В. И. Иванов, В. Ю. Попов. – М. : Едиториал УРСС, 2002. – 324 с.
6. Miyazaki Y. Radio Wave absorption analysis of coated elliptic cylinders with lossy magnetic ferrite films using conformal mapping method / Y. Miyazaki // 3rd International Symposium on Electromagnetic Compatibility. – Beijing (China), 2002. – P. 432–437.

УДК: 538.935, 538.915, 538.953

V. V. Pogosov, E. V. Vasyutin, A. V. Babich

# FEATURE OF MAGIC METAL NANOCCLUSERS IN MOLECULAR TRANSISTOR

*Effects of the charging and single-electron tunneling in a cluster structure are investigated theoretically. In the framework of the particle-in-a-box model for the spherical and disk-shaped gold clusters, the electron spectrum and the temperature dependence of the electron chemical potential are calculated. Difference between the chemical potentials of massive electrodes and island's one leads to its charging. We show that the effective residual charge is equal to the non-integer value of elementary charge  $e$  and depends on the cluster's shape. The equations for the analysis of the current-voltage characteristic are used under restrictions associated with the Coulomb instability of a cluster. For single-electron molecular transistors the non-monotonic size dependences of current gap and its voltage*

© Pogosov V. V., Vasyutin E. V., Babich A. V., 2009

7. Roussigne Y. Spin waves in stripe submitted to a perpendicular applied field / Y. Roussigne, P. Moch // J. Phys. Condense Matter. – 2005. – Vol. 17, № 10. – P. 1645–1652.
8. Гуревич А. Г. Магнитные колебания и волны / А. Г. Гуревич, Г. А. Мелков. – М. : Физматлит., 1994. – 464 с.
9. Зависяк И. В. Поверхностные магнитостатические колебания в эллиптических отверстиях и цилиндрах / И. В. Зависяк, Г. П. Головач, М. А. Попов, В. Ф. Романюк // РЭ. – 2006. – Т. 51, № 2. – С. 213–220.
10. Попов М. А. Поверхностные магнитостатические колебания в ферритовых трубках эллиптического сечения / М. А. Попов, И. В. Зависяк // Известия вузов. Радиоэлектроника. – 2006. – Т. 49, № 6. – С. 3–11.
11. Попов М. А. Modelling of the magnetostatic surface oscillations in elliptical nanotubes / М. А. Попов, И. В. Зависяк // УФЖ. – 2008. – Т. 53, № 7. – С. 706–711.
12. Попов М. А. Поверхностные магнитостатические колебания в эллиптических цилиндрических магнитных доменах / М. А. Попов, И. В. Зависяк // ФТТ. – 2009. – Т. 51, № 1. – С. 81–84.

Надійшла 7.04.2009  
Після доробки 28.05.2009

*Представлена оригинальная модифицированная эллиптическая система координат, рассмотрены ее свойства, приведены полезные соотношения. Предложено использовать эту систему координат при рассмотрении задач математической физики с симметрией эллиптического цилиндра, в частности, задач про нахождение собственных волн и колебаний в волноводах и резонаторах эллиптического сечения, продемонстрированы примеры ее применения.*

*The original modified elliptical coordinate system is submitted, its properties are considered, useful relationships are given. It was suggested to use this coordinate system for mathematical physics problems with the symmetry of the elliptical cylinder, in particular for eigenwaves and eigen oscillations problems in waveguides and of elliptic cross-section resonators, the examples of application were demonstrated.*

*asymmetry are computed. We suggest that an overheating of electron subsystem leads to the disappearance of a current gap and gradual smoothing of current-voltage curves that is observed experimentally.*

## 1 INTRODUCTION

The nanodispersed systems are prospective object of nanotechnology [1, 2, 3, 4, 5, 6]. Transport of electrical charge across a nanoscale tunnel junction is accompanied by many effects, such as the Coulomb blockade of the average current, transfer of energy between elec-

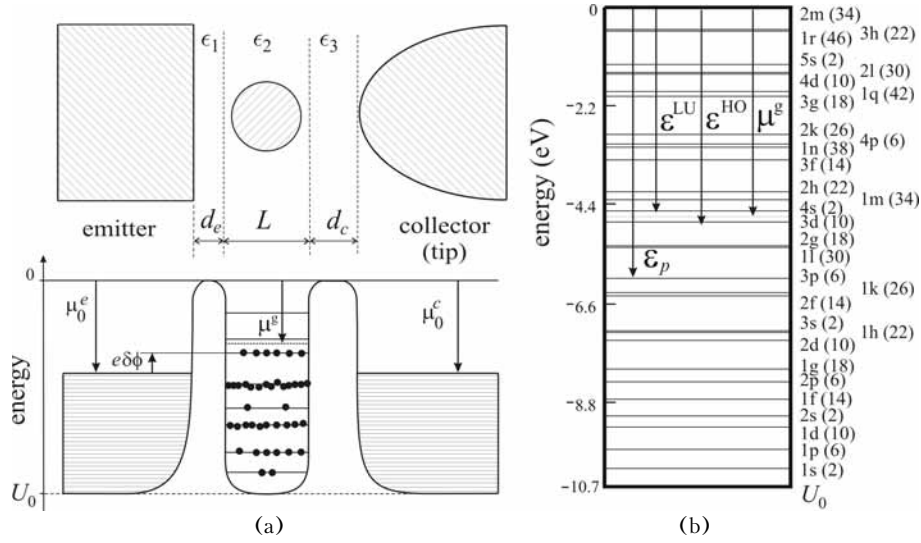


Figure 1:

(a) The energy diagram at  $T = 0$  for the structure Au/Au<sub>40</sub>/Au before application of voltage.

In experiments [11, 12, 13, 14, 15]  $\epsilon_1 = \epsilon$ ,  $\epsilon_2 = \epsilon_3 = 1$ ;

(b) The energy spectrum for spherical magic cluster Au<sub>196</sub>

trons and ions, and consequent heating of the junction [7, 8, 9]. In nanometer scale devices, electron transport can occur through well-resolved quantum states (e. g. single-electron transistor based on CdSe quantum rods [10]). If the temperature is increased, the Coulomb and quantum staircases of current are gradually smeared out by thermal fluctuations.

The tunneling current flowing through two massive electrodes can be controlled, if a cluster is placed between them. At first sight, the probability of electron tunneling (and consequently a value of the current) should be much greater in the presence of a granule between the reservoirs, than in the case of its absence. However, an opposite behavior (see inset in Fig. 2(a) of Ref. [11]) was observed in experiments for the spherical-like [11, 12, 13] and disk-shaped [14, 15] small clusters. Measured  $I$ - $V$  characteristics have a plateau of the zero current (a current gap). The metal-dielectric transition for gold cluster can appear [16].

Simple tunnel construction can be schematically represented by the distinctive “sandwich” [11, 12, 13, 14] (see Fig. 1). It consists of a thick gold film (emitter) covered by a dielectric one (with dielectric constant  $\epsilon \approx 3$ ). Disc-shaped [14] or spherical-like [13] gold clusters are self-organized on the dielectric layer (on the detecting and manipulating single molecules with STM see in Ref. [17]). Also, a tip of STM is used in the capacity of the third electrode (collector).

The experiments demonstrated the following features of the  $I(V)$  behavior:

1. The gap width of the zero conductance is *approximately* proportional to the *inverse radius* of the spheres (Figs. 1(c) and 2(a) in Ref. [11]) and disks

(Fig. 4 in Ref. [14]). This does not allow one to establish unequivocally classical or quantum origin of the gap. On the other hand, out of the current gap, the steps of the staircase are clearly visible (Fig. 3 in Ref. [14] and Fig. 1(b) in Ref. [15]).

2. For a disk, the gap width varies non-monotonically with alteration of the collector-cluster distance under the fixed emitter-cluster one (Fig. 3 in Ref. [15]).

3. The observed current gap decreases significantly as temperature increases from 5 K to 300 K in structure based on disk-shaped cluster of  $2R \approx 40$  Å (Fig. 2 in Ref. [14]).

Some of the experimental features of the  $I$ - $V$  curves were investigated in Ref. [18], however, the fact of smoothing of staircases for granule-molecule at low temperatures is still not understood. Such a smoothing is typical for molecular transistors [19].

Using the measured  $I(V)$  dependence, the capacitances and resistances of the tunnel junctions and the “residual” (fractional) charge  $Q_0$  of a grain were fitted in Ref. [12] according to the circuit approach proposed in “orthodox” theory [1]. We propose another way: to use the self-capacitance and to fit the temperature of granule.

The aim of this work is the computation of the current-voltage characteristic of the molecular transistor based on metal clusters. For this purpose, we use some results of physics for charged metal clusters [20]. The temperature features of the  $I$ - $V$  curves are explained by overheating of electron subsystem.

The structure of a granule (or cluster) changes as more and more atoms condense together. A fundamental

characteristic of metal clusters, that researchers must explain, is why certain sizes occur preferentially. The elements of the periodic table have heightened stability because those atoms possess a special number of electrons (magic numbers). The tendency for clusters to form in exactly magic sizes arises from the rules of quantum mechanics, which dictate that bound electrons can have only certain energies. So the existence of magic numbers for metal clusters makes sense: they correspond to the number of valence electrons which completely fill one or more shells in a cluster and make it especially stable, by analogy with filled proton and neutron energy shells in atomic nuclei [21]. The calculated magic numbers depend from the shape of cluster (i. e. on the configuration of the ions).

## 2 FORMULATION OF THE PROBLEM

We consider spherical gold clusters whose radii vary in the range  $2R \cong \{14, 28\} \text{Å}$ ,  $R = N_0^{1/3} r_0 \Rightarrow N_0 \cong \{100, 600\}$ , where  $N_0$  is the number of atoms,  $r_0$  is the atom density parameter, ( $r_0 = 3.01 a_0$  for gold,  $a_0$  is the Bohr radius). Similarly for disks of monatomic thickness:  $2R \cong \{10, 98\} \text{Å} \Rightarrow N_0 \cong \{14, 10^3\}$ .

The characteristic Coulomb energy of charging is  $E_C = e^2/C$  where  $C$  is self-capacitance of single granule in a vacuum (in the case of a disc, the capacitance can be estimated as for the oblate spheroids of equal volume [20]). The calculations of Ref. [18] demonstrated that these  $C$  are too small for the width of the current gap to be explained. The most obvious example is the case of a disc, since almost half of the disc surface contacts to the dielectric film. Therefore, for these granules we change  $C \Rightarrow (1 + \epsilon)C/2$ . Then, for discs and spheres we have  $E_C \cong \{1.60, 0.21\}$  and  $\{1.82, 1.06\}$  eV, respectively. We note that the value of the capacitance is sensitive to the shape of the granule surface, and even small deviation from the spherical shape can change significantly the capacitance. Temperature of structures (thermostat) is  $T \cong 30$  K. This enhances the importance of quantum mechanical effects.

Let's determine the electron spectrum in spherical and cylindrical wells (see Appendix A). The calculation of the electron spectrum in the cylindrical and spherical wells of the mentioned sizes with finite deepness yields different values for the spectrum discreteness in magic clusters  $\Delta\epsilon_p = \epsilon^{LU} - \epsilon^{HO}$  (see Fig. 2). In the nonmagic clusters the levels of lowest unoccupied states coincide with those of highest occupied ones,  $\epsilon^{LU} = \epsilon^{HO}$  at  $T = 0$ .

Thus, for the whole range of  $R$  in experiments [11, 12, 13, 14] we have to deal with a set of open 0D systems (quantum dots). The resulting inequality,

$$E_C \approx \Delta\epsilon_p \gg k_B T, \quad (1)$$

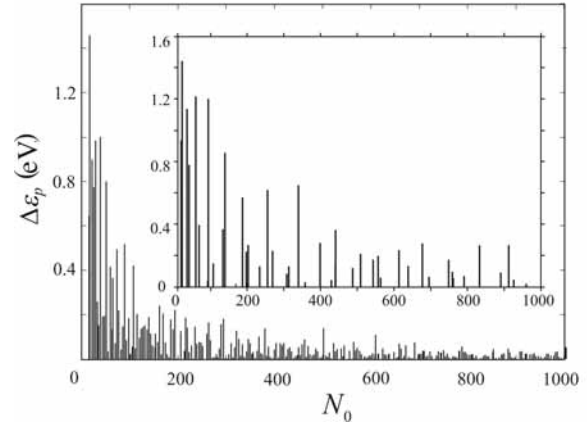


Figure 2 – Calculated specific difference between energies of lowest unoccupied electron state  $\epsilon^{LU}$  and highest occupied one  $\epsilon^{HO}$  in neutral discs and spheres (inset)  $\text{Au}_{N_0}$  at  $T = 0$

corresponds, apparently, to two coexisting structures at  $I$ - $V$  curves: effects of the spectrum quantization and the Coulomb blockade. However, detailed measurements in Refs. [2, 11, 12, 13, 22] performed to date do not yield an unequivocal conclusion about the effect of electron quantization levels upon the  $I(V)$ . In our opinion, the discreteness of the spectrum actually determines the zero conductance gap of the  $I$ - $V$  curves observed in Refs. [11, 12, 13, 14, 15].

## 3 SEMICONDUCTOR-LIKE BEHAVIOR OF MAGIC METAL CLUSTER

The left and right electrodes (emitter and collector) represent the electron reservoirs. Each reservoir is taken to be in thermal equilibrium. A continuum of states is assumed in reservoirs, occupied according to the Fermi-Dirac distribution

$$f(\epsilon^{e,c} - \mu_0^{e,c}) = \{1 + \exp[(\epsilon^{e,c} - \mu_0^{e,c})/k_B T]\}^{-1}, \quad (2)$$

where  $\mu_0 < 0$  is the electron chemical potential for a semi-infinite metal,  $-\mu_0 = W_0$ ,  $W_0$  is the electron work function ( $W_0 = 5.13$  eV for Au). In all cases energy  $U_0 < \epsilon < 0$  is counted off from the vacuum level,  $U_0 < 0$  is the position of conductivity band of a semi-infinite metal [20].

The electron chemical potential  $\mu^g$  of a granule in a quantum case can be defined by the normalization condition

$$\sum_{p=1}^{\infty} f(\epsilon_p - \mu^g) \equiv \sum_{p=1}^{\infty} \{1 + \exp[(\epsilon_p - \mu^g)/k_B T]\}^{-1} = N_0, \quad (3)$$

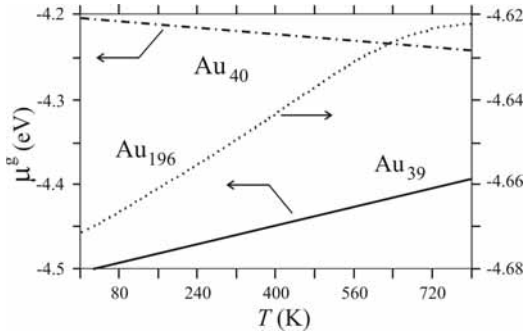


Figure 3 – Temperature dependence of chemical potentials of the neutral spherical gold non-magic ( $Au_{39}$ ) and magic ( $Au_{40}$ ,  $Au_{196}$ ) clusters

where sum runs over all one-electron states,  $N_0$  is the total number of thermalized electrons in a cluster,  $\mu^g \equiv \mu^g(R)$ . If the electron spectrum is known, from Eq. (3) it is possible to calculate  $\mu^g$  of cluster  $Au_{N_0}$  (gold is univalent).

Fig. 3 depicts the chemical potential of some spherical clusters as a function of temperature. Predictably, the dependence is slack and is completely determined by the level hierarchy in dots, and also by the number of electrons.

For the magic clusters as for an intrinsic semiconductor, which has equal numbers of electrons in the conduction band and holes in the valence band, the chemical potential lie halfway between the lowest unoccupied level  $\epsilon^{LU}$  and highest occupied one  $\epsilon^{HO}$ , regardless of the temperature, because each electron promoted to the lowest unoccupied level leaves a hole in the highest occupied term. The Fermi level of non-magic clusters coincides with a real level in a cluster. Calculations show, that the temperature gradient of chemical potential can be both the positive and the negative, and at some temperatures it can change a sign. Similar behavior  $\mu^g(N_0, T)$  for magic clusters  $Na_{N_0}$  have been reported in Ref. [23].

A contact potential difference appears between a cluster and electrodes is

$$\delta\phi = (\mu^g - \mu_0^e)/e. \quad (4)$$

An equilibrium is reached by the charging of a cluster since its capacitance is finite. If  $|\mu^g| < |\mu_0^e|$ , a cluster is charged positively by a charge  $Q_{\text{eff}}^0 = -e(N' - N_0) > 0$ , where  $N'$  is determined by the solution of the Eq. (3) with replacement  $\mu^g \rightarrow -W_0$  for the same spectrum  $\{\epsilon_p\}$  shifted on  $-e\delta\phi$ , according to the Koopmans' theorem [20]. Thus, in thermodynamic limit we have

$$Q_{\text{eff}}^0 = C\delta\phi. \quad (5)$$

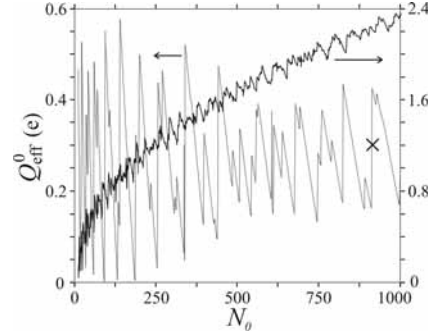


Figure 4 – The calculated size dependence of the residual effective charge  $Q_{\text{eff}}^0$  (5) for the structure  $Au/Au_{N_0}/Au$  based on the clusters of various shape: sphere (dotted line) and disk (solid line). For illustration,  $Q_{\text{eff}}^0$  of magic sphere  $Au_{912}$  is marked as  $\times$

The existence of  $Q_{\text{eff}}^0$  has to do with the transparency of the tunnel barriers before application of voltage. A quasi-classical approximation [24, 25] gives for a metallic sphere of radius  $R$ :  $\mu^g - \mu_0 = \mu_1/R$ ,  $\mu_1 = 1.9 \text{ eV} \times a_0$ ,  $Q_{\text{eff}}^0 = +0.07e$ . In a quantum case, filling levels by electrons we find a highest occupied state,  $\epsilon^{HO} < 0$ , and lowest unoccupied state,  $\epsilon^{LU} < 0$ , counted off from the vacuum level (Fig. 1). Then it is necessary to make a replacement  $\mu^g \rightarrow \epsilon^{HO}$ .

In accordance with our previous results [24],  $|\mu^g| < |\mu_0|$ , therefore the cluster is charged positively before the application of voltage. The size dependence of a charge  $Q_{\text{eff}}^0(N_0)$  for referred gold clusters is demonstrated on Fig. 4. For the above mentioned sizes of spherical clusters,  $Q_{\text{eff}}^0 < e$ . However,  $Q_{\text{eff}}^0$  can accept values larger than  $e$  for the disks of monatomic thickness. Additional charging of the cluster can lead to the Coulomb instability, because the quantity  $Q_{\text{eff}}^0$  is close to a critical charge [20]. Moreover, cluster's anomalous electrostriction is possible as a result of the charging [26].

Otherwise, in the case  $|\mu^g| > |\mu_0|$  (e. g.  $Pb/Au_{N_0}/Pb$ ) the cluster is charged negatively and it is necessary to make a replacement  $\mu^g \rightarrow \epsilon^{LU}$ . Residual effective charge is equal to non-integer elementary charge  $e$ , which is analogous to the charge of cluster in “chemisorption regime”.

The possibility of a fractional charge at tunneling structures was discussed in Ref. [27]. Perhaps, this problem is related to the fractional quantization (or fractional statistics), when the decoupling of the spin and the electron quantum numbers of a charge is important. In the percolation systems it is supposed that the charge at the each granule has a soliton origin. The value of this charge was calculated numerically in Ref. [28].

We consider a central electrode-granule in an external electric field. Between the emitter ( $V = 0$ ) and the

collector the positive voltage  $V$  is applied. In a weak electric field approach we assume, that the ionic subsystem of a granule is not deformed, and the electronic "cloud", generated by the own valence electrons, is deformed only.

#### 4 GRANULE UNDER VOLTAGE $V$

The total energy of a granule is the functional of nonhomogeneous electron concentration,  $\tilde{E}[n(\mathbf{r})]$ . The functional contains a contribution responsible for the interaction of electrons and ions with an external field,

$$e \int [n(\mathbf{r}) - n_i(r)] (\mathbf{E} \cdot \mathbf{r}) d^3r. \quad (6)$$

For simplicity, we suppose that the charge distribution  $n_i(r)$  of the ion subsystem is spherically symmetric.

Let's write down an electron distribution of a granule as

$$n(\mathbf{r}) = n_0(r) + \delta n_1(r) + \delta n_2(\mathbf{r}). \quad (7)$$

Here,  $n_0(r)$  is the electron density of neutral cluster in the absence of the external field,

$$\int n_0(r) d^3r = N_0,$$

$\delta n_1$  is the perturbation arising from the charging of granule,

$$\int \delta n_1(r) d^3r = \Delta N, \quad (8)$$

where  $\Delta N > 0$  and  $\Delta N < 0$  correspond to negatively and positively charged granule, respectively ( $|\Delta N| = N_0$ ).  $\delta n_2(\mathbf{r})$  is the next perturbation arising from the external field which responses for the polarization of a neutral granule,

$$\int \delta n_2(\mathbf{r}) d^3r = 0. \quad (9)$$

We assume, that functions  $n_0(r)$  and  $n_1(r)$  are spherically symmetrical, and  $n_2(\mathbf{r})$  is axially symmetrical. Then one can expand the  $E[n(\mathbf{r})]$  in the functional Taylor series down to the second order of smallness with respect to  $\delta n_1$  and  $\delta n_2$ ,

$$\begin{aligned} \tilde{E}[n(\mathbf{r})] = & \tilde{E}[n_0(\mathbf{r})] + \sum_j \int \frac{\delta \tilde{E}}{\delta n(\mathbf{r})} \delta n_j(\mathbf{r}) d^3r + \\ & + \frac{1}{4} \sum_{j,k} \iint \frac{\delta^2 \tilde{E}}{\delta n_j(\mathbf{r}) \delta n_k(\mathbf{r}')} \delta_j n(\mathbf{r}) \delta_k n(\mathbf{r}') d^3r d^3r' + \dots \end{aligned} \quad (10)$$

Here the functional derivatives are taken at  $n(\mathbf{r}) = n_0(r)$ , and indexes  $j$  and  $k$  runs 1 and 2 according to the definition (7). The zeroth-order expression  $\tilde{E}[n_0(\mathbf{r})] \equiv \tilde{E}_{00}$  is a total energy of a cluster before the

charging ( $\Delta N = 0$ ) and in the absence of the external field ( $E = 0$ ). The functional derivative

$$\delta \tilde{E} / \delta n(\mathbf{r}) = \mu^g + e(\mathbf{E} \cdot \mathbf{r}). \quad (11)$$

In the absence of charging and external field  $-\mu^g \rightarrow W_0$  as  $R \rightarrow \infty$ .

Finally, in the semiclassical approximation (see Appendix B), we get

$$\tilde{E} = \tilde{E}_{00} + \mu^g \Delta N - e \Delta N \eta V + (\Delta N)^2 \tilde{E}_C / 2 - \alpha E^2 / 2. \quad (12)$$

Solving separately the electrostatic problem for the same structure with fraction of the voltage  $\eta V > 0$  (Fig. 1), we obtain

$$\eta = \frac{\epsilon_2 \epsilon_3 (d_e + \epsilon_1 L / 2 \epsilon_2)}{\epsilon_1 \epsilon_2 d_c + \epsilon_1 \epsilon_3 L + \epsilon_2 \epsilon_3 d_e} \equiv \eta^+. \quad (13)$$

Here,  $L \equiv 2R, H$  for a sphere of radius  $R$  and a disk of thickness  $H$ , respectively. We describe the situation for  $\epsilon_1 \equiv \epsilon$ ,  $\epsilon_2 = \epsilon_3 = 1$ . Under Eq. (13) one can find the values  $\eta^+ < 0.65$  and  $\eta^+ < 0.55$  in experiments [11, 12, 13, 14, 15] for the spherical-like and disk-shaped clusters, respectively.

Now we examine the problem of critical surplus charges of a cluster in the presence of an external voltage. For convenience, further we write  $n \equiv \Delta N$ .

#### 5 COULOMB INSTABILITY OF A CLUSTER IN ELECTRIC FIELD

It is necessary to note, that even the vanishing external electric field leads to the instability of a cluster because of the possibility of electron tunneling. We assume, that a cluster relaxed in a metastable state over a period of time which is much smaller, than that between acts of tunneling. As a result of the charging, the intrinsic mechanical stress leads to the Coulomb explosion. This problem was described in Ref. [20] for the solitary spherical cluster in absence of an external electric field. Extending these results, one can write the following expression

$$m \{ (-\mu_{\text{el(ion)}}^g + |e\eta V|) R / e + e/2 \} \quad (14)$$

for the critical electronic or ionic charge in quasi-classical approximation. Here  $\mu_{\text{el(ion)}}^g$  is the size-dependent electron (ion) chemical potential. For the range  $V = (0, 2 \text{ V})$  we have:

1)  $\eta = 1$ . Transitions of electrons between the emitter and the cluster occur more often, than between the cluster and the collector, therefore the electrons are ac-



cumulated on the cluster. In this case their maximal number is

$$n_{\max}; -\mu_{\text{el}}^g R/e^2 + 1/2,$$

where  $\mu_{\text{el}}^g = \mu_0 + \mu_1/R \equiv \mu^g$  and  $n_{\max}; \{+2.5, +6.5\}$  for whole range of sizes  $\{R\}$ .

2)  $\eta \approx 1$ . Transitions of electrons between the cluster and the collector occur more often, than between the cluster and the emitter, therefore on the cluster the deficiency of electrons is observed. Using the definition of ion chemical potential (see Refs. [20, 25]), this number determines as

$$n_{\min} = (\mu_{\text{ion}}^g - |e\eta V|)R/e^2 - 1/2$$

and  $n_{\min} \approx \{-4, -11\}$ .

Similarly, for  $V = (-2, 0 \text{ V})$  we have:

$$1) n_{\min}; \mu_{\text{ion}}^g R/e^2 - 1/2; \{-3.8, -10.6\}.$$

$$2) n_{\max} = (-\mu_{\text{el}}^g + |e\eta V|)R/e^2 + 1/2; \{+3, +8\}.$$

Below, the *whole* numbers  $[n_{\max}]$  and  $[n_{\min}]$  bound the summation in (26). The effect of spectrum quantization can change these numbers no more than in  $\pm 1$  according to (1) (see Ref. [20]).

Effective collision frequency of excited electrons in a cluster is defined as [31]

$$\frac{1}{\tau_e} = \frac{1}{\tau} + \frac{v_F}{R}, \quad v_F = \frac{h}{m r_0} \left(\frac{9\pi}{4}\right)^{1/3}, \quad (15)$$

where  $\tau$  is a relaxation time in the bulk of the metal, caused by electron-electron collisions ( $\tau \times 10^{14} = 6.23 \text{ s}$  for Au at  $T = 75 \text{ K}$  [32]), and  $v_F$  is the electron velocity at the Fermi surface in the bulk. The estimation performed in Ref. [33] gives a preferred electron collision on walls of a dot, therefore  $\tau_e; R/v_F$ . It leads to  $\tau_e \Delta \epsilon_p/h; \{0.52, 0.17\}$ , i. e. to a broadening of levels. We assume that the electron thermalization occurs much faster than acts of tunneling. “New” electrons fill up a number of own electrons, changing their distribution and, accordingly, the chemical potential. This state of the cluster will be a starting state for the next act of tunneling.

## 6 BASIC ENERGY AND KINETIC RELATIONS

We assume, that the *total* energy of all three electrodes  $\tilde{E}$  does not change during the tunneling. In the case of transition of  $\delta N$  electrons from the emitter to the granule (containing  $n$  “surplus” electrons), using Eq. (12) we have

$$\begin{aligned} \delta \tilde{E} &= -\delta N \overset{\rightarrow}{\epsilon}^e + \delta N \epsilon_p + \frac{(-e)^2}{2C} [(n + \delta N)^2 - n^2] - \\ &- e \delta N \eta^+ V = 0. \end{aligned} \quad (16)$$

In this expression we take into account that the  $\delta N$  electrons are ionized from the level  $\overset{\rightarrow}{\epsilon}^e$  on the emitter (whose capacitance is equal to infinity) and then fill up the level  $\epsilon_p$  in a granule with finite capacitance  $C$ .

By analogy with Ref. [34], using Eq. (12) and Eq. (16) for  $\delta N = 1$ , and then taking into account a contact potential difference (4), for emitter-granule transition we have

$$\overset{\rightarrow}{\epsilon}^e = \epsilon'_p + \tilde{E}_C (n + 1/2) - e\eta^+ V, \quad (17)$$

where  $\epsilon'_p \equiv \epsilon_p - e\delta\phi$ . The arrow on the top indicates the energies which are determined by transfers according to Fig. 1. We suppose, that  $n \equiv n(V)$  and  $n = 0$  at  $V = 0$ . However, the granule is charged by the charge  $Q_{\text{eff}}^0$  before voltage applied. Therefore, we assume that  $n$  is the result of the applied voltage only.

For granule-emitter transition we have

$$\overset{\rightarrow}{\epsilon}^e = \epsilon'_p + \tilde{E}_C (n - 1/2) - e\eta^+ V. \quad (18)$$

Similarly, for the granule-collector and collectorgranule transitions we have

$$\overset{\leftarrow}{\epsilon}^c = \epsilon'_p + \tilde{E}_C (n \mp 1/2) + e(1 - \eta^+) V. \quad (19)$$

Here the upper/under arrows at the left correspond to the following signs on the right. Independently of  $n$  the relation

$$\overset{\rightarrow}{\epsilon}^e - \overset{\leftarrow}{\epsilon}^c = \tilde{E}_C = \overset{\leftarrow}{\epsilon}^c - \overset{\rightarrow}{\epsilon}^e$$

takes place. It agrees with well-known quasi-classical relation for the ionization potential and electron affinity,

$$IP - EA = \tilde{E}_C$$

(see, e. g. Ref. [20]). Thus, Eqs. (17)–(19) represent a golden rule approximation.

The tunneling of a single electron through barriers is determined by the tunnel rates  $\Gamma^{e,c}$ , which depend on the junction geometry and the voltage fraction  $\eta$ . In general, their evaluation is far from a trivial problem [2, 27]. We assume that  $\Gamma^{e,c}$  are small and the temperature is not too low, i. e.  $k_B T > h(\Gamma^e + \Gamma^c) = \min\{\Delta \epsilon_p, \tilde{E}_C\}$ .

By analogy with the theory of Ref. [29], we introduce the partial tunneling streams from electrodes to a granule

$$\vec{\omega}_n^e = 2 \sum_p \Gamma(\vec{\epsilon}^e) f(\vec{\epsilon}^e - \mu_V^e) [1 - f(\vec{\epsilon}^e - \mu_C^e)], \quad (20)$$

$$\overleftarrow{\omega}_n^e = 2 \sum_p \Gamma(\overleftarrow{\epsilon}^e) f(\overleftarrow{\epsilon}^e - \mu_V^e) [1 - f(\overleftarrow{\epsilon}^e - \mu_C^e)], \quad (21)$$

and from a granule to the electrodes

$$\overleftarrow{\omega}_n^e = 2 \sum_p \Gamma(\overleftarrow{\epsilon}^e) [1 - f(\overleftarrow{\epsilon}^e - \mu_V^e)] f(\overleftarrow{\epsilon}^e - \mu_C^e), \quad (22)$$

$$\vec{\omega}_n^e = 2 \sum_p \Gamma(\vec{\epsilon}^e) [1 - f(\vec{\epsilon}^e - \mu_V^e)] f(\vec{\epsilon}^e - \mu_C^e), \quad (23)$$

where the factor 2 takes into account the spin degeneration of levels in electrodes. In view of the applied voltage (and charging of a granule) the spectrums (see Eqs. (17)–(19)) and the chemical potentials are shifted in distributions (2) and (3),

$$\begin{aligned} -\mu_V^e &\equiv W_0, \quad \mu_C^e = \mu^g - e\delta\phi + \tilde{E}_C(n \mp 1/2) - e\eta^+V, \\ \mu_C^e &= \mu - e\delta\phi + \tilde{E}_C(n \pm 1/2) - e(1 - \eta^+)V, \\ \mu_V^e &= \mu_0^e - eV. \end{aligned}$$

As the first approximation of the perturbation theory [20], for small  $V$ ,  $\mu^g$  is determined not only by the formal shift of the well depth, but also by the number of conduction electrons in the granule ( $N = N_0 + n_q$ ,  $n_q = n + [Q_{\text{eff}}^0]/e$ ). The use of the chemical potentials is correct in a quasi-equilibrium state, i. e. when the intervals between acts of tunneling are much longer than the relaxation time. It is also supposed, that the external electric field and the Coulomb blockade do not remove degeneration of levels.

Let's denote the total electron streams from/to leads into/out the cluster, as

$$\omega_n^{\text{in}} = \vec{\omega}_n^e + \overleftarrow{\omega}_n^e, \quad \omega_n^{\text{out}} = \overleftarrow{\omega}_n^e + \vec{\omega}_n^e.$$

In the limit of weak tunneling, the probability  $P_n$  of the finding of  $n$  above mentioned electrons at central electrode is defined by the master equation in the stationary limit

$$\begin{aligned} \dot{P}_n &= \omega_{n+1}^{\text{out}} P_{n+1} + \omega_{n-1}^{\text{in}} P_{n-1} - \\ &- (\omega_n^{\text{in}} + \omega_n^{\text{out}}) P_n = 0. \end{aligned} \quad (24)$$

The requirement of the stationarity gives the recurrent relation

$$P_{n+1} = P_n \frac{\omega_n^{\text{in}}}{\omega_{n+1}^{\text{out}}}. \quad (25)$$

The dc current flowing through a metallic quantum dot (with restriction on its instability (14)), is determined as

$$I = -e \sum_{n_{\text{min}} < 0}^{n_{\text{max}} > 0} P_n (\vec{\omega}_n^e - \overleftarrow{\omega}_n^e) = -e \sum_{n_{\text{min}} < 0}^{n_{\text{max}} > 0} P_n (\overleftarrow{\omega}_n^e - \vec{\omega}_n^e). \quad (26)$$

Let's consider the case of "strong quantization" for electron spectrum:

$$\Delta\epsilon_p \gg E_C.$$

This regime is hypothetically reached by a significant increase of the cluster capacitance (the cluster shape must be changed to the needle-like or disk-like one under the condition that its volume is fixed (see, e.g. Ref. [20]). Thus the charge  $Q_{\text{eff}}^0$  in (5), which provides a contact potential difference, is proportional to the capacitance and can have large magnitude. When the voltage is applied, the charge, which is caused by the transferring surplus electrons, is much less than  $Q_{\text{eff}}^0$ . Therefore it insignificantly have influence on the cluster energetics. In reality, the inequality  $\Delta\epsilon_p \gg E_C$  is not possible even for the long atomic chain [20]. Nevertheless, this case is useful from the methodical point of view to analyze the current gap of  $I - V$  characteristics.

As an assumption, we use the fixed tunnel rates at the Fermi level in the emitter. It is correct for the small voltages,  $eV = W_0$ . Neglecting in (17)–(19) terms  $\sim \tilde{E}_C$ , it is easy to obtain the result, similar to Ref. [29]:

$$I = I_0 \sum_p [f(\epsilon^e - \mu_0^e) - f(\epsilon^e - \mu_V^e)], \quad (27)$$

where  $I_0 = 2e\Gamma^e\Gamma^c/(\Gamma^e + \Gamma^c)$ .

The expressions in this section are written down for  $V > 0$ . In the case  $V < 0$ , the  $I(V)$  can be easily received, if we set  $V = 0$  on a collector and  $V > 0$  on the emitter, and use  $\eta^- = 1 - \eta^+$ .

In the general case, for calculation of  $I - V$  (26) it is necessary to know probabilities  $P_n$ . Their statistical determination is a complicated problem [35, 36]. In the experiments, the size of the cluster and its location are known only approximately, therefore detailed calculations of  $P_n$  are not suitable. Using the recurrent relations we can find the ratios  $P_{n \neq 0}/P_0$ .

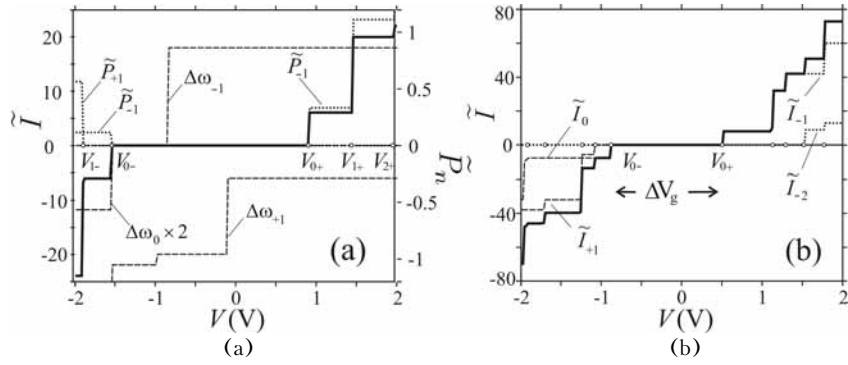


Figure 5 – The current-voltage curves (solid lines) and its components, calculated from Eq. (26) ( $\beta = 1$ ,  $\eta^+ = 0.1$ ,  $T = 30$  K).  $\Delta\omega_n(V)$  is given in  $\Gamma^e$  units:

(a) Au/Au<sub>40</sub>/Au; (b) Au/Au<sub>100</sub>/Au

## 7 APPLICATION AND DISCUSSION

### 7.1 Thermal equilibrium

Setting the collector-granule distance  $d_c$ , parameter  $\beta = \Gamma^e/\Gamma^c$  and using the recurrent relation (25) for Eq. (26), it is possible to calculate the reduced dc current  $\tilde{I} \equiv I/(eP_0\Gamma^e)$ . We do not evaluate separately the threshold voltages, in our scheme it appears automatically.

The results of calculations of the  $I-V$  characteristics for the structures Au/Au <sub>$N_0$</sub> /Au, based on spherical clusters, are presented in Fig. 5. For completeness of analysis, the voltage behavior of the reduced probabilities  $P_n(V) \equiv P_n/P_0$  and the difference of electron streams  $\Delta\omega_n = \overrightarrow{\omega}_n^e - \overleftarrow{\omega}_n^e$  are given also.

The current jumps are stipulated by the jumps of  $\tilde{P}_n(V)$  and  $\Delta\omega_n(V)$ , because the current is formed by their product. As one can surmise, the jump of probability  $\tilde{P}_{-1}(V)$  causes the current jump in the threshold voltage  $V_{0+}$ .

Making use of the equality  $\tilde{I} \equiv \sum_n \tilde{I}_n(V)$  in accordance with Eq. (26) one can fix also the “threshold” values of  $n$ . As is seen from Fig. 5 (b), the role of partial current components  $I_n$  (with  $|n| > 1$ ) grow with increasing  $N_0$ . The charging leads to energy shift of spectrum according to Eqs. (17)–(19). Thus the different parts of a spectrum are involved during tunneling.

The current gap width  $\Delta V_g$  for all structures is determined by values  $n = 0, -1$ . The probability  $P_{-1}$  prevails over  $P_{+1}$ , because the “granule-collector” electron stream is more than a “emitter-granule” one and the granule is charged positively (i. e.  $n < 0$ ). For low temperatures ( $k_B T_{\text{eff}}^g = \Delta\epsilon_F$ ), the current gap width  $\Delta V_g = V_{0+} + |V_{0-}|$  is determined analytically by the conductance gap boundaries  $V_{0-}$  and  $V_{0+}$ . For example,  $V_{0+}$  is defined from the condition of absence of

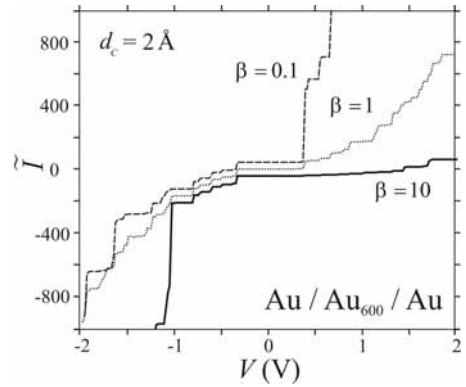


Figure 6 – Calculated  $I-V$  curves at  $T = 30$  K for structure based on spherical clusters. For presentation the curves are shifted slightly on a vertical

collector current of the direct  $\tilde{I}-V$  curve branch ( $V > 0$ ), and finally we have:

$$\Delta V_g = \left( \frac{1}{2e} \tilde{E}_C + \frac{1}{e} \Delta\epsilon \right) \left[ \frac{1}{2 - \eta^+} + \frac{1}{2 - \eta^-} \right], \quad (28)$$

where  $\Delta\epsilon \equiv \mu^g - \epsilon^{\text{HO}} \geq 0$  at  $T = 0$ . Calculated values of  $\Delta V_g$  are in a good agreement with the experimental values based both on spherical and disc-shape clusters. For large granules  $\Delta V_g \rightarrow 0$  as  $R \rightarrow \infty$ .

Within the applied voltage the  $\tilde{I}-V$  characteristics versus  $\eta^+$  are shifted to the right and the gap width decreases a little. The calculated  $\tilde{I}-V$  curves of the structure Au/Au<sub>600</sub>/Au for fixed  $\eta^+$  ( $d_c = 2$  Å) and different  $\beta$  are shown in Fig. 6. The current gap is practically independent on  $\beta$ , however, the current jumps are strongly dependent on the value of  $\beta$ , which, in its turn, has no influence on threshold voltages.

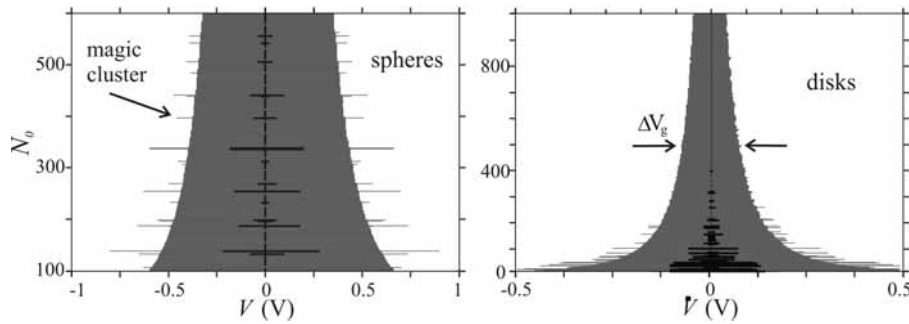


Figure 7 – The current gaps vs  $N_0$  calculated from Eq. (26) ( $d_c = 2 \text{ \AA}$  and  $\beta = 10$ ). Solid lines show the gaps calculated from Eq. (27) for the case of “strong quantization”. For presentation the gaps are placed on a vertical

In order to illustrate our results, in Fig. 7, we compare the size dependences  $\Delta V_g(N_0)$  calculated from Eqs. (26) and (27) for spheres and disks. The largest quantities  $\Delta V_g$  correspond to the magic granules, for which  $\Delta \epsilon \neq 0$ . For the case of “strong quantization” the size of current gap for non-magic clusters is equal to zero explicitly, because the emitter Fermi level is in line to the closed levels in cluster. Calculations demonstrate the non-monotonic dependence  $\Delta V_g(N_0)$ . These results shows also, that a charging leads to the growth of a gap.

The actual forms of dependence  $\Delta V_g(d_c)$  for the structure based on magic disk  $\text{Au}_{178}$  ( $R \approx 35 \text{ \AA}$ ) are plotted in Fig. 8. In experiments [14, 15] the gap varied as  $0.8 \rightarrow 0.4 \rightarrow 0.7 \text{ V}$  for the cyclic variation of  $d_c \approx 1 \rightarrow 2 \rightarrow 1 \text{ \AA}$ . The reasons of such numerical distinction of our results, apparently consists in neglecting the role of nonlinearities in the strong electric field and in energy dependence of tunneling rates. At high rates the capacitance ceases to be classical and can strongly grow ( $E_C \rightarrow 0$ ) [37, 38], showing non-monotonic dependence from  $\Gamma^c$ . It means, that in a reality we deal with the intermediate cases (between limiting estimations from Eqs. (26) and (27) in Fig. 8).

Let’s discuss other features of the tunnel structure. In spite of the fact that the emitter and a collector are made of one material, the chemical potentials of electrons are not equal to each other: the emitter is represented by a thick film of Au (111), and a collector is a polycrystal of Au. Their work functions are different [39]. Except for it the emitter is covered with a dielectric film, that also influences a electron work function. We can estimate this contribution.

Proceeding from indirect measurements [40], the work function decreases with growth dielectric constant  $\epsilon$  of coating. The calculations of the electron work function  $W_d$  for cylindrical nanowires in a dielectric confinement are done in Ref. [41]:  $W_d$  decreases approximately on 20 % at magnitude as  $\epsilon$  rise from 1

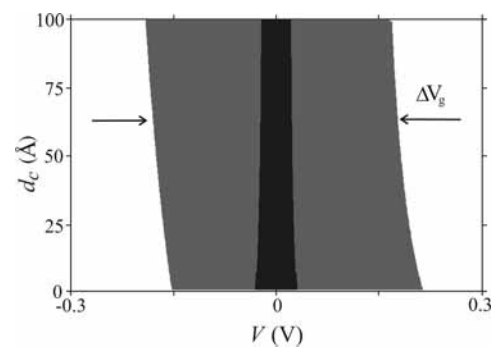


Figure 8 – The current gap vs  $d_c$  calculated from Eqs. (26) and (27) for magic disk  $\text{Au}_{178}$ ,  $\beta = 10$

to 4. The basic contribution thus can be related to the change of electrostatic dipole barrier which contribution to a work function of system gold-vacuum makes up to 30 % [39]. Hence, this contribution also makes a upper limit of the change of  $W_d$  for metal-dielectric-vacuum system. Owing to  $W_d < W_0$  the inequality  $|\mu^g| > W_d$  is possible, that can lead to negative charging of the cluster before the application of voltage.

## 7.2 A role of hot electrons

The consequence of the phonon spectrum deformation of granules is the weakening of the electron-phonon interaction within them:  $v_F/R \gg \omega_D$ , where  $v_F$  is the electron velocity at the Fermi surface in the bulk, and  $\omega_D$  is the Debye frequency. This interaction can be so suppressed that the electron-electron interaction becomes the main mechanism for the dissipation of the energy, which is injected to the particle. This additional energy results in the overheating of the electron subsystem, which is described by the Fermi statistics with some effective (enhanced) temperature  $T_{\text{eff}}^g$ , and the temperature of the ion subsystem only slightly changes [42, 7, 8, 9]. With the increase of the bias voltage  $V$ , the number of electrons, relaxing in the granule, increases significantly.

Among them are all the electrons with energies in the interval  $e\eta V$  below the Fermi level of the granule ( $\eta V$  is the fraction of the bias voltage on the granule), since the “flow” of tunneling electrons increases from below lying levels, thereby, involving large number of conductivity electrons to the relaxation process. At the same time, channels of losses appear, which are related to the generation of holes on the occupied levels and their subsequent recombination. The granule does not fragmentize at the significant overheating of the electron subsystem, because the  $I-V$  curves are reproduced at the cyclic changes of the bias voltage [14, 13].

The estimate of the energy, which is pumped by the conductivity electrons to the granules of discontinuous films, is given in Ref. [7] ( $\sim 0.2, 0.3$  eV). This means that the experiments [14, 13] correspond to the Coulomb blockade regime in the region of current gap at the whole diapason of  $R$  and reasonable values of  $T_{\text{eff}}^g$ . Also, the quantum ladder can be smeared out by the thermal fluctuations,

$$E_C > \Delta\varepsilon_F > k_B T^{e,c,g},$$

where  $\Delta\varepsilon_F$  is the difference between discrete levels in the vicinity of the granule Fermi level, and  $\Delta\varepsilon_F = \Delta\varepsilon_p$  for magic clusters at  $T = 0$  (see Fig. 2). We represent the emitter and the collector as the electron reservoirs with continuum spectrums and temperatures  $T^{e,c}$  equal to the thermostat one. The spectrum of states is calculated in advance and, therefore, the chemical potential of neutral  $\text{Au}_{N_0}$  granules and its temperature dependence can be found from equation (3) at a given temperature  $T_{\text{eff}}^g$ .

For the comparison with the results of Refs. [14, 13], the calculations are done for three temperatures of the collector and emitter  $T^e = T^c = 5, 30, 300$  K, and also  $T_{\text{eff}}^g = T^e, 2000$  K. The values  $\Gamma^c = \Gamma^e = 1$  and  $\eta = 1/2$  are used for all cases.

Fig. 9 shows calculated  $I-V$  curves for disc of radius  $R = 20$  Å (magic cluster  $\text{Au}_{230}$ ) and sphere of radius  $R = 10$  Å (magic cluster  $\text{Au}_{256}$ ). The calculation of the  $\tilde{I}-V$  curves and current gap can be done only numerically at  $k_B T_{\text{eff}}^g \geq \Delta\varepsilon_F$ , when the larger part of the spectrum, compared to  $\Delta\varepsilon_F$ , is responsible for the charge transfer. Our calculations show an evident dependence of  $\tilde{I}-V$  curves flatness on the electron subsystem temperature.

However, in order to obtain an agreement with observed  $I-V$  curves it is necessary to suggest that electrons in the emitter and collector are also heated up to some effective temperature, which is higher than the thermostat one. It is possible, because electrons (the current  $I = 1$  pA is provided by  $I/e \sim 10^6$  number of electrons per second) relax in generally on the free path length in the last electrodes.

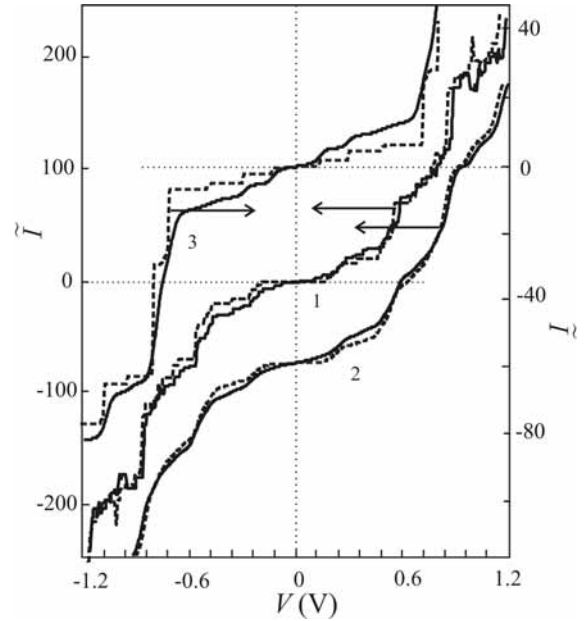


Figure 9 – Calculated  $\tilde{I}-V$  curves of structure based on magic clusters: disc  $\text{Au}_{230}$  and sphere  $\text{Au}_{256}$

- 1 –  $\text{Au}_{230}$ : dotted curve –  $T^{e,c,g} = 5$  K,  
solid –  $T^{e,c} = 5$  and  $T^{e,c,g} = 300$  K
- 2 –  $\text{Au}_{230}$ : dotted curve –  $T^{e,c,g} = 300$  K,  
solid –  $T^{e,c} = 300$  and  $T_{\text{eff}}^g = 2000$  K
- 3 –  $\text{Au}_{256}$ : dotted curve –  $T^{e,c} = 30$  K and  $T_{\text{eff}}^g = 2000$  K,  
solid –  $T_{\text{eff}}^e = 300$  and  $T_{\text{eff}}^g = 2000$  K

For the illustration, we present our result at Fig. 9 for the sphere at  $T^{e,c} = 300$  K and  $T_{\text{eff}}^g = 2000$  K. Only by such a way, we can explain the flattening of the  $I-V$  curves for the metallic cluster structures at low thermostat temperatures. With the increase of the bias voltage, the current flow is accompanied by the increase of the electron gas temperature.

## 8 SUMMARY

In the framework of the particle-in-a-box model for the spherical and disk-shaped gold clusters, the electron spectrum was computed. In this model, the work function of clusters is smaller than that of semi-infinite gold electrodes. It resulted in the appearance of a contact potential difference between a cluster and electrodes. Residual effective charge is equal to non-integer elementary charge  $e$ , which is analogous to the charge of cluster in “chemisorption regime”. For the small spherical clusters, positive charge is less than  $e$ . However, this charge can accept a values larger than  $e$  for the disks of monatomic thickness. Additional charging of the clusters can lead to the Coulomb instability, because it is close to a critical charge. The charging results in the energy shift of the spectrum.

The current-voltage characteristics were analyzed taking into account the contact potential difference. For single-electron molecular transistors based on the small gold clusters the current gap and its voltage asymmetry were computed. The largest gaps correspond to the magic granules. For the case of “strong quantization”, the size of a current gap for non-magic clusters is equal to zero explicitly, because the emitter Fermi level is in line to the closed levels in cluster. We derive the simple expression for the size dependence of the current gap width, which taking into account the Coulomb blockade and discreteness of the energy levels. The results shows, that a charging leads to the growth of a gap.

In conclusion, we have calculated the  $I-V$  characteristics of structure based on magic clusters: disc  $\text{Au}_{230}$  and sphere  $\text{Au}_{256}$ . We have suggested that the overheating of electron subsystem leads to the disappearance of current gap and significant flattening of current-voltage curves. Our results are in a good qualitative agreement with experiment of Ref. [14, 13].

#### APPENDIX A. ELECTRON SPECTRUM IN CYLINDRICAL-LIKE CLUSTERS

As an approximation, the profile of the one-electron effective potential in the cluster can be represented as a potential well of the depth  $U_0 < 0$ . The three-dimensional Schrödinger equation for a quantum box can be separated to the one-dimensional equations. The spectrum of wave numbers in a spherical and cylindrical potential wells are determined from the continuity condition of a logarithmic derivative of the wave function on the boundaries. For a disk of radius  $R$  and thickness  $H$  it is necessary to solve numerically the equation:

$$k_{nm} \frac{I'_m(k_{nm}R)}{I_m(k_{nm}R)} = \kappa_{nm} \frac{K'_m(\kappa_{nm}R)}{K_m(\kappa_{nm}R)}. \quad (\text{A.1})$$

Here  $I_m$  is the Bessel function,  $K_m$  is the McDonald function, the stroke denotes a derivative over an argument,  $k_{nm} = \sqrt{k_0^2 - \kappa_{nm}^2}$ ,  $\kappa_{nm} = \sqrt{k_0^2 - \kappa_{nm}^2}$ ,  $\hbar k_0 = \sqrt{2m_e|U_0|}$ , and  $m_e$  is the electron mass. The  $n = 1, 2, 3, \dots$  number the roots of the Eq. (29) for the fixed  $m = 0, \pm 1, \pm 2, \dots$

Quantization of the wave vector  $k_s$  along the cylinder axis is determined by the solution of the equation:

$$k_s H = s\pi - 2\arcsin(k_s/k_0),$$

where  $s$  is the integer number. Neglecting the area near cylinder edges, the energy spectrum is calculated by a simple way as follows

$$\varepsilon_{nms} = U_0 + \frac{\hbar^2}{2m_e}(k_{nm}^2 + k_s^2).$$

In addition to the spin degeneration, there is a double degeneration with respect to the sign of index  $m$ , since  $k_{n,m} = k_{n,-m}$ . Further, the spectrum of cluster is denoted as  $\varepsilon_p$ ,  $p = 1, 2, 3, \dots$  is the number of *one-electron state*. All levels are numbered in order to increase energies.

#### APPENDIX B. ENERGY OF CLUSTER IN EXTERNAL ELECTRICAL FIELD

Using spherical coordinates, we remove the center point  $z = 0$  from emitter in a center of a granule, and we direct a  $z$  axis from a collector to the emitter under the conservation of the potential difference between them. Then an electric field  $E = |E|z$ , where  $z$  is a unit vector along an  $z$  axis.

As the surplus charge is effectively distributed over a surface, it is quite reasonable for estimation to use the form

$$\delta n_1(r) = \begin{cases} 0, & 0 < r < R - b, \\ \tilde{n}, & R - b < r < R, \\ 0, & R > 0. \end{cases} \quad (\text{B.1})$$

Using condition (8), we obtain the following expression for surface concentration

$$\tilde{n} = \frac{\Delta N}{\Omega(3\xi - 3\xi^2 + \xi^3)},$$

where  $\Omega = 4\pi R^3/3$ , and  $\xi = b/R$  is the small parameter.

Then, we use the linear response approach (see, e. g. Ref. [30])

$$\delta n_2(r, \theta) = Y(r)|E|\cos\theta. \quad (\text{B.2})$$

The spherically symmetric function  $Y(r)$  in (31) are determined from the normalization condition (9) and a global minimum of the functional,  $\delta \tilde{E}[n(r)] \rightarrow 0$ .

One of the terms, interesting for us, is

$$-e \int \delta n_1(r) \varphi(z) d^3r,$$

where  $\varphi$  is an external electrostatic potential. In the case of  $V > 0$  and vacuum collector-emitter space  $\varphi(z) = V(z - d_e - L/2)/d$ , where  $d = d_e + L + d_c$ . After the integration in spherical coordinates, the term, which is proportional to  $z$ , vanishes, and, as a result, we have  $-e\Delta N\eta V$ ,  $\eta$  is a fraction of a voltage.

Other three terms

$$\int \frac{\delta n_1(\mathbf{r})\delta n_2(\mathbf{r}') + \delta n_1(\mathbf{r})\delta n_1(\mathbf{r}') + \delta n_2(\mathbf{r})\delta n_2(\mathbf{r}')}{|\mathbf{r} - \mathbf{r}'|} d^3r d^3r'. \quad (\text{B.3})$$

give a basic contribution to the second order of expansion (10). The first integral in Eq. (32) for the functions (30) and (31) vanishes after the integration on corners and second one equals

$$(\Delta N) \frac{2E_C}{2} \left( 1 + \frac{1}{3}\xi + O(\xi^2) \right).$$

The third integral was calculated earlier for definition of the polarizability of a cluster  $\alpha = -(4\pi/3) \int_0^\infty Y(r)r^3 dr \equiv R_{\text{eff}}^3 \approx R^3$  [30]. Finally, using  $\xi \rightarrow 0$ , we obtain Eq. (12).

## REFERENCES

1. Likharev, K. K. SET: Coulomb blockade devices [Text] / K. K. Likharev // Nano et Micro Technologies. – 2003. – V. 3. – P. 71–114.
2. Von Delft, J. Spectroscopy of discrete energy levels in ultrasmall metallic grains [Text] / J. von Delft, D. C. Ralph // Phys. Rep. – 2001. – V. 345. – № 1. – P. 61.
3. Aleiner, I. L. Quantum Effects in Coulomb Blockade [Text] / I. L. Aleiner, P. W. Brouwer, L. I. Glazman // Phys. Rep. – 2002. – V. 358. – P. 309.
4. Otero, R. Observation of preferred heights in Pb nanoislands: A quantum size effect [Text] / R. Otero, A. L. Vazquez de Parga, R. Miranda // Phys. Rev. B. – 2002. – V. 66. – id. 115401.
5. Semrau, S. Designable electron transport features in one-dimensional arrays of metallic nanoparticles: Monte Carlo study of the relation between shape and transport [Text] / S. Semrau, H. Schoeller, W. Wenzel // Phys. Rev. B. – 2005. – V. 72. – id. 205443.
6. Pogosov, V. V. Introduction to Physics of Charged and Size Effects: Surface, Clusters, and Low-Dimensional Systems [Text] / V. V. Pogosov. – Moscow: Fizmatlit, 2006. – 328 p.
7. Fedorovich, R. D. Electron and light emission from island metal films and generation of hot electrons in nanoparticles [Text] / R. D. Fedorovich, A. G. Naumovets, P. M. Tomchuk // Physics Reports, 2000. – V. 328. – P.73-79.
8. D'Agosta, R. Local Electron Heating in Nanoscale Conductors [Text] / R. D'Agosta, Na Sai, Di Ventra Massimiliano // Nanoletters. – 2006. – V. 6. – № 12. – P. 2935–2938.
9. Galperin, M. Heat conduction in molecular transport junctions [Text] / M. Galperin, M. Ratner, A. Nitzan // Phys. Rev. B. – 2007. – V. 75. – id. 155312.
10. Millo, O. Charging and quantum size effects in tunneling and optical spectroscopy of CdSe nanorods [Text] / O. Millo, D. Katz, D. Steiner et al. // Nanotechnology. – 2004. – V. 15. – P. 1–6.
11. Ohgi, T. Charging effects in gold nanoclusters grown on octanedithiol layers [Text] / T. Ohgi, H. Y. Sheng, Z. C. Dong, H. Nejo, D. Fujita // Appl. Phys. Lett. – 2001. – V. 79. – P. 2453.
12. Ohgi, T. Consistent size dependency of core-level binding energy sifts and single electron tunneling effects in supported Au nanoclusters [Text] / T. Ohgi, D. Fujita // Phys. Rev. B. – 2002. – V. 66. – id. 115410.
13. Ohgi, T. Capacitance dependence of chemical potential distribution in supported nanoclusters [Text] / T. Ohgi, Y. Sakotsubo, D. Fujita, Y. Ootuka // Surface Science. – 2004. – V. 402. – P. 566–568.
14. Wang, B. Single-electron tunneling study of two-dimensional gold clusters [Text] / B. Wang, X. Xiao, X. Huang // Appl. Phys. Lett. – 2000. – V. 77. – № 8. – P. 1179.
15. Hou, J. G. Nonclassical behavior in the capacitance of a nanojunction [Text] / J. G. Hou, B. Wang, J. Yang // Phys. Rev. Lett. – 2001. – V. 86. – № 13. – P. 5321.
16. Boyen, H.-G. Alloy formation of supported gold nanoparticles at their transition from clusters to solids: Does size matter? [Text] / H.-G. Boyen, A. Ethirajan, G. Kastle et al. // Phys. Rev. Lett. – 2005. – V. 94. – id. 016804.
17. Hou, J. G. Detecting and manipulating single molecules with STM [Text] / J. G. Hou, A. Zhao // NANO: Brief Rep. Rev. – 2006. – V. 1. – P. 15.
18. Pogosov, V. Effects of charging and tunnelling in a structure based on magic and non-magic metal clusters [Text] / V. Pogosov, E. Vasyutin // Nanotechnology. – 2006. – V. 17. – P. 3366–3374.
19. Soldatov, E. S. Room temperature molecular single-electron transistor [Text] / E. S. Soldatov, V. V. Khanin, A. S. Trifonov, S. P. Gubin, V. V. Kolesov, D.E. Presnov, S. A. Iakovenko, G. V. Khomutov, A. N. Korotkov // Physics-Uspekhi. – 1998. – V. 41. – P. 202.
20. Pogosov, V. Energetics of metal slabs and clusters: the rectangle-box model [Text] / V. Pogosov, V. Kurbatsky, E. Vasyutin // Phys. Rev. B. – 2005. – V. 71. – id.195410.
21. Brack, M. Metal clusters and magic numbers [Text] / M. Brack // Sci. Amer. – 1997. – V. 277. – P. 30.
22. Gubin, S. P. Molecular clusters as building blocks for nanoelectronics: the rst demonstration of a cluster single-electron tunnelling transistor at room temperature [Text] / S. Gubin, Y. Gulayev, G. Khomutov et al. // Nanotechnology. – 2002. – V. 13. – № 1. – P. 185–194.
23. Kurkina, L. I. Electron structure and chemical potential of small jellium clusters at nonzero temperatures [Text] / L. I. Kurkina, O. V. Farberovich // Sol. St. Commun. – 1996. – V. 98. – P. 469.
24. Pogosov, V. V. Sum-rules and energy characteristics of small metal particle [Text] / V. V. Pogosov // Sol. St. Commun. – 1990. – V. 75. – P. 469.
25. Kiejna, A. On the temperature dependence of the ionization potential of self-compressed solid- and liquid-metallic clusters [Text] / A. Kiejna, V. V. Pogosov // J. Phys.: Condens. Matter. – 1996. – V. 8, № 23. – P. 4245.
26. Pogosov, V. V. On some tenzoemission effects of the small metal particles [Text] / V. V. Pogosov // Sol. St. Commun. – 1992. – V. 81. – P. 129.
27. Azbel', M. Ya. Time, tunneling and turbulence [Text] / M. Ya. Azbel' // Phys. Uspekhi. – 1998. – V. 41. – P. 543.
28. Parthasarathy, R. Percolating through networks of random thresholds: Finite temperature electron tunneling in metal nanocrystal arrays [Text] / R. Parthasarathy, X.-M. Lin, K. Elteto et al. // Phys. Rev. Lett. – 2004. – V. 92, № 7. – id.076801.
29. Averin, D. Theory of single-electron charging of quantum wells and dots [Text] / D. V. Averin, A. N. Korotkov, K. K. Likharev // Phys. Rev. B. – 1991. – V. 44. – № 12. – P. 6199.
30. Snider, D. R. Density-functional calculation of the static electronic polarizability of a small metal sphere [Text] / D. R. Snider, R. S. Sorbello // Phys. Rev. B. – 1983. – V. 28. – № 10. – P. 5702.
31. Wood, D. M. Quantum size effects in the optical properties of small metallic particles [Text] / D. M. Wood, N. W. Ashcroft // Phys. Rev. B. – 1982. – V. 25. – № 10. – P. 6255–6273.
32. Sokolov, A. V. Optical Properties of Metals [Text] / A. V. Sokolov. – New York : American Elsevier, 1967. – 311 p.
33. Kurbatsky, V. P. Optical low-frequency absorption of small metal particles [Text] / V. P. Kurbatsky, V. V. Pogosov // Techn. Phys. Lett. – 2000. – V. 26. – P. 1020.
34. Beenakker, C. W. J. Theory of Coulomb-blockade oscillations in the conductance of a quantum dot [Text] / C.W. J. Beenakker // Phys. Rev. B. – 1991. – V. 44. – № 4. – P. 1646.

35. Brack, M. Thermal properties of the valence electrons in alkali metal clusters [Text] / M. Brack, O. Genzken, K. Hansen // Z. Phys. D. – 1991. – V. 21. – P. 65.
36. Petrov, E. G. Kinetic rectification of charge transmission through a single molecule [Text] / E. G. Petrov, V. May, P. Hänggi // Phys. Rev. B. – 2006. – V. 73. – id.045408.
37. Wang, J. Capacitance of atomic junctions [Text] / J. Wang, H. Guo, J.-L. Mozos et al. // Phys. Rev. Lett. – 1998. – V. 80. – P. 4277.
38. König, J. Strong tunneling in the single-electron box [Text] / J. König, H. Schoeller // Phys. Rev. Lett. – 1998. – V. 81. – P. 3511.
39. Pogosov, V. V. Effect of deformation on surface characteristics of finite metallic crystals [Text] / V. V. Pogosov, O. M. Shtepa // Ukr. Phys. J. – 2002. – V. 47. – № 11. – P. 1065.
40. Modinos, A. Field, Thermionic and Secondary Electron Emission Spectroscopy [Text] / A. Modinos. – New York : Plenum Press, 1984. – 375 p.
41. Smogunov, A. N. Electronic structure of simple metal whiskers [Text] / A. N. Smogunov, L. I. Kurkina, S. I. Kurganskii, O. V. Farberovich // Surf. Sci. – 1997. – V. 391. – P. 245.
42. Shklovskii, V. A. The role of electrons of conductivity in forming of thermal resistance at border metal - dielectric [Text] / V. A. Shklovskii // Let. J. Exper. Theor. Phys. – 1977. – V. 26. – P. 679.

Надійшла 02.02.2009

*Теоретически исследованы эффекты зарядки и одноэлектронного тунелирования в структуре на кластере. В рамках модели бесконечной потенциальной ямы для сферических и дискообразных золотых кластеров вычислены электронный спектр и температурная зависимость химического потенциала. Разница между химическими потенциалами массивных электродов и островков приводит к зарядке последних. Мы показываем, что эффективный остаточный заряд не равен целому значению заряда электрона  $e$  и зависит от формы кластера. Уравнения для анализа вольт-амперной характеристики используются с учетом ограничений, связанных с кулоновской неустойчивостью кластера. Для одноэлектронных молекулярных транзисторов вычислены немономонные размерные зависимости токовой щели и ее асимметрия по напряжению. Мы предполагаем, что перегрев электронной подсистемы приводит к исчезновению токовой щели и постепенному сглаживанию вольт-амперных характеристик, что наблюдается в экспериментах.*

*Теоретично досліджені ефекти зарядки й одноелектронного тунелювання в структурі на кластері. У рамках моделі нескінченної потенційної ями для сферичних і дискообразних золотих кластерів обчислені електронний спектр і температурна залежність хімічного потенціалу. Різниця між хімічними потенціалами масивних електродів і островців приводить до зарядки останніх. Ми показуємо, що ефективний залишковий заряд не дорівнює цілому значенню заряду електрона  $e$  й залежить від форми кластера. Рівняння для аналізу вольт-амперної характеристики використовуються з урахуванням обмежень, пов'язаних з кулоновською нестійкістю кластера. Для одноелектронних молекулярних транзисторів обчислені немономонні розмірні залежності струмової щілини і її асиметрія по напрузі. Ми припускаємо, що перегрів електронної підсистеми приводить до зникнення струмової щілини й поступового згладжування вольт-амперних характеристик, що спостерігається в експериментах.*

*Теоретично досліджені ефекти зарядки й одноелектронного тунелювання в структурі на кластері. У рамках моделі нескінченної потенційної ями для сферичних і дискообразних золотих кластерів обчислені електронний спектр і температурна залежність хімічного потенціалу. Різниця між хімічними потенціалами масивних електродів і островців приводить до зарядки останніх. Ми показуємо, що ефективний залишковий заряд не дорівнює цілому значенню заряду електрона  $e$  й залежить від форми кластера. Рівняння для аналізу вольт-амперної характеристики використовуються з урахуванням обмежень, пов'язаних з кулоновською нестійкістю кластера. Для одноелектронних молекулярних транзисторів обчислені немономонні розмірні залежності струмової щілини і її асиметрія по напрузі. Ми припускаємо, що перегрів електронної підсистеми приводить до зникнення струмової щілини й поступового згладжування вольт-амперних характеристик, що спостерігається в експериментах.*

УДК 537.874.6

Я. В. Чумаченко, В. П. Чумаченко

## К ОБОСНОВАНИЮ ЧИСЛЕННОГО РЕШЕНИЯ ОДНОЙ ЗАДАЧИ РАССЕЯНИЯ ВОЛН ДЛЯ НАГРУЖЕННОГО ИЗЛОМА ПРЯМОУГОЛЬНОГО ВОЛНОВОДА

*Рассматривается полученное ранее решение задачи рассеяния волн в Н-плоскостном  $90^\circ$ -изломе прямоугольного волновода, нагруженном диэлектриком. Исследована бесконечная система линейных алгебраических уравнений задачи. Установлена возможность ее решения методом усечения.*

### ВВЕДЕНИЕ

В последние годы метод произведения областей (ПО) успешно использовался для анализа ряда устройств волноводной техники. Укажем в качестве примера работы [1–3], где исследовались структуры различные по геометрии и назначению. В работе [4] метод был применен для решения задачи рассеяния

волн в нагруженном Н-плоскостном  $90^\circ$ -изломе прямоугольного волновода. Бесконечная система линейных алгебраических уравнений (СЛАУ) задачи решалась численно путем ее замены конечным числом уравнений. В настоящей статье приводится формальное обоснование применимости использованного метода усечения. Такое обоснование представляется важным, так как сходные матричные операторы появляются при применении метода ПО к исследованию и других узлов, в которых граничные поверхности пересекаются под прямым углом (как, например, в Т-соединениях и крестообразных соединениях прямоугольных волноводов).

© Чумаченко Я. В., Чумаченко В. П., 2009



### ИСХОДНЫЕ СООТНОШЕНИЯ

Структура, рассмотренная в [4], представляет собой волноводный уголок с разной шириной плеч  $c$  и  $d$ . Прямоугольная соединительная полость нагружена диэлектриком с проницаемостью  $\epsilon\epsilon_0$ . Волна  $H_{10}$  набегаёт со стороны волновода шириной  $d$ .

Задача сведена к решению бесконечной системы линейных алгебраических уравнений относительно коэффициентов разложения  $B_m^{(1)}$  и  $B_m^{(2)}$ , связанных с соединительной полостью. Система имеет вид:

$$B_m^{(1)} + \frac{1}{\Delta_m^{(x)}} \sum_{n=1}^{\infty} d_{mn}^{(x)} B_n^{(2)} = \frac{2\delta_{1m}\gamma_1^{(1)}}{\Delta_1^{(x)}}, \quad (1)$$

$$\frac{1}{\Delta_m^{(y)}} \sum_{n=1}^{\infty} d_{mn}^{(y)} B_n^{(1)} + B_m^{(2)} = 0, \quad m = 1, 2, \dots, \quad (2)$$

где

$$d_{mn}^{(x)} = (-1)^{m+n} \frac{2mn\pi^2}{cd^2} \frac{e^{-2\gamma_n^{(y)}d} - 1}{(\gamma_n^{(y)})^2 + (m\pi/d)^2}, \quad (3)$$

$$\Delta_m^{(x)} = \gamma_m^{(1)}(1 - e^{-2\gamma_m^{(x)}c}) + \gamma_m^{(x)}(1 + e^{-2\gamma_m^{(x)}c}), \quad (4)$$

$$\gamma_m^{(x)} = \left[ \left( \frac{m\pi}{d} \right)^2 - \chi^2 \right]^{\frac{1}{2}}, \quad \gamma_m^{(1)} = \left[ \left( \frac{m\pi}{d} \right)^2 - k_0^2 \right]^{\frac{1}{2}}, \quad (5)$$

$$\chi = \sqrt{\epsilon}k_0, \quad k_0 = \frac{2\pi}{\lambda}, \quad (6)$$

$\lambda$  – длина волны в свободном пространстве.

Значения  $\gamma_m^{(y)}$ ,  $\gamma_m^{(2)}$ ,  $d_{mn}^{(y)}$ ,  $\Delta_m^{(y)}$  могут быть найдены из (3)–(5) путем замены  $(x) \leftrightarrow (y)$ ,  $c \leftrightarrow d$  и (1)  $\rightarrow$  (2).

### АНАЛИЗ МАТРИЧНЫХ КОЭФФИЦИЕНТОВ СЛАУ

Введем новые искомые величины

$$\tilde{B}_m^{(1)} = B_m^{(1)}m, \quad \tilde{B}_m^{(2)} = B_m^{(2)}m. \quad (7)$$

Тогда из (1), (2) получим:

$$\tilde{B}_m^{(1)} + \frac{1}{\Delta_m^{(x)}} \sum_{n=1}^{\infty} \tilde{d}_{mn}^{(x)} \tilde{B}_n^{(2)} = \frac{2\delta_{1m}\gamma_1^{(1)}m}{\Delta_1^{(x)}}, \quad (8)$$

$$\frac{1}{\Delta_m^{(y)}} \sum_{n=1}^{\infty} \tilde{d}_{mn}^{(y)} \tilde{B}_n^{(1)} + \tilde{B}_m^{(2)} = 0, \quad m = 1, 2, \dots, \quad (9)$$

где

$$\tilde{d}_{mn}^{(x)} = (-1)^{m+n} \frac{2m^2\pi^2}{cd^2} \frac{e^{-2\gamma_n^{(y)}d} - 1}{(\gamma_n^{(y)})^2 + (m\pi/d)^2}. \quad (10)$$

Последнее соотношение легко преобразуется к виду

$$\tilde{d}_{mn}^{(x)} = (-1)^{m+n} \frac{2m^2c}{d^2} \frac{e^{-2\gamma_n^{(y)}} - 1}{n^2 + (\gamma_m^{(x)}c/\pi)^2}. \quad (11)$$

При

$$\frac{m\pi}{d} > \chi \quad (12)$$

$\gamma_m^{(x)}$ ,  $\gamma_m^{(1)}$  являются действительными величинами и  $\gamma_m^{(x)} < \gamma_m^{(1)}$ , что обеспечивает

$$\Delta_m^{(x)} > 2\gamma_m^{(x)}. \quad (13)$$

Оценим теперь сумму ряда  $\sum_{n=1}^{\infty} \left| \frac{\tilde{d}_{mn}^{(x)}}{\Delta_m^{(x)}} \right|$ :

$$\sum_{n=1}^{\infty} \left| \frac{\tilde{d}_{mn}^{(x)}}{\Delta_m^{(x)}} \right| < \sum_{n=1}^{\infty} \frac{|\tilde{d}_{mn}^{(x)}|}{2\gamma_m^{(x)}} \leq \sum_{n=1}^{\infty} \frac{m^2c}{d^2\gamma_m^{(x)}} \frac{1}{n^2 + \left( \gamma_m^{(x)} \frac{c}{\pi} \right)^2} = A. \quad (14)$$

Учитывая известные [5] соотношения

$$\sum_{n=1}^{\infty} \frac{1}{n^2 + a^2} = \frac{\pi}{2a} \left( \text{ctha}\pi - \frac{1}{a\pi} \right), \quad (15)$$

$$\text{cthz} - \frac{1}{z} \leq 1, \quad 0 \leq z < \infty, \quad (16)$$

получим (при  $a = \gamma_m^{(x)} \frac{c}{\pi}$ )

$$A \leq \frac{\pi^2 m^2}{2d^2 \gamma_m^{(x)2}} = \frac{1}{2} \frac{1}{1 - \left( \frac{\chi d}{m\pi} \right)^2}. \quad (17)$$

Из (14), (17) следует, что при

$$m > \sqrt{2} \frac{\chi d}{\pi} \quad (18)$$

выполняется неравенство

$$\sum_{n=1}^{\infty} \left| \frac{\tilde{d}_{mn}^{(x)}}{\Delta_m^{(x)}} \right| < 1. \quad (19)$$

Аналогично можно показать, что при

$$m > \sqrt{2} \frac{\chi c}{\pi} \quad (20)$$

справедливо неравенство

$$\sum_{n=1}^{\infty} \left| \frac{\tilde{d}_{mn}^{(y)}}{\Delta_m^{(y)}} \right| < 1. \quad (21)$$

Переномеруем формально неизвестные как  $A_1 = \tilde{B}_1^{(1)}$ ,  $A_2 = \tilde{B}_1^{(2)}$ ,  $A_3 = \tilde{B}_2^{(1)}$ ,  $A_4 = \tilde{B}_2^{(2)}$ , ... и перепишем в соответствующем порядке уравнения системы (8), (9). В силу (19) и (21) эта система является квазирегулярной [6] и ее решение известным образом сводится к решению конечной системы алгебраических уравнений.

### ЗАКЛЮЧЕНИЕ

Выполнен анализ матричных коэффициентов СЛАУ, возникающей при применении метода ПО к задаче рассеяния волн в Н-плоскостном волноводном изломе с соединительной полостью, заполненной диэлектриком. Показано, что для решения системы применим метод усечения. Полученные результаты могут быть также использованы при исследовании нагруженных Т-соединений и крестообразных соединений прямоугольных волноводов.

### ПЕРЕЧЕНЬ ССЫЛОК

1. Chumachenko V. P. Simple full-wave model of E-plane waveguide star junction // V. P. Chumachenko Journal of Electromagnetic Waves and Application. – Vol. 16. – Sept. 2002. – P. 1223–1232.

2. Petrusenko I. V. Scattering by inductive post in uniformly curved rectangular waveguide // I. V. Petrusenko, V. P. Chumachenko. – IEE Proceedings. Microwaves, Antennas and Propagation. – Vol. 150. – Dec. 2003. – P. 498–504.
3. Ващенко В. В. Решение методом произведения областей Н-плоскостной задачи дифракции волн на наклонной границе раздела диэлектрических сред в прямоугольном волноводе // В. В. Ващенко, В. П. Чумаченко. – Радиоелектроніка. Інформатика. Управління. – № 1. – 2007. – С. 5–9.
4. Chumachenko V. P. Accurate analysis of waveguide junctions with rectangular coupling cavity // V. P. Chumachenko, E. Karacuha, I. V. Petrusenko. – Microwave and Optical Technology Letters. – Vol. 31. – Nov. 2001. – P. 305–308.
5. Партон В. З. Методы математической теории упругости / В. З. Партон, П. И. Перлин – М. : Наука, 1981. – 688 с.
6. Канторович Л. В. Приближенные методы высшего анализа / Л. В. Канторович, В. И. Крылов. – М. : Физматгиз, 1962. – 708 с.

Надійшла 27.04.2009  
Після доробки 14.05.2009

*Розглядається отриманий раніше розв'язок задачі розсіювання хвиль в Н-площинному  $90^\circ$  зламі прямокутного хвилеводу, навантаженого діелектриком. Досліджена нескінченна система лінійних алгебраїчних рівнянь задачі. Встановлена можливість знаходження її розв'язку методом зрізання.*

*Solution of the H-plane scattering problem for a dielectric-loaded  $90^\circ$  bend of a rectangular waveguide is considered. A previously derived infinite system of algebraic equations has been analyzed. It has been found that the system can be solved using a truncation procedure.*

# РАДІОЕЛЕКТРОНІКА ТА ТЕЛЕКОМУНІКАЦІЇ

## РАДИОЕЛЕКТРОНИКА И ТЕЛЕКОММУНИКАЦИИ

### RADIO ELECTRONICS AND TELECOMMUNICATIONS

---

---

УДК 621.391

А. Н. Буханько, В. М. Безрук, Е. В. Дуравкин

#### ПРИМЕНЕНИЕ Е-СЕТЕЙ ДЛЯ МОДЕЛИРОВАНИЯ АЛГОРИТМОВ УПРАВЛЕНИЯ УЧАСТКОМ ТЕЛЕКОММУНИКАЦИОННОЙ СЕТИ

---

*В статье разработаны имитационные модели алгоритмов работы управляющего агента для участка сложной телекоммуникационной сети на основе выбранного аппарата Е-сетей.*

#### **ВВЕДЕНИЕ**

В настоящее время в связи с бурным развитием ресурсоемких сетевых приложений, повышением скорости и объемов передачи информации актуальным является вопрос создания эффективных децентрализованных систем управления сложными телекоммуникационными сетями (ТКС) [1]. С другой стороны, разработчикам отводится не так много времени для разработки и внедрения новых систем управления информационными потоками, предназначенных для удовлетворения требований, которые предъявляются к современным ТКС. Следовательно, в их арсенале должны быть средства, позволяющие за достаточно короткий срок получить информацию об разрабатываемой системе управления с заданной полнотой и использовать ее при их разработке.

Исследование показателей работы системы управления можно провести путем испытаний созданного образца либо путем испытаний на ее математической модели. Сложность и значительная стоимость постро-

ения реальных систем управления ТКС указывает на трудности, связанные с реализацией первого подхода к их исследованию. Поэтому доступным методом исследования разрабатываемых или модернизируемых систем управления является их моделирование.

Также необходимо отметить, что система управления информационным обменом в сложной гетерогенной ТКС является сложной и многоуровневой. В настоящее время на практике происходит переход от централизованных систем управления ТКС к децентрализованным [1], в которых управления участками ТКС осуществляется управляющими агентами (УА). Управляемый участок сети ТКС представляет собой совокупность каналов связи (КС), инцидентных (принадлежащих) определенному УА. При этом возникает необходимость создания математических моделей УА для участков ТКС с целью исследования их характеристик. Однако, построение аналитических моделей агентов систем управления является нецелесообразным в связи с большой трудоемкостью их разработки и сложностью анализа. В этом случае рациональным является построение моделей агентов с использованием имитационного аппарата, в частности, Е-сетей [2, 3]. Е-сети являются наиболее подходящим

аппаратом имитационного моделирования, позволяющим легко моделировать параллельные взаимодействующие процессы, отражая общую динамику работы децентрализованной агентной системы управления. Аппарат Е-сетей допускает различную трактовку своих элементов по уровню абстракции (детализации), что позволяет строить иерархические модели УА.

Целью данной статьи является разработка имитационных моделей алгоритмов работы УА, включающих процедуры определения пропускных способностей (ПС) каналов связи КС, принадлежащих агенту, и выбора подходящих КС по критерию QoS.

## ПОСТАНОВКА ЗАДАЧИ

Построение имитационной модели алгоритмов управления участком сложной гетерогенной ТКС на основе математического аппарата Е-сетей, что включает следующие этапы:

- 1) определение дискриптивной алгоритмической модели процессов управления;
- 2) построение формализованной модели процессов управления, в частности, представление состояний процесса и переходов между ними в соответствии с основными положениями теории Е-сетей;
- 3) построение математической модели, а именно, математическое описание элементов графа Е-сети (предикатные переходы, метки, генераторы, места (состояния), места-поглотители и т. д.);
- 4) проверка степени адекватности построенной имитационной модели путем проверки выполнения основных алгоритмических свойств Е-сетей.

## ПОСТРОЕНИЕ ИМИТАЦИОННЫХ МОДЕЛЕЙ РАБОТЫ УА

### 1 Процедура определения пропускных способностей каналов связи

Рассмотрим первую процедуру работы УА [4]. Ее задачей является анализ ПС КС и динамики ее изменения на основе численных характеристик выбранных критериев QoS. В функции процедуры входит:

- получение номеров КС с постоянной либо динамически увеличивающейся ПС – в случае сохранения либо улучшения характеристик участка ТКС;
- получение номеров КС с динамически уменьшающейся ПС – в случае ухудшения характеристик участка ТКС;
- выдача полученных результатов следующей процедуре выбора КС по критерию QoS.

Указанное управление КС базируется на основе анализа следующих критериев качества QoS: ПС в текущий и предыдущий такт работы процедуры ( $N_k$ ,  $N_{k-1}$ , соответственно) и длительности управляющего интервала  $t_{\text{зад}}$ .

Полученные значения ПС текущего и предыдущего интервала работы  $N_k$ ,  $N_{k-1}$  позволяют сделать вывод об изменении ПС канала, что дает возможность оценить динамику изменения состояния данного канала. Критерий качества  $t_{\text{зад}}$  позволяет оценить время реакции (оперативность) управления КС на основе полученной задержки такта управления.

Исходя из вышесказанного, в предложенной имитационной модели предлагается использовать нагруженную метку, моделирующую КС, следующего вида:

$$R[t_{\text{зад}}, N_k, N_{k-1}, T_k, m, C, i], \quad (1)$$

где  $R$  – вектор атрибутов метки;  $t_{\text{зад}}$  – время жизни метки;  $N_k$ ,  $N_{k-1}$  – ПС на  $k$ -м и  $(k-1)$ -м интервале времени  $m$ -го канала, соответственно;  $T_k$  – счетчик тактов управления;  $m$  – номер КС;  $C$  – относительная стоимость использования канала;  $i$  – первый информационный атрибут (принимает значения 0 или 1; при инициализации – 0);  $j$  – второй информационный атрибут (принимает значение 0, 1 или 2; при инициализации – 0).

Задача определения номеров КС с недостаточной ПС сводится к анализу динамики разности критериев качества ПС предыдущего и данного шагов работы агента. Расчет данных критериев происходит по значению времени задержки управляющего пакета обратной связи (например,  $t_{\text{tcp}}$ ), используя соотношение (2)

$$N_k = \frac{n}{t_{\text{зад}}}, \quad (2)$$

где  $n$  – размер пакета  $t_{\text{tcp}}$ ;  $t_{\text{зад}}$  – значение задержки пакета.

После анализа выражения (2) производится проверка разности ПС двух последовательных тактов управления на принадлежность некоторому экспертно заданному кризисному интервалу  $\Delta n_k^m$ . Если полученное значение находится в данном интервале более заранее заданного времени  $T_3$  (расчет по атрибуту  $T_k$ ) и, при этом, ПС меньше некоторого минимального значения  $N_k^m(\text{min})$ , то информационным атрибутам метки присваивается индекс 1. В этом случае канал, моделируемый данной меткой, не используется в течение следующего такта работы УА. Упрощенный вариант Е-сети, моделирующей данную процедуру, представлен на рис. 1.

*Управляющие переходы* [2]:  $MY_1$  – решающее правило доступа КС на анализ ПС;  $MX_2$  – анализ суммарного количества тактов управления для выбранного КС (один или больше одного);  $MX_3$  – проверка канала на нахождение разницы его ПС в критическом интервале;  $MX_4$  – анализ количества временных тактов управления, при котором сохраняется критический интервал;  $MX_5$ ,  $MX'_5$  – сравнение текущей ПС КС с неким его минимальным значением

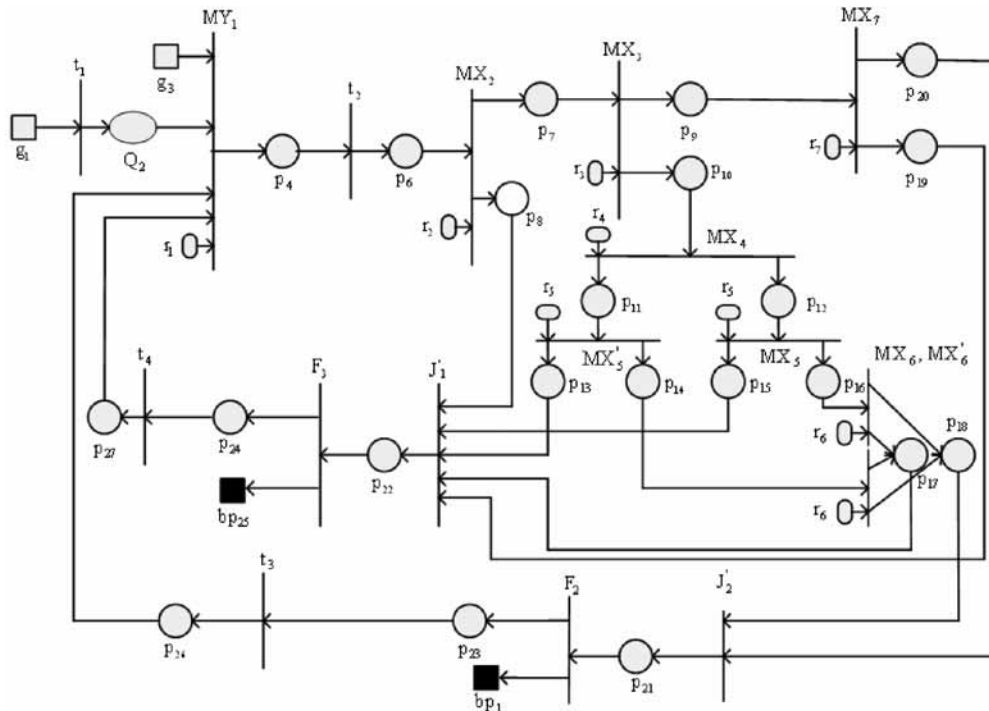


Рисунок 1 – Е-сеть первой процедуры

(критическим), сохраняющим его работоспособность;  $MX_6, MX'_6$  – определение количества интервалов управления, при которых текущая ПС КС была меньше критической;  $MX_7$  – анализ канала на многократно увеличенную разницу ПС.

*Переходы:*  $t_1$  – передача КС на анализ процедуре;  $t_2$  – формирование, отправка и получение ПС КС по значению временных параметров обратного пакета  $ttsp$ ;  $t_3, t_4$  – изменение информационных параметров QoS КС при переходе его на последующий шаг работы алгоритма;  $J'_1$  – обработка КС с достаточной ПС;  $J'_2$  – обработка КС с недостаточной ПС;  $F_2, F_3$  – разветвители потоков меток.

*Предикаты управляющих переходов* определяются следующим образом:

1. Управляющие переходы типа MX:

$$r_2 = \begin{cases} 0, M(p_6(R(T_k))) = 0, p_8; \\ 1, M(p_6(R(T_k))) > 1, p_7, \end{cases}$$

$$r_3 = \begin{cases} 0, M(p_7(R(N_k))) - M(p_7(R(N_{k-1}))) \neq \Delta n_k^m, p_9; \\ 1, M(p_7(R(N_k))) - M(p_7(R(N_{k-1}))) = \Delta n_k^m, p_{10}, \end{cases} \quad (3)$$

где  $N_k, N_{k-1}$  – ПС канала на  $k$ -м и  $(k-1)$ -м интервале времени;

$$r_4 = \begin{cases} 0, M(p_{10}(R(T_k))) < T_3 \rightarrow M(p_{11}(R(T_k))) = \\ = M(p_{10}(R(T_k))) + 1, p_{11}; \\ 1, M(p(R(T))) \geq T_3 \rightarrow M(p_{12}(R(i))) = 1, p_{12}, \end{cases} \quad (4)$$

где  $T_3$  – предельное значение счетчика временных интервалов,  $i$  – информационный параметр, указывающий, что КС с данным параметром становится «претендентом» на исключение из списка работающих;

$$r_5 = \begin{cases} 0, M(p_{12}(R(N_k^m))) < \neq N_k^m(\min), p_{15}; \\ 1, M(p_{12}(R(N_k^m))) < N_k^m(\min) \cup \\ \cup M(p_{12}(R(j))) = 0 \rightarrow M(p_{16}(R(j))) = 1; \\ M(p_{12}(R(N_k^m))) < N_k^m(\min) \cup \\ \cup M(p_{12}(R(j))) = 1 \rightarrow M(p_{16}(R(j))) = 2, p_{16}, \end{cases}$$

$$r_5 = \begin{cases} 0, M(p_{11}(R(N_k^m))) < \neq N_k^m(\min), p_{13}; \\ 1, M(p_{11}(R(N_k^m))) < N_k^m(\min) \cup \\ \cup M(p_{11}(R(j))) = 0 \rightarrow M(p_{14}(R(j))) = 1; \\ M(p_{11}(R(N_k^m))) < N_k^m(\min) \cup \\ \cup M(p_{14}(R(j))) = 1 \rightarrow M(p_{14}(R(j))) = 2, p_{14}, \end{cases} \quad (5)$$

где  $N_k^m(\min)$  – минимальное значение ПС для  $m$ -го КС;

$$r_6 = \begin{cases} 0, M(p_{16}(R(j))) = 0, 1, p_{17}; \\ 1, M(p_{16}(R(j))) = 2, p_{18}, \\ 0, M(p_{14}(R(j))) = 0, 1, p_{17}; \\ 1, M(p_{14}(R(j))) = 2, p_{18}. \end{cases} \quad (6)$$

$$r_7 = \begin{cases} 0, M(p_9(R(N_k))) - M(p_9(R(N_{k-1}))) >> \neq \Delta n_k^m, p_{19}; \\ 1, M(p_9(R(N_k))) - M(p_9(R(N_{k-1}))) >> \Delta n_k^m, p_{20}. \end{cases} \quad (7)$$

2. Управляющие переходы типа МУ:

$$r_1 = \begin{cases} Q_1, M(p(R(N_k))) = 0; \\ Q_{30}, M(p(R(N_k))) = 0 \cup g_{32} = \\ = 1, 2\tau_{12}, M(p(R(t_{\text{зад}}))) = 0; \\ Q_{31}, M(p(R(N_k))) = 0 \cup g_{32} = \\ = 1, 2\tau_{12} + 1, M(p(R(t_{\text{зад}}))) = 0, \end{cases} \quad (8)$$

где  $2\tau_{12}$ ,  $2\tau_{12} + 1$  – четный и нечетный период времени работы перехода  $МУ_1$ ,  $g_3$  – управляющий генератор, работающий в следующих промежутках времени  $[(T, (T+1)/2), (T+1, (T+2)/2), \dots, (T+(n-1)), (T+n)/2]$ .

Полученная Е-сеть работает в тактовом режиме анализа КС. Цикл управления данной сетью  $T$  определяет время управления КС, что включает в себя следующие операции: получение ПС согласно (2), анализ критериев качества QoS по полученным предикатам (3)–(8), вынесение решения по использованию каждого КС в последующем такте работы, выдачу результатов следующей процедуре алгоритма управления и обнуление результатов. Значение  $T$  определяется согласно экспертным оценкам или, исходя из статистики работы системы управления, и задается заранее.

Из данной имитационной модели следует, что при начальной неопределенности количества КС метки обрабатываются и продвигаются по сети последовательно. Данная модель считается адекватной, если время срабатывания любого перехода полученной Е-сети намного меньше общего времени такта. При этом Е-сеть должна иметь свойство строгой ограниченности – сохраняемости [5]. В этом случае количество меток в сети интерпретируется как некий постоянный ресурс. В разработанной модели количество меток равно количеству КС.

Следовательно, ПС  $m$ -го КС находится с помощью соотношения (2), где параметр  $t_{\text{зад}}$  определяется задержкой перехода  $t_2$ . Исходя из логики процедуры алгоритма управления [5], результатом работы Е-сети являются два потока меток, моделирующих два возможных типа КС: с достаточной ПС для передачи информации и игнорируемых на следующем такте работы сети. Номера соответствующих КС выдаются Е-сетью базе данных процедуры определения работоспособных путей передачи информации с использованием мест-поглотителей, соответственно,  $bp_{25}$  и  $bp_1$ . Параллельно, обнуляются позиции  $N_{k-1}$  для всех КС и информационных параметров  $i, j$  для игнорируемых КС и метки передаются на переход  $МУ_1$ , моделирующий следующий такт работы процедуры алгоритма управления.

Результатом моделирования является определение перечня КС с достаточной ПС для передачи информации, а также КС с недостаточной ПС с возможно-

стью их исключения из списка работающих в следующий такт работы процедуры. При этом такая модель позволяет обработать ситуации, при которых происходит игнорирование КС при минимальном уменьшении ПС в них или случайном увеличении производительности КС после долгого постепенного спада.

## 2 Процедура выбора подходящих каналов связи по критерию QoS

Рассмотренная имитационная модель определяет численные характеристики для процедуры нахождения оптимальных КС для передачи информации [4]. Функцией данной процедуры является получение номера(ов) канала(ов), удовлетворяющих некоторым критериям QoS для передачи исходного потока данных с максимально возможным качеством. Другими словами, из множества КС  $M$ , подходящих для передачи информации, полученного с использованием процедуры 1, выделяется подмножество  $M' \subseteq M$  оптимальных КС для передачи заданного информационного потока данных. В данном случае, критерием оптимальности является минимизация количества КС, используемых для передачи, при минимальных затратах. Затраты определяются в виде минимальной ПС и исходной стоимости использования КС. Подмножество  $M'$  в частном случае может содержать один канал.

Как и для процедуры 1, построим имитационную модель на основе аппарата Е-сетей, поскольку использование аналитического аппарата в данном случае приводит к громоздким математическим моделям и невозможности их изменения и адаптации к новым условиям работы.

Функциями предлагаемой Е-сети являются:

- получение списка КС, удовлетворяющих задаче передачи потока данных с заданной ПС  $N_n$ ;
- определение оптимального набора КС, удовлетворяющих критерию  $\min(N_m, N_k^m, c)$ , то есть минимального количества КС с минимальными ПС и стоимостью.

Имитационная модель процедуры 2 приведена на рис. 2.

*Управляющие переходы:*  $МХ_1$  – сравнение ПС КС с ПС, необходимой для передачи поступающего потока данных и запрет использования игнорируемых КС;  $МХ_2$  – определение количества КС с удовлетворяющей потоку данных ПС;  $МХ_3$  – определение отношения критериев качества QoS пропускная способность/стоимость для КС с удовлетворяющей ПС, нахождение КС с минимальной ПС и стоимостью, достаточной для передачи заданного потока данных;  $МХ_4$  – определение соотношения критериев качества QoS пропускная способность/стоимость для КС с неудовлетворяющей ПС;  $МУ_5$  – нахождение опти-

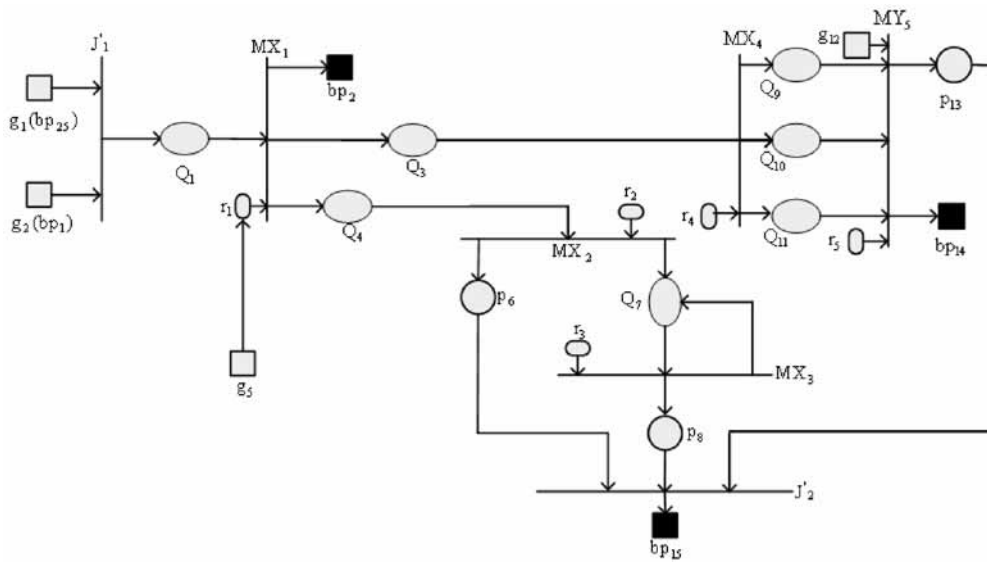


Рисунок 2 – Е-сеть процедуры 2

мального набора КС для передачи информации при их максимальной ПС и минимальной стоимости.

*Неуправляемые переходы:*  $J'_2$  – определение способа передачи заданного потока данных.

*Предикаты управляющих переходов* определяются следующим образом:

1. Управляющие переходы MX:

$$r_1 = \begin{cases} bp_2, M(Q_1(R(i))) = 1 \cup M(Q_1(R(j))) = 2; \\ Q_3, M(Q_1(R(N_k^m))) < N_{II}; \\ Q_4, M(Q_1(R(N_k^m))) \geq N_{II}, \end{cases} \quad (9)$$

где  $N_{II}$  – ПС, необходимая для передачи потока данных, генерируемая  $g_5$ ;

$$r_2(\tau) = \begin{cases} 0, N(M(Q_4(R(N_k^m \geq N_{II}))) > 1, p_6; \\ 1, N(M(Q_4(R(N_k^m \leq N_{II}))) = 1, Q_7, \end{cases} \quad (10)$$

где  $N(M(Q_4(R(N_k^m \leq N_{II})))$  – количество меток, переходящих через переход со значением ПС больше исходной;  $\tau$  – время, в течении которого активен данный переход:  $\tau \ll T$ , где  $T$  – интервал генерации (поступления) меток с мест  $g_1$  и  $g_2$ ;

$$r_3 = \begin{cases} P_8, M(Q_7(R(N_k^m))) \in A = \{N_k^m: \min_{N_k^m} (N_k^m \leq N_{II})\} \cup \\ \cup M(Q_7(R(c))) \in B = \{c_k^m: \min(c_k^m - c_{\min})\}; \\ Q_7, M(Q_7(R(N_k^m))) \in A^C \cup M(Q_7(R(c))) \in B^C, \end{cases} \quad (11)$$

где  $A$  – множество ПС, максимально приближенных (приблизительно равных) ПС потока данных  $N_{II}$ ;

$(N_k^m \leq N_{II})$ ;  $B$  – множество стоимостей, максимально приближенных к минимальной выбранной стоимости  $c_{\min}$ :  $(c_k^m - c_{\min})$ ;  $A^C$  – дополнение к множеству  $A$ ;  $B^C$  – дополнение к множеству  $B$ ;

$$r_4 = \begin{cases} Q_9, M(Q_3(R(N_k^m))) \in [0, N_{II}/4); \\ Q_{10}, M(Q_3(R(N_k^m))) \in [N_{II}/4, N_{II}/2); \\ Q_{11}, M(Q_3(R(N_k^m))) \in [N_{II}/2, N_{II}], \end{cases} \quad (12)$$

где  $[0, N_{II}/4)$ ,  $[N_{II}/4, N_{II}/2)$ ,  $[N_{II}/2, N_{II})$  – интервалы несоответствия ПС каналов и потока данных  $N_{II}$ .

2. Управляющие переходы MY:

$$r_5 = \begin{cases} N_k \cdot C_k = \{(N_k^1, C_k^1) \dots (N_k^i, C_k^i) \dots (N_k^m, C_k^m)\} \rightarrow \\ \rightarrow M(p_{13}(R(M_m))) = \{m_k^1, \dots, m_k^m\} \sum_{i=1, \dots, m} N_k^i \geq \\ \geq N_{II}, M(p_{13}(R(t_{\text{зад}}^i))) = Q_9, Q_{10}, Q_{11}(M(R(t_{\text{зад}}^i + \tau_5))), \\ m \leq n; \\ \text{не\_найден} \sum_{i=1, \dots, m} N_k^i \geq N \rightarrow bp_{14}, \end{cases} \quad (13)$$

где  $N_k = \{M(R(N_k^1)), M(R(N_k^2)), \dots, M(R(N_k^i))\}$  – множество меток с мест  $Q_9, Q_{10}, Q_{11}$ , рассортированные по убыванию значения ПС  $N_k$ ;  $C_k = \{M(R(c_k^1)), M(R(c_k^2)), \dots, M(R(c_k^m))\}$  – множество меток с мест  $Q_9, Q_{10}, Q_{11}$ , рассортированные по убыванию значения стоимости  $C_k$ ;  $N_k \cdot C_k$  – прямое произведение двух множеств;  $\{m_k^1, \dots, m_k^m\}$  – множество КС, находящихся в оптимальном наборе для передачи;  $\tau_5$  – время работы управляющего перехода  $MY_5$ .

В данной модели циркулируют метки, полученные с использованием мест-поглотителей  $bp_{25}$  (КС

с удовлетворяющей ПС) и  $bp_1$  (игнорируемые КС) процедуры 1. В рассматриваемой модели эти места представляются в виде генераторов  $g_1$  и  $g_2$ , которые выдают последовательности меток с интервалом времени  $T$ .

Предикат  $r_1$  (9) производит анализ состояния ПС каналов методом сравнения их с заданной ПС потока  $N_{\text{п}}$ . Результатом этого сравнения выступает два потока меток: каналы с ПС, превышающей заданную, и каналы с меньшей ПС. Каналы с идентификаторами игнорирования не используются (место  $bp_2$ ). Условие  $r_2$  (10) определяет количество КС, пропускная способность которых превышает заданную ПС потока. Таких каналов может быть один или же некоторое конечное множество  $M' \subseteq M$ . Предикат  $r_3$  определяет в своем условии два множества: множество каналов с ПС, максимально приближенной к ПС потока с минимальной стоимостью и множество остальных КС. Первое множество определяет КС с минимальной относительной стоимостью, у которых ПС в данный момент незначительно отличается от требуемой, что дает возможность полностью использовать ресурсы ПС данного КС без использования ресурсов других каналов. В данной модели предполагается, что такой канал единственный.

Для меток с недостаточной ПС определяется степень ее отличия от номинальной требуемой  $N_{\text{п}}$ . Управляющий переход  $MY_5$  задает метку нового типа, используя генератор  $g_{12}$ . Данная метка имеет вид:

$$R[M_m, t_{\text{зад}}], \quad (14)$$

где  $M_m$  – множество КС для передачи с недостаточной ПС  $\{m_k^1, \dots, m_k^m\}$ .

Данный предикат выполняет функции нахождения оптимального множества КС, которые имеют ПС, входящие в промежутки  $N_k^m \in [0, N_{\text{п}}/4]$ ,  $N_k^m \in [N_{\text{п}}/4, N_{\text{п}}/2]$  и  $N_k^m \in [N_{\text{п}}/2, N_{\text{п}}]$ , то есть КС, которые в отдельности не имеют достаточно сетевых ресурсов для предоставления потоку данных гарантированного качества при передаче. Критерий оптимальности определяется минимальным количеством узлов для передачи с максимальной ПС  $\min(N_m, \max N_k^m, c)$  при минимальной стоимости КС. Предикат  $r_5$  определяет множество упорядоченных по убыванию ПС каналов  $N_k = \{N_k^1, N_k^2, \dots, N_k^n\}$  и их стоимости  $C_k = \{c_k^a, c_k^b, \dots, c_k^n\}$ . В общем случае номера в данных множествах не совпадают.

Для множества  $N_k$  определяется сумма ПС ее составляющих, равная или в малой степени превышающая заданную ПС потока данных  $\sum_{i=1, \dots, m} N_k^i \geq N_{\text{п}}$  при условии минимизации суммы стоимостей ее элементов

$$C_{\text{min}} = \sum_{i=1}^m c_k^i. \text{ ПС полученных каналов } N_k^1, N_k^2, \dots,$$

$N_k^m$  в данной сумме будут использованы для удовлетворения заданных параметров потока данных по ПС и стоимости. Полученные номера составляют множество  $M_m$ , реализованное вектором атрибутов метки (14). Если необходимый оптимальный набор КС не найден, то происходит отказ в обслуживании исходного потока данных в текущий такт работы алгоритма управления КС агентом.

## ВЫВОДЫ

Полученные имитационные модели алгоритмов управления участками ТКС дают возможность получить оценки численных характеристик систем управления ТКС, таких как среднее время реакции УА, количество работоспособных и игнорируемых КС, вероятность выхода КС из строя, а также зависимость данных характеристик от параметров сетевых потоков. Практическая ценность полученных моделей заключается в возможности их программной реализации в УА с использованием любых современных транспортных технологий.

Перспективой дальнейших исследований является разработка имитационных моделей рассмотренных процедур алгоритмов управления с использованием нечеткой логики в предикатах управляющих переходов E-сетей.

## ПЕРЕЧЕНЬ ССЫЛОК

1. Гребешков А. Ю. Стандарты и технологии управления сетями связи / Гребешков А. Ю. – Москва : Экотрендз, 2003. – 287 с.
2. Лосев Ю. И. Применение E-сетей для моделирования процесса функционирования СОД / Лосев Ю. И., Шматков С. И., Дуравкин Е. В. // Радиотехника: Всеукр. межвед. науч.-техн. сб. – 2001. – Вып. 123. – С. 99–103.
3. Иванов В. Г. Разработка и исследование имитационных моделей распределенных информационно-вычислительных систем : автореферат дисс. на соискание ученой степени к. т. н. / Иванов В. Г. – Харьков : ХИРЭ, 1989. – 23 с.
4. Буханько А. Н. Алгоритмы управляющего агента участка сети связи / Буханько А. Н. // Мат. 3-го Международного радиоэлектронного форума «Прикладная радиоэлектроника. Состояние и перспективы развития». – Харьков : ХНУРЭ, 2008. – С. 37–40.
5. Лосев Ю. И. Применение методов анализа E-сетей к моделям СОД / Лосев Ю. И., Шматков С. И., Дуравкин Е. В. // Радиотехника : всеукр. межвед. науч.-техн. сб. – 2002. – Вып. 132. – С. 136–141.

Надійшла 12.05.2009  
Після доробки 18.05.2009

*У статті розроблені імітаційні моделі алгоритмів роботи агента, що управляє ділянкою складної телекомунікаційної мережі на основі вибраного апарату E-мереж.*

*In the article the simulation models of managing agent's algorithms based on the chosen method of E-networks for the area of difficult telecommunication network are developed.*



О. М. Ибрагимов

## ИСПОЛЬЗОВАНИЕ ЧАСТОТНОГО РАДИОВЫСОТОМЕРА ДЛЯ ОПРЕДЕЛЕНИЯ ШЕРОХОВАТОСТИ ЗЕМНОЙ ПОВЕРХНОСТИ КОРРЕЛЯЦИОННЫМ МЕТОДОМ

---

*Рассматривается задача определения среднеквадратической высоты неровностей рельефа земной поверхности с помощью частотного радиовысотомера. Приводятся основные результаты теоретического анализа корреляционной функции принимаемого сигнала, полученной с учетом влияния параметров движения носителя.*

### ВВЕДЕНИЕ

Создание методов получения наиболее полной информации, имеющейся в радиолокационных сигналах и помехах, остается важным направлением современной радиолокации. Дистанционное зондирование земной поверхности является частью этого направления. В качестве дополнительной информации об отражающей поверхности при активной локации используются характеристики поля вторичного излучения, как правило, зависящие также от угла визирования, параметров движения носителя, характеристик зондирующих сигналов, способа обработки отраженных сигналов. Информация, получаемая с помощью дистанционных методов, отличается оперативностью, возможностью ее получения из больших и труднодоступных районов.

Результаты многочисленных исследований [1–4] показывают возможность применения для дистанционного зондирования параметров отражающей поверхности устройств автономной радионавигации летательных аппаратов (доплеровские измерители путевой скорости и угла сноса, радиовысотомеры).

Зондирование отражающей поверхности с помощью доплеровских измерителей скорости, как правило, связано с использованием крупногабаритных остронаправленных антенн или антенн с модифицированной формой диаграммы направленности (ДН) [4], стабильно ориентированной в пространстве, что не всегда удобно при их размещении на борту авиационного носителя в заданных условиях полета.

В настоящей статье представлены результаты теоретического анализа информационных возможностей отраженного от шероховатой поверхности сигнала частотного радиовысотомера (ЧМ-РЛС) со штатной антенной, ориентированной вертикально в надир. В результате анализа рассеянного отражающей поверхностью поля, выполненного в приближении Кирхгофа,

устанавливаются аналитические связи между геометрическими параметрами отражающей поверхности и статистическими характеристиками отраженного сигнала.

ЧМ-РЛС, в которой применяется корреляционный метод обработки сигнала биений, представляет собой систему аналогичную двухчастотной РЛС (двухчастотный радиолокационный интерферометр), где применяется известный «к»-метод [3]. Трудности практической реализации «к»-метода заключаются в том, что необходимо обеспечить высокую степень развязки между каналами РЛС, работающими на близких частотах. При использовании ЧМ-РЛС представляется возможной более простая практическая реализация корреляционного метода измерения высоты неровностей отражающей поверхности, поскольку применяется одноканальная РЛС.

### ПОСТАНОВКА ЗАДАЧИ

В большинстве практических случаев величина облучаемой площадки поверхности оказывается заметно больше геометрических размеров ее неровностей, что обусловлено высотой полета и шириной ДН. Отраженное поле в этом случае формируется от множества неровностей поверхности и при формировании оценки высоты в радиовысотомере происходит статистическое усреднение по всем наклонным дальностям. Рельеф поверхности при этом в показаниях высотомера сглаживается и высотомер его не воспроизводит. Поэтому определение шероховатости отражающей поверхности требует дополнительной обработки принимаемого сигнала, флуктуации которого вызваны неровностями рельефа.

Далее рассматривается движение носителя РЛС в декартовой системе координат  $XYZ$  параллельно плоскости  $XOY$ , вектор скорости горизонтального полета  $V_0$  составляет с осью  $X$  угол  $\alpha$ , лучи антенн коллинеарны вертикальной оси  $Z$ , а центр облучаемой площадки совпадает с началом координат ( $x = 0$ ;  $y = 0$ ;  $z = 0$ ). Функцию  $h(x, y)$ , описывающую рельеф поверхности в любой заданный момент времени  $t$ , можно представить в виде суммы двух составляющих,

каждая из которых описывается независимой случайной функцией с нулевым средним значением:

$$h(x, y) = h_1(x, y) + h_2(x, y). \quad (1)$$

Функция  $h_1(x, y)$  описывает высоты крупных неровностей земной поверхности (крупных волн, применительно к морю), а функция  $h_2(x, y)$  – высоты всех остальных более мелких. В этом случае крупные и мелкие неровности (волны) будут характеризоваться, соответственно, пространственными корреляционными функциями вида:

$$\begin{aligned} B_{h_1}(\Delta x, \Delta y) &= \sigma_{h_1}^2 \rho_{h_1}(\Delta x, \Delta y) = \\ &= \sigma_{h_1}^2 \exp[-(\Delta x^2 + \Delta y^2)/l_{h_1}^2] \cos(2\pi\Delta x/\Lambda_x) \cos(2\pi\Delta y/\Lambda_y), \end{aligned} \quad (2)$$

$$\begin{aligned} B_{h_2}(\Delta x, \Delta y) &= \sigma_{h_2}^2 \rho_{h_2}(\Delta x, \Delta y) = \\ &= \sigma_{h_2}^2 \exp[-(\Delta x^2 + \Delta y^2)/l_{h_2}^2], \end{aligned} \quad (3)$$

где  $\Delta_x, \Delta_y$  – смещение в горизонтальной плоскости;  $\Lambda_x, \Lambda_y$  – средняя длина периодичности рельефа поверхности в направлении координат  $x, y$ ;  $l_{h_1}, l_{h_2}$  – интервалы корреляции крупных и мелких неровностей рельефа, соответственно;  $\sigma_{h_1}, \sigma_{h_2}$  – среднеквадратические значения высот крупных и мелких неровностей. На функции  $h_1(x, y)$  и  $h_2(x, y)$  накладываются следующие ограничения, которые в большинстве практических случаев могут быть приняты:

$$\left| \frac{dh_1}{dx} \right| \leq 0,5; \quad \left| \frac{dh_2}{dx} \right| \leq 0,2; \quad \left| \frac{dh_2}{dy} \right| \leq 0,2 \quad (4)$$

и 
$$\sigma_{h_1} \gg \sigma_{h_2} > \lambda \setminus 4, \quad (5)$$

где  $\lambda$  – длина излучаемой волны. Так же полагается, что размеры облучаемой площадки превышают радиус корреляции функции  $h(x, y)$ . Указанные условия позволяют при решении задачи воспользоваться приближением Кирхгофа [5], считать когерентную составляющую отраженного сигнала отсутствующей, а поля основных поляризаций совпадающими. В этом случае квадратурные составляющие отраженного сигнала можно считать нормальными случайными процессами с нулевым средним значением. Задача состоит в том, чтобы определить функциональную связь параметров шероховатости поверхности со статистическими характеристиками отраженного сигнала с учетом влияния параметров движения носителя (влияние скорости, флуктуации углов крена и тангажа).

### РЕШЕНИЕ ЗАДАЧИ

При выполнении условий (1)–(5) отраженный сигнал можно считать нормальным случайным процессом, который полностью характеризуется корреляционными  $B_{11}(\tau) = B_{22}(\tau)$  и взаимокорреляционными функциями  $B_{12}(\tau) = B_{21}(\tau)$  его квадратурных составляющих. В общем виде эти функции описываются выражениями:

$$\begin{aligned} B_{11}(\tau) &= \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} U_1(h_1, h_2) \cdot U_1(h_1^*, h_2^*) \times \\ &\times W_4(h_1, h_2, h_1^*, h_2^*) \cdot dh_1 \cdot dh_2 \cdot dh_1^* \cdot dh_2^*, \end{aligned} \quad (6)$$

$$\begin{aligned} B_{12}(\tau) &= \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} U_1(h_1, h_2) \cdot U_2(h_1^*, h_2^*) \times \\ &\times W_4(h_1, h_2, h_1^*, h_2^*) \cdot dh_1 \cdot dh_2 \cdot dh_1^* \cdot dh_2^*, \end{aligned} \quad (7)$$

где  $U_1(h_1, h_2), U_2(h_1, h_2)$  – квадратурные составляющие отраженного сигнала;  $h_1 = h_1(t)$  и  $h_2 = h_2(t)$  – высота крупной и мелкой, соответственно, неровности рельефа отражающей поверхности в момент времени  $t$ ;  $h_1^* = h_1(t + \tau)$ ; и  $h_2^* = h_2(t + \tau)$ ; – высота крупной и мелкой неровности в момент времени  $(t + \tau)$ ;  $\tau$  – временной сдвиг;  $W_4(h_1, h_2, h_1^*, h_2^*)$  – совместный закон распределения плотности вероятности случайных величин  $h_1, h_2, h_1^*, h_2^*$ . Квадратурные составляющие определяются следующим образом [5]:

$$\begin{aligned} U_1(h) &= Q\sqrt{P} \int_S G_{\text{пм}}(x, y) \cdot G_{\text{пц}}(x, y) \times \\ &\times \cos[4\pi(\Delta R f_1 - h f_1)/c] dx dy; \end{aligned} \quad (8)$$

$$\begin{aligned} U_2(h) &= Q\sqrt{P} \int_S G_{\text{пм}}(x, y) \cdot G_{\text{пц}}(x, y) \times \\ &\times \sin[4\pi(\Delta R f_1 - h f_1)/c] dx dy, \end{aligned} \quad (9)$$

где  $Q$  – постоянная, зависящая от параметров радиолокатора и коэффициента отражения Френеля  $K_{f_0}$ ;  $P$  – излучаемая мощность;  $\Delta R = (R_1 - R_0)$  – разность хода лучей;  $c$  – скорость света;  $R_1$  – текущее расстояние между излучателем и элементом средней плоскости поверхности;  $R_0$  – расстояние между излучателем и центром облучаемой площадки (при вертикальном облучении  $R_0 = H_0$  – высота полета);  $f_1 = [\omega_0 - f'(t)] \cdot (2\pi)^{-1}$ ;  $f(t)$  – модулирующая функция;  $\omega_0$  – круговая частота излучения;  $G_{\text{пм}}(x, y), G_{\text{пц}}(x, y)$  – ДН приемной, передающей антенны. В рассматриваемом случае приемная и передающая антенны полагаются идентичными, осесимметричными, имеющими гауссову аппроксимацию ДН вида:

$$G_{\text{пм}}^2(x, y) \cdot G_{\text{пц}}^2(x, y) = \exp[-5,5(x^2 + y^2)/H_0^2 \cdot \theta_A^2], \quad (10)$$

где  $\theta_A$  – ширина ДН по уровню половинной мощности. Выражения (6), (7) после подстановки (8)–(10) приводятся к виду [6]:

$$B_{11}(\tau) = \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} W_2(h, h^*) \frac{Q^2 P}{2} \int_S \int_S G_{\text{нм}}^2(x, y) \cdot G_{\text{нд}}^2(x, y) \times \\ \times \left\{ \frac{\sin \left[ \frac{4\pi}{c} (\Delta R f_1 - \Delta R^* f_2) - \frac{4\pi}{c} (h f_1 - h^* f_2) \right]}{\cos \left[ \frac{4\pi}{c} (\Delta R f_1 - \Delta R^* f_2) - \frac{4\pi}{c} (h f_1 - h^* f_2) \right]} + \right. \\ \left. + \frac{\sin \left[ \frac{4\pi}{c} (\Delta R f_1 + \Delta R^* f_2) - \frac{4\pi}{c} (h f_1 + h^* f_2) \right]}{\cos \left[ \frac{4\pi}{c} (\Delta R f_1 + \Delta R^* f_2) - \frac{4\pi}{c} (h f_1 + h^* f_2) \right]} \right\} \times \\ \times dx dx^* dy dy^* dh dh^*, \quad (11)$$

где  $f_2 = [\omega_0 - f'(t + \tau)] \cdot (2\pi)^{-1}$ ;  $h = h_1(t) + h_2(t)$ ;  $h^* = h(t + \tau)$ ,  $\Delta R^*$  – значение функции  $\Delta R(t)$  в момент времени  $(t + \tau)$ ;  $x^* = x + \Delta x$ ;  $y^* = y + \Delta y$ .

Двумерная плотность вероятности  $W_2(h, h^*)$ , как плотность вероятности двух независимых случайных функций  $h_1(x, y)$  и  $h_2(x, y)$  с нормальным распределением, также будет подчиняться нормальному закону распределения с корреляционной функцией:

$$B_h(\Delta x, \Delta y) = \sigma_h^2 \cdot \rho_h(\Delta x, \Delta y) = \\ = \sigma_{h_1}^2 \cdot \rho_{h_1}(\Delta x, \Delta y) + \sigma_{h_2}^2 \cdot \rho_{h_2}(\Delta x, \Delta y). \quad (12)$$

Так, как по определению  $\sigma_{h_1}^2 \gg \sigma_{h_2}^2$ , можно записать:

$$\sigma_h^2 \cdot \rho_h(\Delta x, \Delta y) \approx \sigma_{h_1}^2 \cdot \rho_{h_1}(\Delta x, \Delta y). \quad (13)$$

С учетом (13) в результате интегрирования выражение (11) преобразуется к виду:

$$B_{11}(\tau) = \frac{Q^2 P}{2} \int_S \int_S G_{\text{нм}}^2(x, y) \cdot G_{\text{нд}}^2(x, y) \cdot B(\Delta x, \Delta y) \times \\ \times \frac{\sin \left[ \frac{4\pi}{c} (\Delta R f_1 - \Delta R^* f_2) \right]}{\cos \left[ \frac{4\pi}{c} (\Delta R f_1 - \Delta R^* f_2) \right]} dx dx^* dy dy^* + \\ + \frac{Q^2 P}{2} \int_S \int_S G_{\text{нм}}^2(x, y) \cdot G_{\text{нд}}^2(x, y) \cdot B_1(\Delta x, \Delta y) \times \\ \times \frac{\sin \left[ \frac{4\pi}{c} (\Delta R f_1 + \Delta R^* f_2) \right]}{\cos \left[ \frac{4\pi}{c} (\Delta R f_1 + \Delta R^* f_2) \right]} dx dx^* dy dy^*, \quad (14)$$

где

$$B(\Delta x, \Delta y) = \exp\left(-\frac{\sigma_h^2 \cdot \Delta k^2}{2}\right) \cdot \exp\{-\sigma_h^2 k_1 k_2 [1 - \rho_h(\Delta x, \Delta y)]\}; \\ B_1(\Delta x, \Delta y) = \exp\left(-\frac{\sigma_h^2 \cdot \Delta k^2}{2}\right) \cdot \exp\{-\sigma_h^2 k_1 k_2 [1 + \rho_h(\Delta x, \Delta y)]\}; \\ k_1 = 4\pi \cdot f_1 / c; \quad k_2 = 4\pi \cdot f_2 / c; \\ \Delta k = k_1 - k_2 = 4\pi \cdot \Delta F / c, \quad (15)$$

$\Delta F = W\tau / T_M$ ;  $\Delta F = 2W\tau / T_M$  – при модуляции частоты излучения ЧМ-РЛС по пилообразному закону несимметричному и симметричному, соответственно;  $W$  – девиация частоты;  $T_M$  – период модуляции частоты излучаемого сигнала

Выражение (14) можно упростить, пренебрегая вторым слагаемым, поскольку оно сводится к интегралу от гармонической функции с периодом, значительно меньшим, чем период изменения остальных сомножителей подынтегральной функции, а также в связи с малостью величины сомножителя:  $\exp\{-\sigma_h^2 k_1 k_2 [1 + \rho_h(\Delta x, \Delta y)]\}$  подынтегральной функции  $B_1(\Delta x, \Delta y)$ .

В диапазоне СВЧ всегда справедливо соотношение  $\sigma_h \gg \lambda / 4$ , поэтому произведение  $\sigma_h \sqrt{k_1 k_2}$  достаточно велико и множитель  $\exp\{-\sigma_h^2 k_1 k_2 [1 + \rho_h(\Delta x, \Delta y)]\}$  стремится к нулю даже в случае близких к единице значений коэффициента корреляции  $\rho_h(\Delta x, \Delta y)$ . Это позволяет разложить функцию  $\rho_h(\Delta x, \Delta y)$  в ряд Тейлора по степеням отклонений  $\Delta x$  и  $\Delta y$  и отбросить члены высшего порядка малости. Учитывая, равенство нулю первых производных  $(\rho_h)'_x|_{\Delta x=0} = 0$  и  $(\rho_h)'_y|_{\Delta y=0} = 0$ , можно записать:

$$\rho_h(\Delta x, \Delta y) \approx 1 + \frac{1}{2} (\rho_h)''_x \cdot \Delta x^2 + \frac{1}{2} (\rho_h)''_y \cdot \Delta y^2. \quad (16)$$

С учетом (16), выражение (15) приводится к виду:

$$B(\Delta x, \Delta y) = \exp[-2\sigma^2 \pi \cdot \Delta(F/c)] \times \\ \times \exp\left[-\frac{8\pi^2 \sigma_h^2}{\lambda^2} \left(\frac{2}{l_h^2} + \frac{4\pi^2}{\Lambda^2}\right)\right] (\Delta x^2 + \Delta y^2), \quad (17)$$

где поверхность полагается изотропной  $\Lambda = \Lambda_x = \Lambda_y$ .

Так как центр облучаемой площадки совпадает с началом координат, для величины  $(\Delta R f_1 - \Delta R^* f_2)$ , входящей в подынтегральное выражение (14), имеем:

$$(\Delta R f_1 - \Delta R^* f_2) = \frac{x^2 \Delta F}{2H_0} + \frac{y^2 \Delta F}{2H_0} - \frac{f_0}{H_0} \Delta x \cdot x - \\ - \frac{f_0}{H_0} \Delta y \cdot y + \frac{f_0}{H_0} V_x \cdot x \cdot \tau + \frac{f_0}{H_0} V_y \cdot y \cdot \tau, \quad (18)$$

где  $V_x = V_0 \cdot \cos \alpha$ ;  $V_y = V_0 \cdot \sin \alpha$ ;  $V_z = 0$  составляющие вектора скорости носителя ЧМ-РЛС.

В результате подстановки выражений (17) и (18) в формулу (14) и достаточно громоздких вычислений интеграла, функции  $B_{11}(\tau)$  и  $B_{22}(\tau)$  приводятся к виду:

$$B_{11}(\tau) = B_{22}(\tau) = K \cdot A(\tau) \cdot \cos[\varphi(\tau)]; \quad (19)$$

где  $K = \frac{G \cdot A_{\text{эф}} \cdot R_A \cdot P \cdot \theta_A^2 \cdot K_{f_0}^2}{16\pi \cdot H_0^2 (5,5 \cdot a_{\text{м}}^2 + \theta_A^2)}$ ;  $a_{\text{м}}^2 = \sigma_h^2 \left(\frac{4}{j_h^2} + \frac{8\pi^2}{\Lambda^2}\right)$ ;  $G$  – коэффициент усиления антенны;  $A_{\text{эф}}$  – эффектив-

ная поверхность антенны;  $R_A$  – сопротивление антенны;  $a_m^2$  – параметр шероховатости поверхности [5];  $g_1 = 2\pi \times a_m^2 \cdot \xi \cdot H_0 \cdot \Delta F / c$ ;  $g_2 = 2\pi \cdot V_0 \cdot a_m \cdot \sqrt{\xi} / \lambda$ ;  $\xi = \theta_A^2 \div (5,5 \cdot a_m^2 + \theta_A^2)$ .

После аналогичных вычислений выражения для взаимокорреляционных функций квадратурных составляющих приводятся к виду:

$$B_{12}(\tau) = -B_{21}(\tau) = K \cdot A(\tau) \cdot \sin[\varphi(\tau)]. \quad (20)$$

Корреляционная функция комплексной огибающей сигнала определяется следующим образом:

$$B_E(\tau) = B_{11}(\tau) + B_{22}(\tau) + j[B_{12}(\tau) - B_{21}(\tau)]. \quad (21)$$

Тогда с учетом соотношений (19), (20) модуль нормированной корреляционной функции комплексной огибающей отраженного сигнала имеет вид:

$$|\rho(\tau)| = \frac{\sqrt{B_{11}^2(\tau) + B_{12}^2(\tau)}}{B_{11}(0)} = \exp\left[-2\sigma_h^2 \left(\frac{2\pi \cdot \Delta F}{c}\right)^2\right] \cdot (1 + g_1^2)^{-\frac{1}{2}} \exp\left(-\frac{g_2^2 \cdot \tau^2}{1 + g_1^2}\right). \quad (22)$$

Для линейного и квадратичного режима детектирования, используя известные соотношения, можно получить, соответственно, коэффициент корреляции огибающей и коэффициент корреляции квадрата огибающей выходного сигнала:

$$R_E(\tau) = \rho^2(\tau); \quad R_{E^2}(\tau) = \rho^2(\tau) \cdot \pi/4(4 - \pi). \quad (23)$$

Эти коэффициенты отличаются между собой только постоянными множителями. Таким образом, оба режима детектирования практически равноценны в алгоритме обработки отраженного сигнала.

Как следует из выражения (22), точность рассматриваемого корреляционного метода определения неровностей отражающей поверхности определяется вторым и третьим декоррелирующими множителями. Второй множитель учитывает декорреляцию отраженного сигнала, которая обусловлена шириной ДН, а третий множитель, кроме того, учитывает декорреляцию, вызванную движением носителя РЛС. При уменьшении ширины ДН, как видно из полученных соотношений (19), декоррелирующие множители устремляются к единице и точность измерений повышается, поскольку для заданного фиксированного значения  $\Delta F$  величина модуля нормированной корреляционной функции комплексной огибающей отра-

женного сигнала будет определяться только средне-квадратической величиной неровностей рельефа отражающей поверхности.

Используя рассмотренный подход к решению поставленной задачи, можно показать и учесть влияние других параметров движения носителя на точность метода, в частности, случайных флуктуаций углов крена и тангажа.

## ВЫВОДЫ

В результате проведенного анализа получены функциональные связи, являющиеся основополагающими для выбора алгоритма обработки отраженного сигнала. Показано, что в условиях движения носителя декорреляция отраженного сигнала определяется шириной ДН и высотой полета аналогично тому, как это происходит в статических условиях.

Для уменьшения погрешности измерений следует использовать антенны с узкими ДН. С ростом высоты полета требование к сужению ДН ужесточаются, необходимо принятие мер по стабилизации ориентации антенн в пространстве.

Движение носителя вызывает дополнительно декорреляцию отраженного сигнала, что приводит к увеличению общей погрешности. Эта составляющая погрешности также уменьшается при сужении ДНА.

Представляется возможным в дальнейшем путем пространственной фильтрации, обрабатывая не весь спектр сигнала биений частотного высотомера, а только несколько первых гармонических составляющих, которые формируются основной дальномерной зоной, уменьшить влияние ширины ДН и обеспечить стабильную ориентацию угла визирования по нормали к поверхности.

## ПЕРЕЧЕНЬ ССЫЛОК

1. Вопросы перспективной радиолокации : коллективная монография / под ред. А. В. Соколова. – М. : Радиотехника, 2003. – 512 с.
2. Мироненко П. П. Параметрический алгоритм реконструкции эхо-сигналов в задачах локации / П. П. Мироненко, В. А. Безуглый // Аэрофизика и космические исследования : тр. науч. конф. МФТИ, 2006. – М. : Долгопрудный, 2006. – С. 8.
3. Plant W. J. Remote sensing of the sea surface using one and two-frequency microwave techniques / W. J. Plant, D. L. Schuler // Radio Science. – 1980. – V. 15, No. 3. – P. 605615.
4. Некрасов А. В. Измерение скорости и направления ветра над морской поверхностью с борта гидросамолета в целях обеспечения его безопасной посадки на воду / А. В. Некрасов // Информационное противодействие угрозам терроризма. – 2007. – № 9. – С. 1735.
5. Зубкович С. Г. Статистические характеристики радиосигналов, отраженных от земной поверхности / С. Г. Зубкович. – М. : Сов. радио, 1968. – 224 с.
6. Ибрагимов О. М. Использование радиовысотомеров для измерения шероховатости земной поверхности / О. М. Ибрагимов, В. Н. Парфентьев, В. П. Прахов //

Методы обработки сигналов : сб. науч. тр. – Вып. № 94. – Моск. энерг. ин-т. – М. : 1986. – С. 41–44.

Надійшла 27.03.2009  
Після доробки 29.04.2009

*Розглядається задача визначення середньоквадратичної висоти нерівностей рельєфу земної поверхні за допомогою частотного радіо вимірювача. Подані основні результати теоретичного аналізу кореляційної функції*

*відбитого сигналу, отриманої з урахуванням впливу параметрів руху носія.*

*A problem is considered for the terrain roughness root mean square height measurement with the FM radio altimeter. Basic results of theoretical analysis of received signal correlation function in view of the carrier motion parameters influence are presented.*

УДК 621.384.3

С. В. Морщавка, Д. М. Пиза, Е. Л. Белоусов

## ИНФОРМАЦИОННАЯ ТЕХНОЛОГИЯ ОБРАБОТКИ ПРОПАШНЫХ КУЛЬТУР. РАДИОМЕТРИЧЕСКИЕ АСПЕКТЫ<sup>1</sup>

*Проведены полигонные экспериментальные исследования спектральных характеристик растений с высокой разрешающей способностью. Произведен выбор приемлемого алгоритма распознавания. Показано, что предложенный метод дискриминационного анализа позволяет реализовать среднюю вероятность правильной классификации более 80 %. Полученные результаты показывают принципиальную возможность реализации интеллектуальных технологий в сельскохозяйственном производстве.*

### ВВЕДЕНИЕ

Известно, что существующие технологии защиты пропашных культур используют механический способ обработки междурядий и механический метод обработки растений в рядке. При этом распыление гербицидов для уничтожения сорных растений производят непрерывно, независимо от того, есть ли на данном участке рядка сорные растения или нет. Недостатки такой технологии очевидны. Во-первых, неоправданно растет количество используемых гербицидов, что приводит к существенному росту затрат на обработку пропашных культур, а также к тотальному загрязнению окружающей среды (грунты, водоемы, атмосфера Земли). Во-вторых, распыление гербицидов на участках, где нет сорных, но есть полезные растения, приводит к накоплению в продуктах питания применяемых гербицидов и продуктов их разложения.

Вполне очевидно, что при современном развитии информационных технологий, теории распознавания, адаптивных методов управления представляется возможным существенным образом изменить технологию возделывания пропашных культур.

В работах [1–3] предложена и запатентована информационная технология обработки пропашных культур. Предложенная технология базируется на идее распознавания растений с классификацией их на два класса: полезные и сорные. Апостериорная информация, полученная в результате распознавания, используется для реализации адаптивного управления рабочими органами сельскохозяйственной машины. При этом при химическом методе обработки гербициды вносятся только на тех интервалах, где есть сорняки.

### ПОСТАНОВКА ЗАДАЧИ

В работе [1] была разработана методика получения данных о спектральных характеристиках отражения от растений. Она заключается в том, что для получения этих характеристик растение в рядке зондируется с бокового направления световым излучением разных длин волн (рис. 1). Отраженное от растения излучение принимается и обрабатывается приемником. Из рис. 1 следует, что система является би-статической [3]. Пространственное разрешение достигается за счет использования перекрещивающихся диаграмм направленности излучателя и приемника. В ранее опубликованных работах [1, 2] предпринимались попытки оценить как возможность классификации растений, так и определить вероятностные характеристики качества распознавания. Однако, разрешающая способность экспериментальной установки, с помощью которой осуществлялся набор отраженных сигналов, была недостаточной. Поэтому, а также

© Морщавка С. В., Пиза Д. М., Белоусов Е. Л., 2009

1. Статья написана по результатам совместного украинско-египетского научно-технического проекта «Информационная технология обработки пропашных культур с использованием методов распознавания растительных объектов».

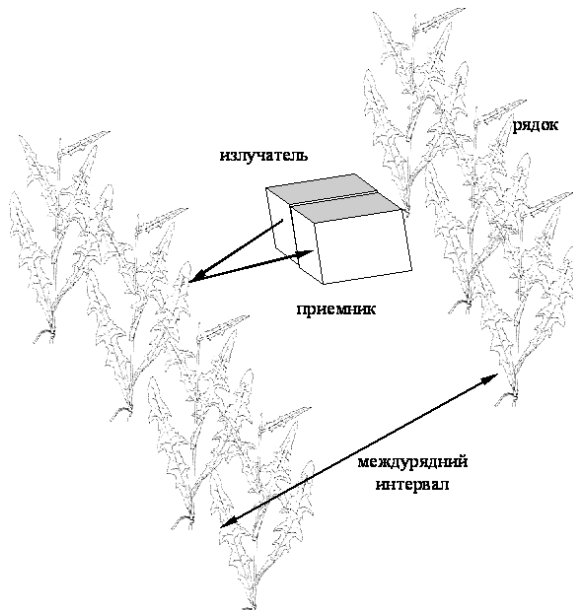


Рисунок 1 – Принцип исследований спектральных характеристик отражения в полевых условиях

в связи с появившейся возможностью повысить разрешающую способность установки до 2 нм, целесообразно провести более точную оценку потенциальных возможностей распознавания растительных объектов.

### АППАРАТНОЕ ОБЕСПЕЧЕНИЕ ИССЛЕДОВАНИЙ

Используемое для полевых исследований спектральных характеристиках растений устройство не обеспечивало работу в режиме реального времени. Однако это было допустимым, поскольку исследуемые характеристики использовались для выбора оптимального алгоритма распознавания. Структурная схема такой установки показана на рис. 2.

Видеокамера использовалась для автоматической маркировки выборок (полезные и сорняки). Необходимость маркировки объясняется тем, что для распознавания классов могут быть использованы алгоритмы с обучением, а для оценки качества – вероятности правильного распознавания, при этом появляется необходимость в априорных знаниях о классе растений.

В качестве датчика изображения использована USB камера – Genius VideoCamEye. При компактных размерах камера обеспечивает разрешающую способность до 640x480 с возможностью программного управления яркостью и контрастностью и выдачей видеоданных в разных форматах.

Датчик местоположения предназначен, с одной стороны, также для маркировки, а с другой стороны, наличие геопространственных отметок позволит в по-

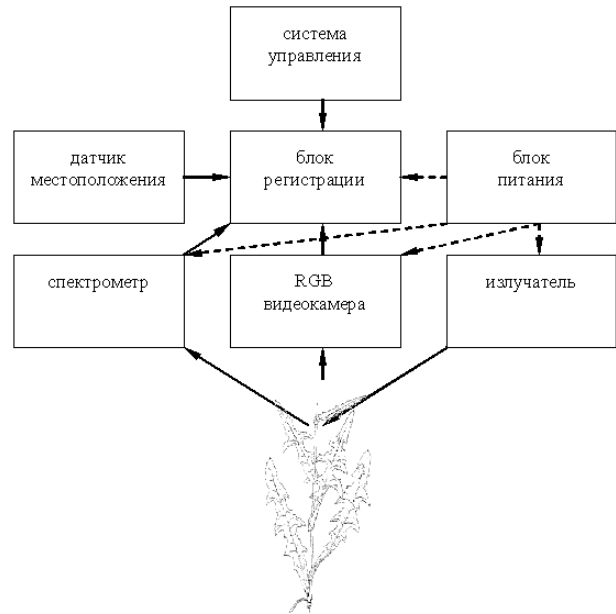


Рисунок 2 – Структурная схема устройства

следующем по результатам экспериментальных исследований создать локальную карту засоренности посевов. Для реализации датчика местоположения использован приемник BT-Q920 фирмы QSatarg. Он имеет компактные размеры и позволяет передавать информацию к блоку регистрации данных с помощью Bluetooth интерфейса.

В качестве излучателя использована компактная галогенная лампа с рефлектором мощностью 20 Вт и напряжением питания 12 В.

Спектрометр USB650 фирмы Ocean Optics необходим для измерения мощности спектральных компонент излучения, отраженных от растений. Кроме самого спектрометра предусмотрены интерфейсы к устройству регистрации и оптическая система для улучшения энергетических характеристик устройства (чувствительности, в первую очередь) и согласования поля зрения с размерами растений.

В оптическую систему вошли: волоконно-оптический кабель и коллимирующая линза. Оптический кабель типа P200-2-UV-VIS (на волокне диаметром 200 мкм), производства фирмы Ocean Optics предназначен для соединения спектрометра с линзой 74-VIS. Рабочий диапазон оптического кабеля и линзы избран таким же, как и у спектрометра – 350–1100 нм.

Устройство регистрации необходимо для сохранения полученных характеристик с дополнительной информацией о виде растения и его местоположении. Устройство питания использовано для автономного питания всех узлов и блоков схемы.

Внешний вид собранного устройства изображен на рис. 3.



Рисунок 3 – Внешний вид устройства

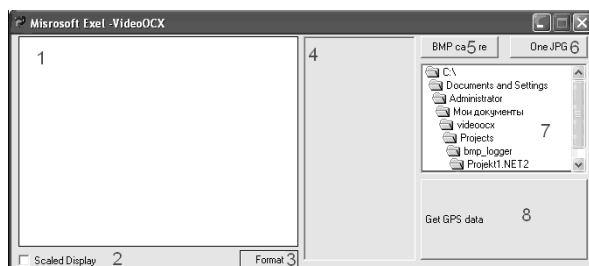


Рисунок 4 – Основное окно программы регистрации данных

При разработке программы регистрации использовались программные компоненты, входящие в базовый комплект программного обеспечения спектрометра USB Red Tide 650 фирмы Ocean Optics. В такой комплект вошли как готовые программы SpectraSuite и OOI\_base32, так и динамическая библиотека, которая содержит модули для управления спектрометром.

Окно, с помощью которого можно управлять программой регистрации информации, изображено на рис. 4.

Оно содержит следующие поля:

- окно видеoinформации (поз. 1);
- поле управления масштабом изображения в поле 1 (поз. 2);
- кнопка вызова окна настройки формата видеoinформации (поз. 3);
- поле для вывода спектрограмм в графическом виде (поз. 4);
- кнопка запуска автоматического сохранения информации (поз. 5);
- кнопка однократного сохранения изображения в формате JFIF (поз. 6);

- поле выбора папки для сохранения файлов в автоматическом режиме (поз. 7);

- кнопка включения/отключения отображения геопространственной информации (поз. 8).

Программная оболочка активизируется лишь двумя кнопками 5 и 8 (если нужно сохранять геопространственную информацию).

## ЭКСПЕРИМЕНТАЛЬНЫЕ ИССЛЕДОВАНИЯ

Экспериментальные исследования проводились в два этапа:

- полевые исследования спектров отражения от растений;
- обработка результатов исследований спектров.

На первом этапе с помощью разработанного и изготовленного устройства для исследования спектров растений измерялись относительные спектральные характеристики в автоматическом режиме с периодом 1 с с сохранением как текущей спектрограммы, так и дополнительной информации о местоположении и внешнем виде растения.

Характеристики растений в строке снимались последовательно, объектив приемника (спектрометрического блока) перемещался вдоль рядка плавно, без рывков и остановок.

Полученные спектральные зависимости содержат в себе, кроме характеристик, которые с точностью до множителя и постоянной составляющей отвечают спектральным характеристикам отражения, еще и информацию о спектральной яркости излучателя и пропускной способности оптического тракта.

Для того чтобы максимально уменьшить влияние сдвига за счет темнового заряда матрицы ПЗС и линейных искажений, проводилась предварительная обработка полученных спектрограмм:

$$K_{\lambda_i n} = \frac{K_{\lambda_i} - \min(K_{\lambda})}{\max(K_{\lambda}) - \min(K_{\lambda})}, \quad (1)$$

где  $K_{\lambda_i n}$ ,  $K_{\lambda_i}$  – нормируемый и измеренный коэффициенты яркости на длине волны  $\lambda_i$ , соответственно;  $\max(K_{\lambda})$ ,  $\min(K_{\lambda})$  – максимальный и минимальный измеренные коэффициенты, соответственно.

Поскольку реальное распределение спектральных коэффициентов яркости растений в видимом и ближнем инфракрасном диапазоне можно достаточно точно описать с помощью нормального распределения, то использование непараметрических методов неоправданно из-за плохого соотношения скорости и качества классификации.

Поэтому был выбран класс алгоритмов, которые относятся к оптимальным за Байесом. Одним из таких классификаторов [4], является классификатор, использующий метрику Махаланобиса. В этом случае

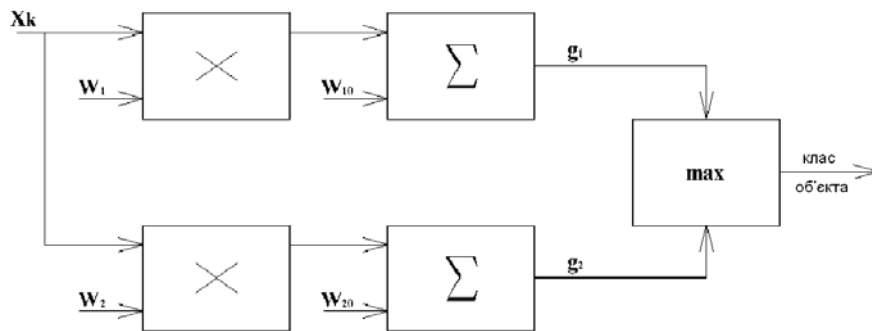


Рисунок 5 – Линейный классификатор с использованием метрики Махаланобиса

расстояние между вектором класса, который классифицируется  $X_k$ , и средним вектором  $U_i$  определяется как:

$$r_i(X_k) = (X_k - U_i)^T \cdot V_i^{-1} (X_k - U_i), \quad i = 1, 2. \quad (2)$$

Если допустить, что ковариационные матрицы для обоих классов (полезное растение и сорняки) одинаковы, то решающие границы будут линейными (для многомерного пространства – определяются гиперплоскостью). Тогда вычисления можно существенно упростить.

После несложных превращений можно выделить дискриминирующую функцию в виде:

$$g_i(X_k) = W_i^t \cdot X_k + W_{i0}, \quad i = 1, 2, \quad (3)$$

где

$$W_i = V^{-1} \cdot U_i, \quad i = 1, 2, \quad W_{i0} = -0,5 \cdot U_i^t \cdot V^{-1} \cdot U_i. \quad (4)$$

Структурная схема такого классификатора, известного, как классификатор по метрике Махаланобиса (рис. 5), аналогична линейному классификатору, при этом используется такое же количество операций и нужен такой же объем памяти. С точки зрения соотношения скорости работы и качества распознавания применение такого классификатора является наиболее приемлемым [2], так как по количеству операций он приближается к линейному классификатору, а качество его работы заметно лучше.

### РЕЗУЛЬТАТЫ ЭКСПЕРИМЕНТАЛЬНЫХ ИССЛЕДОВАНИЙ

Для проведения экспериментальных исследований был избран тестовый полигон, который расположен вблизи г. Запорожья (рис. 6). Полигон представляет собой несколько полей, на которых подсолнечник высевался в разное время и потому присутствовали растения разных сроков вегетации.

В последующем из полученных экспериментальных данных было отобрано 1355 наборов по 2048 спек-

тральным точкам. Критерием отбора была возможность дешифрации вида растений по изображениям и форме спектральной кривой (отсутствие ограничений и избыточной зашумленности).

Виды растений, которые попали в объектив, их полные названия и количество образцов каждого класса приведены в табл. 1.

В отобранных выборках содержится информация о трех классах, два из которых – растения, а третий – чистая почва. Решение не отсеивать выборки третьего класса на первом этапе было вызвано тем, что наличие трех классов отвечает реальной ситуации при работе сельхозтехники. При работе в поле классификация сначала будет осуществляться на две категории (наличие растений первого и второго класса, их отсутствие – третий класс). И только потом, первая категория разделяется на сорняки и полезные растения.

Есть фактор, который налагает ограничение на количество учебных выборок, – необходимость получить невырожденную ковариационную матрицу. Для этого необходимо, чтобы количество выборок было не меньше количества признаков. При этом количество спектральных коэффициентов будет больше или



Рисунок 6 – Тестовый полигон



Таблица 1 – Видовой состав исследуемых образцов

Вид растения	Латинское название	Класс и количество выборок
Подсолнух	Heliantus cultus	1 (полезная) – 401
Амарант	Amaranthus retroflexus	2 (сорняк) – 576
Лебеда	Shenopodium album	
Молочай	Euphorbia falcate	
Мишій	Setaria viridis	
Осот	Sonchus arvensis	
Пырей	Agropyrum repens	
Участки без растений	–	3 (почва) – 378
Вместе:	Всего	1355

равняться квадрату признаков и, если между выборками не будет линейных зависимостей, то можно будет без дополнительных мероприятий найти обратную ковариационную матрицу. Как видно из табл. 1, для всей совокупности сорняков можно использовать не больше 576 признаков, для культурных растений не больше 401 признака.

Поскольку аналогово-цифровой преобразователь спектрометра выдает 2048 спектральных коэффициентов, то практически имеет смысл использовать из них не больше 400. С другой стороны, оптика спектрометра обеспечивает разграничение на уровне 2 нм (650 независимых спектральных коэффициентов). Поэтому было решено уменьшить количество коэффициентов до 256 путем усреднения значений соседних 8 спектральных точек. При этом спектральное разрешение снижается до 3,048 нм и, что существенно, уменьшается влияние шумовых компонент.

Нормированные спектральные выборки коэффициентов яркости были подвергнуты разведывательному анализу для того, чтобы выяснить, как соотносится

предложенное количество классов и их реальное количество. Для этих целей проводилась кластеризация данных с использованием разных метрик и объединительных правил. Результаты кластеризации в виде зависимостей объединительных расстояний от количества классов показаны на рис. 7. Как следует из этих графиков, результаты экспериментальных исследований правильнее всего отнести к трем или четырем классам (наибольшие приросты на графике объединительных расстояний). Это значит, что классификация на три класса может дать хорошие результаты.

Еще одним оценочным методом является кластеризация по методу к-ближайших соседей. Средние значения классов, которые были сформированы с помощью такого метода, изображены на рис. 8. Один из классов, которые образовались, четко отвечает участкам без растительности (Cluster 1), а два других приблизительно отвечают растениям разных классов – сорняков (Cluster 2) и полезных (Cluster 3).

Сверив результаты кластеризации с информацией о внешнем виде растений можно оценить качество

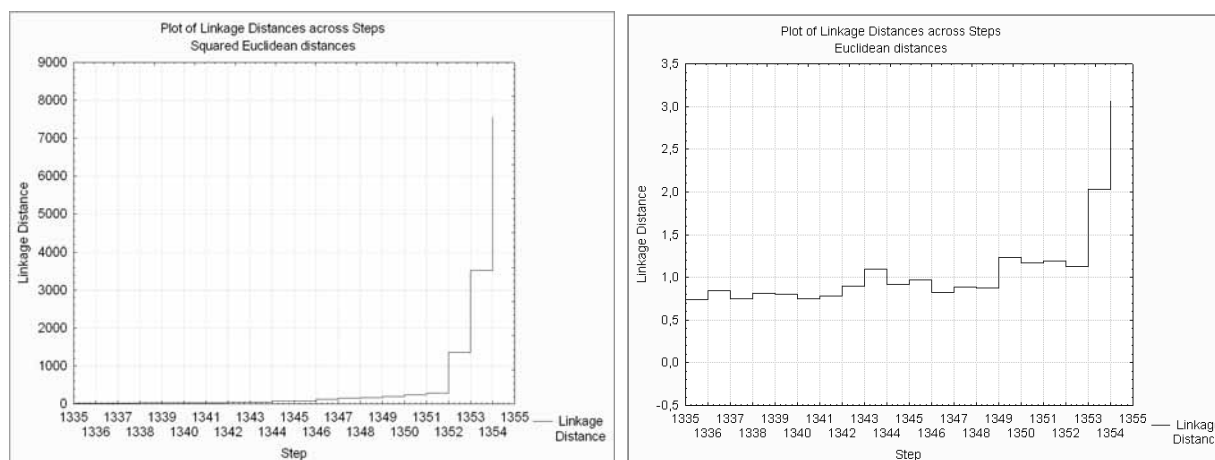


Рисунок 7 – Результаты кластеризации:

а – по правилу Уорда; б – по простому объединительному правилу

Таблиця 2 – Классификационная матрица

Апостериорные данные (в результате классификации)		1(cluster 2)	2 (cluster 3)	3 (cluster 1)	Всего
	1	259	307	6	572
	2	124	255	0	379
3	18	14	372	404	
Априорные данные	Всего	401	576	378	1355
Достоверность правильной классификации		64,5 %	44,3 %	98,4 %	

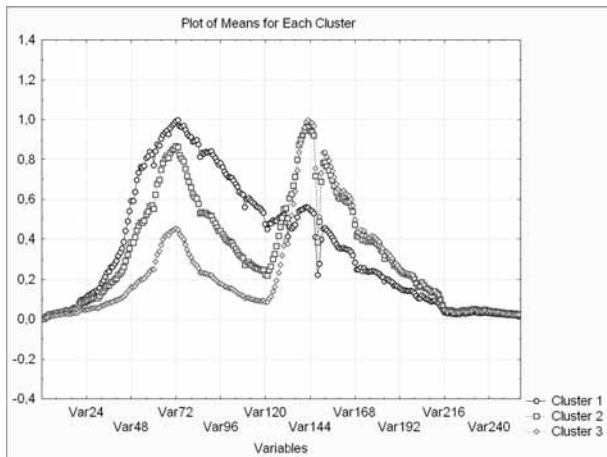


Рисунок 8 – Результаты кластеризации по методу к-ближайших соседей

распознавания «вслепую» без априорных данных о растениях. Результаты проверки показаны в табл. 2. Из таблицы видно, что такое распознавание возможно, но качество будет низким. Исключением является класс, который отвечает участкам без растений с показателем в 98,4 %. Такая точность определения этих участков позволяет разделить одну задачу классификации с тремя классами на две двухклассовые задачи.

Проверка закона распределения вероятности на нормальность была проведена для оценки возможности применения методов дискриминационного анализа, модели Фишера, метрики Махаланобиса. Иногда в задачах классификации, даже если не совсем удов-

летворяются все обязательные условия, экспериментальные данные позволяют достаточно успешно классифицировать выборки [5]. Для проверки возможности классифицировать растительные объекты по экспериментальным данным использован блок дискриминационного анализа программы Statistica. Проведение дискриминационного анализа необходимо для того, чтобы определить возможность классификации растительных объектов по спектральным характеристикам, а также оценить качество распознавания. Сначала, согласно методике проведения анализа, проводился отбор заданного количества наиболее значимых признаков.

Поведение характеристик качества распознавания в зависимости от количества использованных признаков обобщено на рис. 9. Из приведенной иллюстрации видно, что качество распознавания до определенного момента слабо изменяется при уменьшении количества использованных признаков, а затем начинается резкое ухудшение вероятности правильного распознавания.

Исходя из этого, можно определить оптимальное количество признаков для распознавания с точки зрения соотношения скорости и качества обработки – 64 (ближайшая степень двойки). Для улучшения качества распознавания количество признаков следует увеличивать, хотя при этом вероятность правильного распознавания увеличивается практически линейно.

На рис. 10 показаны максимально достигаемые показатели распознавания. Если построить дерево классификации так, чтобы сначала отделить участки без растений (где разбрызгивать гербициды нет необходимости), а затем разделять растения на два класса,

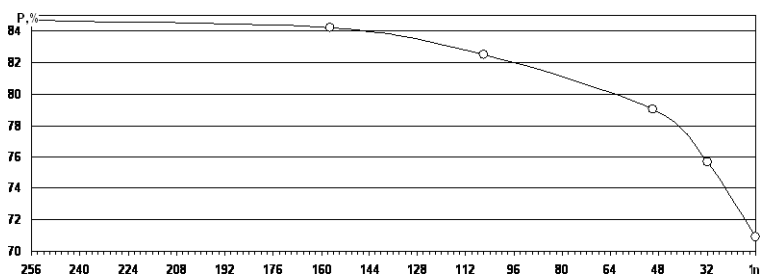


Рисунок 9 – Изменение вероятности правильного распознавания в зависимости от количества признаков

Classification Matrix (data in Workbook2.stw)				
Rows: Observed classifications				
Columns: Predicted classifications				
Class	Percent Correct	1	2	3
		p=,2959	p=,4251	p=,2790
1	73,31671	294,0000	104,0000	3,0000
2	88,02083	63,0000	507,0000	6,0000
3	98,14815	5,0000	2,0000	371,0000
Total	86,49446	362,0000	613,0000	380,0000

Classification Matrix (data in Workbook2.stw)				
Rows: Observed classifications				
Columns: Predicted classifications				
Class	Percent Correct	1	2	
		p=,4104	p=,5896	
1	76,80798	308,0000	93,0000	
2	90,27778	56,0000	520,0000	
Total	84,74923	364,0000	613,0000	

Рисунок 10 – Классификационные матрицы для распознавания по полной системе признаков:

a – на три класса; б – на два класса

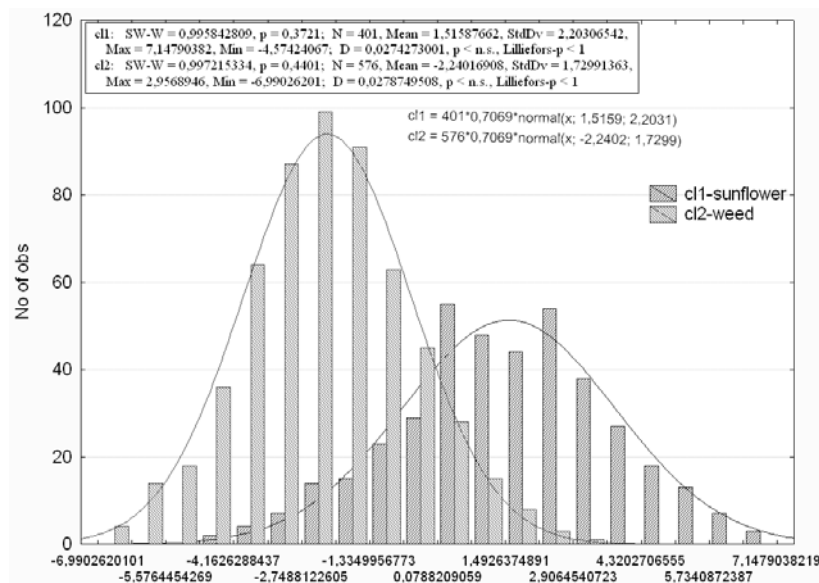


Рисунок 11 – Гистограммы значений дискриминирующих функций

то можно достичь высоких показателей качества распознавания. Участки, на которых отсутствуют растения, будут распознаваться с достоверностью близкой к единице (98,1%), сорняки с достоверностью 90%, а полезные с достоверностью 76,8%.

На практике это значит, что гербициды практически всегда будут экономиться на участках без растений и только 10% сорняков могут не быть обработанными. Однако, количество необработанных сорняков может быть и меньше, поскольку большая часть из этих 10% относится к смешанным случаям, когда в объектив спектрометра попадает отражение от смеси полезных растений и сорняков. Причем соотношение между их вкладами в суммарный сигнал приближается к единице, то есть сорные растения не доминируют. При приближении к участкам, где соотношение смещается в сторону сорняков, классификатор относит их к правильному классу.

В процессе исследований оценены распределения значений дискриминирующих функций для двух классов, которые показаны на гистограммах (рис. 11). Очевидно, что они достаточно точно отвечают нормальному.

Поэтому, можно считать, что использование методов линейного дискриминационного анализа в этой задаче правомерно. Кроме того, классы достаточно четко разделяются в пространстве значений дискриминирующих функций.

## ВЫВОДЫ

Проверка возможности использования линейного дискриминационного анализа (тесты на нормальность распределений та возможность обращения ковариационной матрицы) обнаружила, что не на всех участках спектра распределение можно считать нормальным, но многомерное распределение, в целом, отвечает нормальному закону, что и позволяет получить обратную матрицу.

Исследования с помощью дискриминационного анализа обнаружили принци

пиальную возможность выделения участков без растительности с достоверностью свыше 98% и последующей классификации на полезные и сорные со средней вероятностью правильной классификации 84% (90% для сорняков и 75% для подсолнечни-

ка). Такие результаты показывают принципиальную возможность реализации интеллектуальных сельскохозяйственных технологий, использующих классификацию растений на поле.

### ПЕРЕЧЕНЬ ССЫЛОК

1. Пиза Д. М. Повышение вероятности распознавания растительных объектов за счет увеличения разрешающей способности лидаров / Д. М. Пиза, С. В. Морщавка // Кибернетика и вычислительная техника. – 1999 г. – Вып. 117. – С. 14–17.
2. Пат. 33759 Україна. МПК (2006) А01 В 39/00. Спосіб прополки просапних культур / Пиза Д. М., Морщавка С. В.; заявник та патентовласник Запорізький національний технічний університет; заявл. 03.03.07; опубл. 10.07.08, Бюл. № 13.
3. Пиза Д. М. Выбор наборов признаков для распознавания растительных объектов / Д. М. Пиза, С. В. Морщавка // Системний аналіз, управління і інформаційні технології: Вісник Харківського державного політехнічного університету. Збірка наукових праць. Випуск 97. – Харків: ХДПУ. – 2000. – С. 69–74.
4. Кузнецов А. С. Синтез формы радиолокационного сигнала для распознавания объектов / Кузнецов А. С.

// Зарубежная радиоэлектроника. – 1992. – № 10. – С. 49–60.

5. Боровиков В. П. Популярное введение в программу STATISTICA / Боровиков В. П. – М. : Компьютер Пресс, 1998. – 238 с.

Надійшла 14.05.2009  
Після доробки 18.05.2009

*Проведено полігонні експериментальні дослідження спектральних характеристик рослин з високою роздільною здатністю. Виконано вибір прийняттого алгоритму розпізнавання. Показано, що запропонований метод дискримінаційного аналізу дозволяє реалізувати середню вірогідність правильної класифікації більше ніж 80 %. Отримані результати показали принципову можливість реалізації інтелектуальних технологій в сільськогосподарському виробництві.*

*Proving ground experimental studies of high resolution spectral characteristics of plants have been conducted. Appropriate algorithm of recognition has been successfully selected. It has been shown that the proposed method of discriminatory analysis allows to achieve average probability of proper classification exceeding 80 %. The results show that intellectual technologies can be basically used in agricultural production.*

**МАТЕМАТИЧНЕ  
ТА КОМП'ЮТЕРНЕ  
МОДЕЛЮВАННЯ**

**МАТЕМАТИЧЕСКОЕ  
И КОМПЬЮТЕРНОЕ  
МОДЕЛИРОВАНИЕ**

**MATHEMATICAL  
AND COMPUTER MODELLING**

---

---

УДК 681.391

О. О. Архипова, В. М. Журавльов

**ЧАСТОТНИЙ АНАЛІЗ ВИКОРИСТАННЯ БУКВ УКРАЇНСЬКОЇ МОВИ**

---

*В статті розглянуті невирішені задачі в галузі дослідження якості каналів мовної комунікації. Проведено частотний аналіз використання букв української мови для текстів загальним обсягом біля 580 тисяч знаків художнього, публіцистичного та технічного спрямування. Вперше отримано гістограму частот використання букв алфавіту для сучасної української мови.*

**ВСТУП. ПОСТАНОВКА ЗАДАЧІ**

Якість передачі мовлення – це одна із головних характеристик каналу мовного зв'язку. Згідно [1], головними критеріями якості каналів мовної комунікації є:

- 1) розбірливість (зрозумілість, ясність);
- 2) гучність (голосність);
- 3) природність (натуральність).

*Розбірливість*, безперечно, можна назвати головним параметром, оскільки вона відображає виконання системою прийому-передачі мови свого головного призначення – забезпечення того, щоб слухач правильно зрозумів зміст переданого.

*Гучність* – визначає бажаний рівень прийнятих сигналів, який для оптимальних умов має бути таким, щоб не викликати стомленості та перенапруження слухового апарата. Цей параметр не є само-

статнім і використовується разом із першим, а в умовах використання спеціальної техніки, що регулює гучність, втрачає сенс.

*Природність* – оцінює здатність системи відтворювати не тільки зміст мови, що передається, але й її індивідуальні особливості, притаманні різним мовцям. Цей параметр не такий важливий як розбірливість. Виключенням є випадки спеціальних систем зв'язку, наприклад, систем, у яких потрібне визначення особи (ідентифікація) мовця за голосом, або для художнього відтворення мови та музики.

Усі відомі на цей час методи оцінки якості передачі мовлення можуть бути розділені на дві великі групи: *суб'єктивні експертні методи* і *об'єктивні методи* [1].

Артикуляційні випробування є суб'єктивним методом оцінки розбірливості – це найбільш прямий й очевидний, а іноді й єдиний, шлях дослідження якості каналу мовного зв'язку. Головними перевагами методу артикуляційних випробувань є його універсальність та відносна простота. Однак процедура організації суб'єктивних експертиз за оцінкою розбірливості мови – справа громіздка, тривала й досить дорога.

Серйозною й самостійною складною проблемою методу артикуляційних випробувань є створення спеціальних артикуляційних таблиць. Як показує практика вимірювань, тип таблиць, що використовуються, істотно впливає на результати вимірювань. Артикуляційні таблиці складаються за певними правилами [2]. Ці правила враховують лінгвістичні (мовні) і технічні вимоги до таблиць. Лінгвістичні вимоги полягають у тому, щоб таблиці достатньою мірою відображали фонетичну структуру мови. Технічні вимоги передбачають забезпечення максимальної економності під час виконання вимірів, мінімальної надмірності із максимально можливою однорідністю для того, щоб зменшити розкид результатів одиночних вимірів. Поєднання в таблицях лінгвістичних і технічних вимог можливе тільки у разі розумного компромісу, оскільки вони взаємно суперечливі.

Для української мови не складено артикуляційних таблиць, їх складання є актуальною науково-технічною задачею. Для забезпечення лінгвістичних вимог до артикуляційних таблиць для української мови необхідним є буквенний та, у подальшому, фонемний частотні аналізи.

### ЧАСТОТНИЙ АНАЛІЗ ВИКОРИСТАННЯ БУКВ УКРАЇНСЬКОЇ МОВИ

Кількість різних букв, як і фонем, у кожній мові обмежена. Важливими характеристиками мови є по-

вторюваність букв (монограм), пар букв (біграм) і взагалі  $m$ -грам, сполучуваність букв одна з одною, чергування голосних і приголосних тощо. Примітно, що ці характеристики є досить стійкими [3].

Якщо  $\vartheta(a_{i_1}a_{i_2}\dots a_{i_m})$  – кількість появ  $m$ -грами  $a_{i_1}a_{i_2}\dots a_{i_m}$  у тексті  $T$ , а  $L$  – загальне число підрахованих  $m$ -грам, то при досить великих  $L$  частоти

$$\frac{\vartheta(a_{i_1}a_{i_2}\dots a_{i_m})}{L} \quad (1)$$

для даної  $m$ -грами мало відрізняються одна від одної.

Виходячи з цього, відносну частоту (1) вважають наближеною ймовірністю  $P(a_{i_1}a_{i_2}\dots a_{i_m})$  появи даної  $m$ -грами у випадково обраному місці тексту (за статистичним визначенням імовірності).

Частотний аналіз використання букв проведений для ряду європейських мов, його результати наведені у книзі [4]. Необхідно зазначити, що частота використання букв для французької, німецької, англійської іспанської та італійської мов різна. Деяка різниця значень частот у таблицях, які наводяться з різних джерел, пояснюється тим, що частоти істотно залежать не тільки від довжини тексту, але й від його характеру. Наприклад, у технічних текстах рідка буква  $\Phi$  може стати досить частою у зв'язку із частим використанням таких слів, як функція, диференціал, дифузія, коефіцієнт і т. п.

Таблиця 1 – Частоти використання букв російської мови

–	О	Е, Ё	А	И	Т	Н	С
0,175	0,09	0,072	0,062	0,062	0,053	0,053	0,045
Р	В	Л	К	М	Д	П	У
0,04	0,038	0,035	0,028	0,026	0,025	0,023	0,021
Я	Ы	З	Ь, Ъ	Б	Г	Ч	Й
0,018	0,016	0,016	0,014	0,014	0,013	0,012	0,01
Х	Ж	Ю	Ш	Ц	Щ	Э	Ф
0,009	0,007	0,006	0,006	0,004	0,003	0,003	0,002

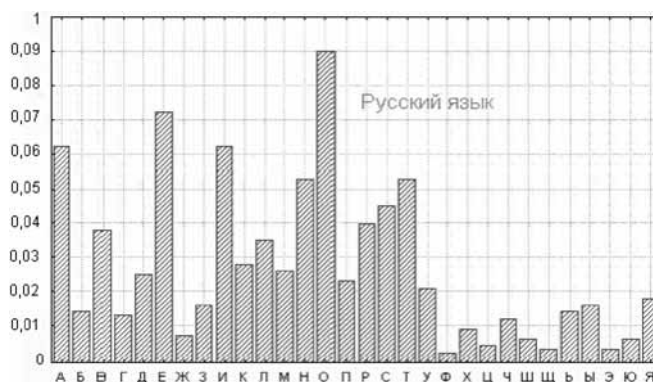


Рисунок 1 – Гістограма частот використання букв алфавіту російської мови

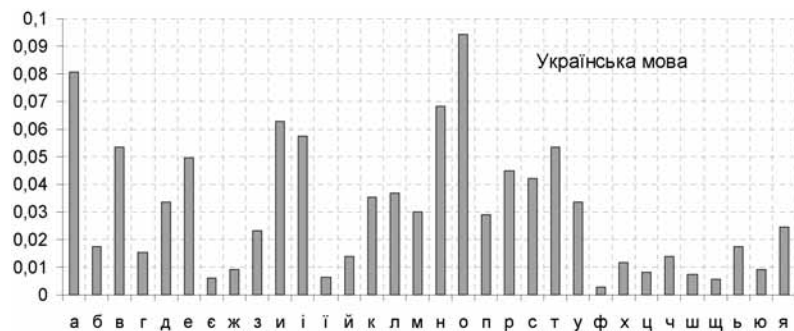


Рисунок 2 – Гістограма частот використання букв алфавіту української мови

Таблиця 2 – Осереднені частоти та дисперсії використання букв української мови

	Технічні тексти		Вірші		Гуманітарні твори	
	частота	дисперсія	частота	дисперсія	частота	дисперсія
А	0,0709	2,350E-05	0,0871	1,564E-04	0,0840	2,870E-05
Б	0,0136	1,092E-05	0,0211	6,685E-08	0,0183	4,918E-06
В	0,0533	2,066E-05	0,0468	1,573E-05	0,0604	9,286E-05
Г	0,0142	7,898E-06	0,0150	1,450E-05	0,0173	2,897E-06
Д	0,0350	3,305E-05	0,0332	9,669E-07	0,0332	3,025E-06
Е	0,0458	7,780E-05	0,0544	3,143E-07	0,0483	3,821E-05
Ж	0,0066	2,227E-06	0,0108	1,457E-06	0,0106	3,962E-06
З	0,0238	2,096E-05	0,0220	3,176E-06	0,0239	7,000E-06
И	0,0613	3,346E-05	0,0633	3,170E-05	0,0632	9,140E-07
Й	0,0114	8,932E-06	0,0162	4,442E-08	0,0139	7,501E-06
К	0,0358	1,231E-05	0,0330	1,674E-07	0,0373	4,538E-07
Л	0,0305	8,335E-06	0,0432	7,376E-06	0,0370	2,138E-05
М	0,0283	3,545E-05	0,0328	4,833E-08	0,0297	8,720E-06
Н	0,0836	1,155E-05	0,0562	6,087E-06	0,0645	3,203E-05
О	0,0950	2,055E-05	0,0920	3,702E-05	0,0956	9,878E-06
П	0,0304	3,369E-05	0,0264	1,494E-06	0,0303	2,258E-06
Р	0,0499	5,291E-05	0,0400	4,206E-06	0,0445	9,543E-06
С	0,0406	4,418E-05	0,0458	4,125E-06	0,0407	6,013E-06
Т	0,0546	8,941E-06	0,0576	6,375E-06	0,0483	5,564E-05
У	0,0321	2,617E-05	0,0333	1,083E-06	0,0353	5,434E-06
Ф	0,0061	1,096E-05	0,0004	1,239E-08	0,0020	1,871E-06
Х	0,0129	1,185E-05	0,0123	1,771E-06	0,0105	5,424E-07
Ц	0,0108	8,155E-06	0,0061	2,982E-06	0,0080	6,574E-06
Ч	0,0136	7,865E-06	0,0144	4,283E-09	0,0144	3,963E-06
Ш	0,0060	2,921E-06	0,0083	7,629E-07	0,0085	1,252E-06
Щ	0,0040	1,147E-06	0,0063	3,813E-06	0,0065	8,882E-07
І	0,0628	3,037E-05	0,0530	2,239E-04	0,0568	2,204E-05
Ї	0,0076	4,060E-06	0,0060	1,413E-06	0,0058	2,993E-06
Ь	0,0167	7,304E-06	0,0211	2,969E-06	0,0153	6,865E-06
Є	0,0080	5,295E-06	0,0062	1,243E-06	0,0042	3,730E-06
Ю	0,0082	4,378E-06	0,0116	4,322E-08	0,0081	1,980E-06
Я	0,0268	5,537E-05	0,0240	2,921E-06	0,0236	3,870E-06

Таблиця 3 – Ранжовані частоти використання букв української мови

О	0,0942	р	0,0448	я	0,0248	ж	0,0093
А	0,0807	с	0,0424	з	0,0232	ю	0,0093
Н	0,0681	л	0,0369	б	0,0177	ц	0,0083
И	0,0626	к	0,0354	ь	0,0177	ш	0,0076
І	0,0575	д	0,0338	г	0,0155	ї	0,0065
В	0,0535	у	0,0336	ч	0,0141	є	0,0061
Т	0,0535	м	0,0303	й	0,0138	щ	0,0056
Е	0,0495	п	0,0290	х	0,0119	ф	0,0028

Ще більші відхилення від норми в частоті вживання окремих букв спостерігаються в деяких художніх творах, особливо у віршах. Тому для надійного визначення середньої частоти букв бажано мати набір різних текстів, запозичених з різних джерел. Разом із тим, як правило, подібні відхилення незначні і, в першому наближенні, ними можна знехтувати.

Для російської мови частоти знаків алфавіту (у порядку зменшення), де ототожнено Е з Ё, Ь з Ъ, а також є знак пробілу (–) між словами, наведені в табл. 1 (див. [5]), або у вигляді наочної діаграми, приведеної на рис. 1.

Нами був проведений частотний аналіз повторюваності букв української мови за допомогою програми, написаної на мові програмування С++ та пакету Excel. У ході аналізу оброблено біля 580 тисяч знаків українських текстів. Серед них 260 тисяч – сучасні тексти технічного спрямування (роботи та статті по захисту інформації, математичній статистиці, диференційним рівнянням) 76 тисяч знаків – вірші (Т. Шевченко та сучасна поетеса Н. Доценко) та 253 тисяч – відома художня проза та публіцистика (твори Д. Лондона, Е. По, А. Гофмана, газета «Дзеркало тижня»). Осереднені частоти та дисперсії використання букв української мови наведені у табл. 2.

У табл. 3 містяться середні ймовірності появи букв українського алфавіту ранжовані у порядку спадання, а на рис. 2 – діаграма частот використання букв алфавіту української мови. У зв'язку із тим, що у багатьох текстах не розрізняються літери Г і Ґ, їх було ототожнено.

Порівнюючи гістограми частот використання букв російської, української та європейських мов [4], помітно різницю у їх розподілі, яка досягає десяти відсотків. Даний факт свідчить про ймовірне виникнення методичної похибки визначення розбірливості за словами та складами при використанні російськомовних артикуляційних таблиць для оцінки якості україномовних каналів мовної комунікації.

## ВИСНОВКИ

1. Вперше сформульована науково-технічна задача створення артикуляційних таблиць для української мови.

2. Виконано необхідний проміжний етап цієї задачі, що забезпечує лінгвістичні вимоги до артикуляційних таблиць (відображення структури мови), – побудовано гістограму частот використання букв алфавіту сучасної української мови.

3. Отримана діаграма дозволяє скласти кілька артикуляційних таблиць з різним характером наборів слів (в залежності від тематичного спрямування) для проведення подальших артикуляційних випробувань. У майбутньому необхідно зробити фонемний частотний аналіз, також бажаним є мовний аналіз біграм і триграм.

## ПЕРЕЛІК ПОСИЛАНЬ

1. Покровский Н. Б. Расчет и измерение разборчивости речи / Н. Б. Покровский. – М. : Связьиздат, 1962. – 392 с.
2. Вемян Г. В. Передача речи по сетям электросвязи / Г. В. Вемян. – М. : Радио и связь, 1985. – 272 с.
3. Обмен опытом [Электронный ресурс]. – Режим доступа: <http://www.statsoft.ru/home/portal/exchange/textanalysis.htm>. – Назва з екрана.
4. Baudouin C. Elements de cryptographie / C. Baudouin, Ed. A. Pedone. – Paris, 1939. – 214 p.
5. Яглом А. М. Вероятность и информация / А. М. Яглом, И. М. Яглом. – М. : Наука, 1973. – 374 с.

Надійшла 16.02.2009  
Після доробки 27.04.2009

*В статье рассмотрены нерешенные задачи в области исследования качества речевой коммуникации. Проведен частотный анализ использования букв украинского языка для текстов общим объемом 580 тысяч знаков художественного, публицистического и технического характера. Впервые получено гистограмму частот использования букв алфавита для современного украинского языка.*

*In this paper is considered unresolved tasks in the research field of the communications speech quality. The frequency analysis of the Ukrainian letters usage for total amount of texts about 580 thousand signs in art, publicistic and technical area is carried out. For the first time it is received the frequency histogram of letters usage of the Ukrainian alphabet for modern language.*



И. В. Гребенник, А. Ю. Хабаров, Ю. А. Гриценко, Д. В. Иванов

## ОПТИМИЗАЦИЯ ПУТИ СОЕДИНЕНИЯ ДВУХ ТОЧЕК ПРОВОДНОЙ СЕТИ ЭЛЕКТРОСВЯЗИ ПРИ ИНТЕРВАЛЬНОЙ НЕОПРЕДЕЛЕННОСТИ ДАННЫХ

*Анализируется задача выбора наилучшего соединения двух точек проводной сети электросвязи по множеству критериев в условиях неопределенности. Строится интервальная многокритериальная математическая модель задачи, приводится метод ее решения.*

### ВВЕДЕНИЕ

Проводная сеть электросвязи как объект является сложной технической системой, бесперебойное и качественное функционирование которой обеспечивается постоянным решением множества задач самого разного уровня сложности и компетенции лица, принимающего решение (ЛПР). Одной из наиболее часто решаемых задач является определение технической возможности организации линии определенного качества между двумя точками сети электросвязи и выбора наилучшего (в смысле используемых критериев) пути для ее организации, если такая возможность есть [1].

С другой стороны, особенности реальной сети электросвязи связаны с высокой неопределенностью значений многих ее параметров. Это обусловлено невозможностью проведения точных измерений ряда характеристик сети. Во многих случаях информация о состоянии сети, содержащаяся в технической документации, стареет и в значительной степени отличается от описания реальной ситуации. Текущие изменения распределительной сети, связанные с ежедневной деятельностью линейных монтажников, не всегда своевременно и в полном объеме отражаются в технической документации, что существенно снижает общую достоверность сведений о состоянии сети [2, 3]. Указанные особенности, связанные с недостаточной информированностью ЛПР о реальном состоянии сети электросвязи, необходимо учитывать при проведении расчетов. Это в полной мере относится к решению задачи оптимизации пути соединения двух точек сети.

Для учета неопределенности при проведении расчетов на телефонной сети могут быть использованы различные средства. Одним из них является интервальный анализ [4], хорошо зарекомендовавший себя при математическом моделировании и решении задач различных классов в условиях неопределенности.

Настоящая статья посвящена продолжению исследований по оптимизации путей соединения двух точек сети электросвязи, описанных в [1], в случае учета неопределенности данных методами интервального анализа. Ее целью является решение задачи выбора оптимального пути соединения двух точек сети электросвязи при интервальной неопределенности исходных данных. Здесь и далее под интервальной неопределенностью понимается случай, когда некоторые исходные данные задачи заданы в интервальном виде.

### ПОСТАНОВКА ЗАДАЧИ

Общую задачу построения соединения сформулируем следующим образом. Построить соединение заданного типа между двумя точками сети, удовлетворяющее набору ограничений и наиболее эффективное в смысле выбранных критериев  $\{k_1, k_2, \dots, k_n\}$ . При этом значения критериев известны с точностью до некоторых интервалов  $[a_i, b_i] \subset R^1, i \in J_n = \{1, 2, \dots, n\}$ .

В статье [1] данная задача решалась для случая, когда значения всех критериев эффективности точно известны. В качестве модели телефонной сети  $G$  предложено использовать объединение трех графов – графа капитальных сооружений ( $G_1$ ), кабельного графа ( $G_2$ ) и сигнального графа ( $G_3$ ),  $G = G_1 \cup G_2 \cup G_3$ . Точкам сети электросвязи соответствуют вершины графа  $G_i, i \in J_3$ , а соединению точек – простая цепь в графе  $G_i$ , соединяющая две заданные вершины.

Система критериев оценки соединений, предложенная в [1], содержит электрические характеристики ребер и соединений, длины ребер, количество точек коммутации для соединения. При этом необходимо соблюдать ограничения технологического и экономического характера, подробно описанные в [1]. Далее будем полагать, что значения критериев точно не известны и находятся внутри некоторых интервалов.

Рассмотрим на примере сигнального графа  $G_3$  следующие варианты задачи построения эффективного соединения между двумя вершинами сети:

а) соединить эффективным образом две заданные вершины сигнального графа, добавив отдельные ребра в допустимых местах. В терминах предметной области постановка задачи звучит так: организовать

наиболее качественную в смысле выбранных интервальных критериев связь между двумя контакторами с учетом ограничений, при необходимости устанавливая кроссировки;

б) для заданной начальной вершины сигнального графа определить конечную вершину таким образом, чтобы для них существовало эффективное в соответствии с интервальными критериями соединение, а также собственно путь между ними с учетом ограничений. В терминах предметной области постановка задачи звучит так: определить возможность и трассу подключения абонента к одной из телефонных станций, имеющихся в сети, при этом трасса должна обладать характеристиками не хуже заданных;

в) для двух заданных вершин сигнального графа, для которых уже существует соединение, используя интервальные критерии, построить более эффективное, которое может включать в себя отдельные фрагменты прежнего соединения. В терминах предметной области постановка задачи звучит так: определить возможность улучшения характеристик рабочей линии до заданных величин путем расчета другой схемы коммутации с учетом ограничений.

Как видно из приведенных вариантов постановки задачи, их различие состоит только в способе задания соединяемых вершин (в случае (б) требуется предварительный расчет конечной вершины) и ограничений на характеристики проектируемого пути их соединения (в случае (в) они задаются с помощью значений соответствующих характеристик реально существующей линии).

Для решения сформулированных задач с другими типами соединений можно использовать аналогичный подход, реализуя подобные методы на соответствующих графах с применением выбранных критериев и учетом ограничений.

Поставленная задача на графе  $G_3$  решается в несколько последовательных этапов:

1) предварительный (для задачи (б)). Определить по заданной начальной вершине множества вершин, обладающих свойствами, необходимыми для конечной вершины;

2) построение множества альтернатив соединения начальной и конечной вершин с возможным добавлением ребер-кроссировок в допустимых местах с учетом ограничений. Для задачи (б) по выбору ЛПР обрабатываются все допустимые пары или конкретная пара таких вершин;

3) определение наилучшего пути соединения начальной и конечной вершин в смысле системы интервальных критериев (с одновременным выбором конечной вершины для задачи (б), если она не была выбрана ЛПР на этапе 2).

## ЭТАПЫ ОПТИМИЗАЦИИ

Рассмотрим подробнее указанные этапы для графа  $G_3$ .

**Этап 1** может быть описан следующей последовательностью операций поиска конечной вершины:

а) проверить, удовлетворяет ли текущая вершина требованиям, предъявляемым для конечной вершины. Если да, то запомнить ее;

б) найти все вершины, достижимые из текущей после установки ребер-кроссировок и допустимые для использования в соединении;

в) найти все вершины, достижимые из текущей по кабельным сигнальным связям, допустимым для использования в соединении;

г) для найденных на шагах (б) и (в) вершин выполнить операции (а)–(г).

Результатом этапа является множество вершин графа  $G_3$ , которые могут выступать в качестве конечной для искомого соединения. ЛПР может выбрать одну или несколько из них для дальнейшего рассмотрения.

**Этап 2** представляет собой построение множества путей (альтернатив) для последующего выделения на этапе 3 лучшего пути. В связи с чрезвычайно большой размерностью задачи нецелесообразно сразу выделять все допустимые пути. Рациональнее применить эвристику, суть которой состоит в следующем. Не анализируя качественных характеристик отдельных ребер сигнального графа, следует определиться с «направлениями» возможных путей. Под направлением понимается подмножество ребер графа  $G_3$ , соединяющее два подмножества вершин таких, что они могут быть сгруппированы по признаку соответствия вершине кабельного графа  $G_2$ . В терминах предметной области направлением является группа сигнальных связей (реализуемых кабелями между двумя кабельными терминаторами) между контактами, входящими в контакторы, принадлежащие указанным кабельным терминаторам. Поиск направлений происходит следующим образом:

а) проверить, входит ли конечная вершина в текущий контактор. Если да, то запомнить упорядоченное множество направлений до нее;

б) найти все контакторы, достижимые от текущего с помощью установки допустимых ребер-кроссировок;

в) найти все контакторы, достижимые от текущего с помощью ранее не анализированных направлений;

г) для найденных на шагах (б) и (в) контакторов выполнить операции (а)–(г).

Результатом этапа является множество направлений, из элементов которых могут быть построены допустимые соединения.

**Этап 3** предназначен для итоговой оценки и сравнения альтернатив с целью поиска наилучшего соединения или соединения с характеристиками лучше

заданных. Поскольку альтернативы представлены множеством направлений, то задачи поиска наилучшего пути в смысле системы интервальных критериев и поиска пути с характеристиками лучше заданных отличаются по своим постановкам и подходам к решению.

Для решения первой из задач применим следующий подход.

С целью математического моделирования задачи поиска наилучшего пути воспользуемся элементами теории интервального анализа и подходом, описанным в [5].

Пусть  $X$  – множество всех путей, соединяющих две выбранные вершины сети,  $\mathbf{K} = \{\mathbf{k}_1(x), \mathbf{k}_2(x), \dots, \mathbf{k}_n(x)\}$  – множество интервальных отображений [4] вида  $\mathbf{k}_i: X \rightarrow \mathbf{E}_i \subset \mathbf{I}_s\mathbf{R}$ ,  $i \in J_n$ . Здесь  $\mathbf{I}_s\mathbf{R} = I_sR \cup \overline{I_sR}$  – пространство центрированных интервалов, где

$$I_sR = \left\{ \langle X \rangle = \langle x, v_x \rangle \mid x = \frac{a+b}{2} \in R^1, v_x = \frac{b-a}{2} \in R^1, [a, b] \subset R^1 \right\},$$

$$\overline{I_sR} = \{ \langle \bar{X} \rangle = \langle x, -v_x \rangle \mid \forall \langle X \rangle \in I_sR \},$$

$$\mathbf{E}_i = \{ \langle X \rangle \in \mathbf{I}_s\mathbf{R} : x \in R^1, |v_x| \leq \delta_i \},$$

где  $\delta_i$  – верхняя оценка «неопределенности» задания критерия  $k_i(x)$ .

Тогда задачу выбора единственного варианта организации соединения из множества  $X$  по множеству критериев в условиях интервальной неопределенности можно сформулировать в виде следующей оптимизационной задачи:

$$x^0 = \arg \operatorname{extr}_{x \in X} \{ \mathbf{k}_1(x), \mathbf{k}_2(x), \dots, \mathbf{k}_n(x) \}, \quad (1)$$

где экстремум понимается в смысле отношения порядка и способа определения максимума и минимума в пространстве  $\mathbf{I}_s\mathbf{R}$ .

**Замечание.** В случае, когда для всех  $\mathbf{k}_i(x) = \langle k_i(x), v_{k_i(x)} \rangle$  выполняется соотношение  $v_{k_i(x)} = 0$ , задача (1) представляет собой известную многокритериальную задачу выбора единственного элемента из множества  $X$ .

Отношение порядка в пространстве  $\mathbf{I}_s\mathbf{R}$  вводится следующим образом [6]:

$$\begin{aligned} \forall \langle X \rangle = \langle x, v_x \rangle \in \mathbf{I}_s\mathbf{R}, \forall \langle Y \rangle = \langle y, v_y \rangle \in \mathbf{I}_s\mathbf{R}, \\ (\langle X \rangle < \langle Y \rangle) &\Leftrightarrow ((x < y) \vee (x = y) \wedge (v_x < v_y)), \\ (\langle X \rangle > \langle Y \rangle) &\Leftrightarrow ((x > y) \vee (x = y) \wedge (v_x > v_y)), \\ (\langle X \rangle > \langle Y \rangle) &\Leftrightarrow ((x > y) \vee (x = y) \wedge (v_x > v_y)), \\ (\langle X \rangle = \langle Y \rangle) &\Leftrightarrow ((x = y) \wedge (v_x = v_y)). \end{aligned} \quad (2)$$

На основе (2) минимум из  $n$  интервальных чисел  $\langle Z_1 \rangle, \langle Z_2 \rangle, \dots, \langle Z_n \rangle$  определяется так:

$$\langle Z^* \rangle = \min \{ \langle Z_1 \rangle, \langle Z_2 \rangle, \dots, \langle Z_n \rangle \}, \quad (3)$$

где

$$\langle Z^* \rangle = \langle z^*, v_z^* = \min_{i \in J_n} v_{z_i} \rangle, \\ \text{если } z_1 = z_2 = \dots = z_n = z^*,$$

$$\langle Z^* \rangle = \langle z^* = z_{i_j} = \min_{i \in J_n} z_{i_j}, v_{z^*} = \min_{k \in J_r} v_{z_{i_k}} \rangle, \quad i_j \in J_n, j \in J_n, \\ \text{если } z_{i_j} \neq z_{i_k}, i \in J_k, j \neq k,$$

$$\langle Z^* \rangle = \langle z^* = \min_{i \in J_n} z_{i_j}, v_z^* = v_{z_{i_j}} = \min_{k \in J_r} v_{z_{i_k}} \rangle, \\ \text{если } z_{i_1} = z_{i_2} = \dots = z_{i_r} = z^*,$$

$$\{z_{i_1}, z_{i_2}, \dots, z_{i_r}\} \subset \{z_1, z_2, \dots, z_n\}, \quad 1 \leq r < n.$$

Аналогичным образом определяется максимум из  $n$  интервалов. Данный подход к определению минимума (максимума) распространяется и на случай бесконечного множества центрированных интервалов.

Для поиска *наилучшего пути* можно использовать один из способов:

**Способ 1.** С использованием всех выбранных направлений строятся множество альтернатив  $X$ , в которое включаются все возможные варианты соединения заданной пары вершин, после чего все пути  $x \in X$  оцениваются по интервальным критериям  $\{\mathbf{k}_1(x), \mathbf{k}_2(x), \dots, \mathbf{k}_n(x)\}$ .

Суть решения задач оптимизации с интервальными функциями цели состоит в уменьшении диаметра (радиуса) интервала – значения функции цели. Для перехода к задачам минимизации интервальных функций введем интервальные отображения полезности  $\mathbf{p}_i(x)$  частных критериев  $\mathbf{k}_i \in \mathbf{K}$ ,  $i \in J_n$ , задачи (1) с целью нормализации интервальных критериев, следуя [6]. Интервальное отображение полезности  $\mathbf{p}_i(x)$  частного критерия  $\mathbf{k}_i(x)$  должно удовлетворять следующим требованиям:

- 1)  $\mathbf{Y}_i = \{ \mathbf{p}_i(x) \mid \langle 0, 0 \rangle \leq \mathbf{p}_i(x) \leq \langle 1, v_{p_i} \rangle, x \in \mathbf{X} \}$ ;
- 2)  $\mathbf{p}_i(x)$  инвариантно к размерности частного критерия  $\mathbf{k}_i(x)$ ;
- 3)  $\mathbf{p}_i(x)$  инвариантно к виду экстремума частного критерия  $\mathbf{k}_i(x)$ .

Последнее требование означает, что независимо от вида экстремума (минимум или максимум) частного критерия  $\mathbf{k}_i(x)$ , его наилучшему значению на множестве  $\mathbf{X}$  должен соответствовать максимальный  $((\mathbf{p}_i)(x) = \langle 1, v_{p_i} \rangle)$ , а наихудшему – минимальный  $(\mathbf{p}_i(x) = \langle 0, 0 \rangle)$  результат отображения  $\mathbf{p}_i(x)$ . Здесь  $v_{p_i}$  определяется как радиус интервала, задающего максимальную полезность решения  $x \in \mathbf{X}$ , соответствующий  $\delta_i$ .

Указанным требованиям отвечает интервальное отображение  $\mathbf{p}_i(x)$  вида:

$$p_i(x) = \left( \frac{k_i(x) - \bar{k}_i^-}{k_i^+ - \bar{k}_i^-} \right)^{\alpha_i}, \quad (4)$$

где  $k_i(x)$  – значение частного критерия,  $k_i^+$ ,  $\bar{k}_i^-$  – соответственно наилучшее и наихудшее значение частного критерия  $k_i(x)$  на области допустимых решений  $x \in X$ , при этом

$$k_i^+ = \begin{cases} \max_{x \in X} k_i(x), & \text{если } k_i(x) \rightarrow \max, \\ \min_{x \in X} k_i(x), & \text{если } k_i(x) \rightarrow \min; \end{cases}$$

$$k_i^- = \begin{cases} \max_{x \in X} k_i(x), & \text{если } k_i(x) \rightarrow \max, \\ \min_{x \in X} k_i(x), & \text{если } k_i(x) \rightarrow \min; \end{cases}$$

$\alpha_i \in R^1$  определяет характер нелинейности функции полезности  $p_i(x)$  в детерминированном случае [6],  $\bar{k}_i^-$  – интервал, сопряженный интервалу  $k_i^-$  [4].

Рассмотрим интервальное отображение  $\hat{p}_i(x)$  потери полезности частного критерия  $k_i \in K$  как  $\hat{p}_i(x) = \langle 1, v_{p_i} \rangle - \bar{p}_i(x)$ ,  $\bar{p}_i(x)$  – отображение, сопряженное отображению  $p_i(x)$  [4]. Очевидно, что независимо от вида экстремума частного критерия  $k_i(x)$ , наилучшему результату отображения  $\hat{p}_i(x)$  соответствует минимальное значение  $\langle 0, 0 \rangle$ , а наихудшему –  $\langle 1, v_{p_i} \rangle$ . В дальнейшем, ориентируясь на решение задач минимизации интервальных отображений, будем осуществлять выбор наилучшего решения из множества  $X$  с помощью интервальных оценок потери полезности  $\hat{p}_i(x)$ .

Качество каждого соединения  $x \in X$  может быть оценено путем формирования интервальных многофакторных оценок альтернатив на основе отображений обобщенной полезности альтернатив. При этом может быть учтено наличие у ЛПР данных о важности критериев. Рассмотрим следующие варианты [5]:

а) известны точные значения  $a_i$  коэффициентов относительной важности интервальных критериев  $k_1, k_2, \dots, k_n$ , а, следовательно, их отображений локальной потери полезности  $\hat{p}_i(x)$ . Здесь

$$a_i \in R^1, \quad 0 \leq a_i \leq 1, \quad i \in J_n, \quad \sum_{i=1}^n a_i = 1. \quad (5)$$

Тогда аддитивная обобщенная интервальная оценка потери полезности примет вид:

$$\hat{P}(x) = \sum_{i=1}^n a_i \cdot \hat{p}_i(x), \quad (6)$$

а решение задачи определяется как:

$$x^0 = \arg \min_{x \in X} \hat{P}(x). \quad (7)$$

Полученное решение  $x^0$  соответствует минимальной в смысле соотношения (2) интервальной обобщенной оценке потери полезности альтернатив  $x \in X$ ;

б) значения коэффициентов относительной важности интервальных критериев  $k_1, k_2, \dots, k_n$  известны с точностью до интервалов вида

$$\alpha_i = [\alpha_{i_{\min}}, \alpha_{i_{\max}}], \quad i \in J_n. \quad (8)$$

Каждому интервалу вида (8) поставим в соответствие центрированный интервал  $\langle A_i \rangle = \langle a_i, v_{a_i} \rangle \in I_s R$ ,  $i \in J_n$ , [4] следующим образом:

$$\langle A_i \rangle \Leftrightarrow [a_i - v_{a_i}, a_i + v_{a_i}] = [\alpha_{i_{\min}}, \alpha_{i_{\max}}],$$

где

$$a_i = \frac{1}{2}(\alpha_{i_{\min}} + \alpha_{i_{\max}}), \quad v_{a_i} = \frac{1}{2}(\alpha_{i_{\max}} - \alpha_{i_{\min}}).$$

С учетом нормализованных интервальных коэффициентов относительной важности интервальных критериев соотношения, аналогичные (6), (7), примут вид

$$\hat{P}(x, A^H) = \sum_{i=1}^n \langle A_i^H \rangle * \hat{p}_i(x),$$

$$x^0 = \arg \min_{x \in X} \hat{P}(x, A^H),$$

где  $A^H = (\langle A_1^H \rangle, \langle A_2^H \rangle, \dots, \langle A_n^H \rangle) \in I_s^n R$ ,  $\langle A_i^H \rangle = \langle a_i^H, v_{a_i^H} \rangle$ ,

$$a_i^H = \frac{1}{\Delta_i} \left( a_i \sum_{i=1}^n a_i + v_{a_i} \sum_{i=1}^n v_{a_i} \right), \quad v_{a_i^H} = \frac{1}{\Delta_i} \left( v_{a_i} \sum_{i=1}^n a_i + a_i \sum_{i=1}^n v_{a_i} \right),$$

$$\Delta_i = \frac{1}{\left( \sum_{i=1}^n a_i \right)^2 - \left( \sum_{i=1}^n v_{a_i} \right)^2}, \quad \langle 0, 0 \rangle \leq \langle A_i^H \rangle \leq \langle a_i^H, v_{a_i^H} \rangle,$$

$i \in J_n$ ;

в) количественные значения весовых коэффициентов относительной важности критериев неизвестны, но интервальные критерии упорядочены по важности, например, следующим образом:

$$k_1 \succ k_2 \succ \dots \succ k_n.$$

Такое задание предпочтений частных критериев означает, что  $\langle A_1 \rangle \succ \langle A_2 \rangle \succ \dots \succ \langle A_n \rangle$ .

В этой ситуации используется принцип последовательной оптимизации, в соответствии с которым из двух решений  $u \in X$ ,  $v \in X$  первое предпочтительно, т. е.  $u \succ v$ , если

$$p_1(u) < p_1(v)$$

или  $(\mathbf{p}_1(u) = \mathbf{p}_1(v)) \wedge (\mathbf{p}_2(u) < \mathbf{p}_2(v))$

или  $(\mathbf{p}_1(u) = \mathbf{p}_1(v)) \wedge (\mathbf{p}_2(u) = \mathbf{p}_2(v)) \wedge (\mathbf{p}_3(u) < \mathbf{p}_3(v))$

и т. д.

$\exists t \in J_{n-1}$ , такое что  $(\ddot{\mathbf{p}}_j(u) = \ddot{\mathbf{p}}_j(v), j \in J_t) \wedge \ddot{\mathbf{p}}_{t+1}(u) < \wedge (\ddot{\mathbf{p}}_{t+1}(v))$ .

Выбор решения сводится к решению последовательности однокритериальных задач:

$$x_i^0 = \text{Arg} \min_{x \in X_{i-1}^0} \hat{\mathbf{p}}_i(x),$$

здесь  $i \in J_n$ ,  $X_0^0 \equiv X$ ;

г) информация о предпочтениях относительно критериев  $\mathbf{k}_1, \mathbf{k}_2, \dots, \mathbf{k}_n$ , а, следовательно, и о коэффициентах  $a_i, i \in J_n$ , отсутствует.

В этом случае следует использовать модель вида

$$x^0 = \arg \min_{x \in X} \max_{i=1,2,\dots,n} \hat{\mathbf{p}}_i(x).$$

**Способ 2.** В отличие от способа 1, который заключается в многокритериальной оценке соединений заданных вершин, данный способ предполагает предварительную оценку ребер по множеству соответствующих критериев. На основе значений множества критериев для каждого ребра формируется оценка многофакторной эффективности, которая далее рассматривается как вес ребра. Формирование интервальной многофакторной оценки производится по тем же правилам, что и в способе 1. Затем с использованием всех выбранных направлений в сигнальный граф добавляются все допустимые ребра-кроссировки. Их веса полагаются равными нулю, поскольку кроссировки не оказывают какого-либо заметного влияния на качество всей линии из-за их небольшой длины и хороших электрических характеристик. На полученном графе для заданной пары вершин с помощью известных методов [7] ищется критический путь, после чего он проверяется на соответствие ограничениям. В случае, когда найденный критический путь является недопустимым, можно попытаться смягчить некоторые из технологических ограничений. Описанный способ, хотя и не всегда приводит к решению, требует значительно меньших вычислительных затрат, чем способ 1.

Таким образом, способ 2 рекомендуется для быстрого приближенного решения задачи, тогда как способ 1 используется для отыскания лучшего из всех допустимых решений.

Для поиска пути с характеристиками лучше заданных производится последовательная генерация путей

в рамках направлений с одновременной оценкой и сравнением с лучшим из найденных ранее. Если рассматриваемый путь лучше, он запоминается и перебор продолжается.

## ВЫВОДЫ

Таким образом, в работе проведен анализ требований, возникающих при решении задачи поиска технической возможности организации линии между двумя точками сети и выбора наилучшего пути для ее подключения в условиях неопределенности. Разработана интервальная математическая модель на основе известных интервальных многофакторных оценок альтернатив.

В отличие от существующих частных моделей, предложенная интервальная модель позволяет применить единую методологическую и алгоритмическую базу при решении задач оптимизации соединений проводных сетей связи в условиях неопределенности.

## ПЕРЕЧЕНЬ ССЫЛОК

1. Гребенник И. В. Оптимизация пути соединения двух точек проводной сети электросвязи / Гребенник И. В., Хабаров А. Ю. // АСУ и приборы автоматики. – 2002. – № 120. – С. 38–44.
2. Драмашко И. А. Актуальные проблемы автоматизации технического учета / Драмашко И. А., Крижевский В. С., Кокорин А. Б. // Зв'язок. – 1998. – № 3. – С. 43–46.
3. Воробієнко П. П. Формування технічних вимог до міських мереж / Воробієнко П. П., Лейбзон А. Я., Нечипорук О. А., Чумак М. О. // Зв'язок. – 2002. – № 2. – С. 16–21.
4. Стоян Ю. Г. Метрическое пространство централизованных интервалов / Стоян Ю. Г. // Докл. НАН Украины. Сер. А. – 1996. – № 7. – С. 23–25.
5. Гребенник И. В. Интервальное оценивание альтернатив в задачах принятия решений / Гребенник И. В., Романова Т. Е., Шеховцов С. Б. // Кибернетика и системный анализ. – 2009. – № 2. – С. 106–115.
6. Овезгельдыев А. О. Синтез и идентификация моделей многофакторного оценивания и оптимизации / Овезгельдыев А. О., Петров Э. Г., Петров К. Э. – Киев : Наук. думка, 2002. – 164 с.
7. Алгоритмы и программы решения задач на графах и сетях / Под ред. М. И. Нечепуренко. – Новосибирск : Наука, Сиб. отд., 1990. – 520 с.

Надійшла 27.03.2009

*Аналізується задача вибору найкращого з'єднання двох точок провідної мережі електрозв'язку за множиною критеріїв в умовах невизначеності. Будується інтервальна багатокритеріальна математична модель задачі, наводиться метод її розв'язання.*

*The problem of selection of the best connection between two points of the communication network by the set of criteria under uncertainty is analyzed. The multi-criterion interval model of the problem is constructed; the method of its solving is given.*

В. Н. Журавлєв, В. С. Кабак

## ИНФОРМАЦИОННАЯ МОДЕЛЬ ИМПУЛЬСНОГО РЕЧЕСЛУХОВОГО ПРОЦЕССА

*Рассматривается задача, суть которой состоит в раскрытии и анализе импульсного характера психофизиологического информационного управления речеслуховым процессом. Разработан метод анализа связи пространства управляющих параметров центральной нервной системы речевыми исполнительными органами и механизм снятия энтропии с ламинарного потока воздуха в процессе синтеза речевого сигнала.*

### ВВЕДЕНИЕ. ПОСТАНОВКА ЗАДАЧИ

Принято считать [1], что процессы синтеза и анализа речевого сигнала (РС) осуществляются центральной нервной системой (ЦНС) человека в виде импульсной последовательности двоичных сигналов нейронов. В настоящее время доказано [2], что функционирование эволюционных природных биологических объектов определяется нелинейными стохастическими пространственно-временными диссипативными процессами. Синергетическая методология синтеза и анализа РС изложена в работе [3], однако в ней не раскрыт механизм информационного управления энтропией источника и приемника речевых сообщений. Решение задачи, суть которой состоит в раскрытии и анализе импульсного характера психофизиологического информационного управления речеслуховым процессом, является целью исследований настоящей статьи. Объектом исследований являются процессы энергоинформационного обмена биологических подсистем в помещении объекта информационной деятельности (ОИД). Предмет исследований – информационные параметры и функции импульсного процесса речеслухового обмена энергией, информацией и энтропией сообщения.

### СИНЕРГЕТИЧЕСКИЙ АНАЛИЗ ИНФОРМАЦИИ РЕЧЕСЛУХОВОГО ПРОЦЕССА

Информационное и энергетическое взаимодействие исследуемых речеслуховых систем диктора и аудитора целесообразно анализировать, беря в основу методологию синергетики [2], и открытых систем [4, 5], которая позволяет сформулировать гипотезу речеслухового импульсного процесса энергоинформационного обмена открытых пространственно-времен-

ных нелинейных диссипативных психофизиологических информационных подсистем объекта исследований.

1. Психофизиологические речеслуховые подсистемы диктора и аудитора во время речеслухового процесса (РП) характеризуются импульсным обменом информацией  $I(t, \Delta t)$  и энергией  $E(t, \Delta t)$  между подсистемами в среде функционирования. Функция изменения энергии РС  $E(t, \Delta t)$  во времени  $t$  чувствительна к параметрам расстояния  $l_{da}$  между диктором и аудитором и эффективной энергии помех  $E_{eff}^{\xi}$  в канале связи. Переносчиком информации  $I_i(t, \Delta t)$   $i$ -й фонемы, на интервале времени  $Tr$  ее синтеза, являются импульсы приращений энергии  $\Delta E_i(t, \Delta t)$ . Связь между ними определяется функцией распределения плотности вероятностей (ФРПВ):

$$I_i(t, \Delta t) = w_i [\Delta E_i(t, \Delta t), \Delta t(t)] \\ t \in [\Delta t] \in [Tr], \quad (1)$$

где  $\Delta t$  – информационный интервал времени активности импульса энергии;

2. Внутри выделенного помещения ОИД действует закон сохранения энергии  $E(t, \Delta t)$ , информации  $I(t, \Delta t)$  и энтропии  $H(t, \Delta t)$

$$\sum E_i(t, \Delta t) = \text{const}, \\ \sum I_i(t, \Delta t) = H_0(t) - \sum H_i(t, \Delta t), \quad (2)$$

где  $H_0(t)$  – начальная энтропия приемника речевых сообщений;

3. Результатом обмена энергией и информацией является процесс самоорганизации и уменьшения энтропии  $H_0(t, \Delta t)$  приемника сообщений – ЦНС аудитора;

4. Психофизиологические процессы речеслуховых подсистем источника и приемника сообщений определяются теорией биологических функциональных систем П. К. Анохина [6].

Методология функционирования открытых биологических функциональных систем в границах объекта исследований предполагают следующую концепцию:

1. Основой процессов самоорганизации, характеризующих энергоинформационный обмен между диссипативными открытыми биологическими информационными системами в неравновесном состоянии, являются *системообразующие факторы и функции*;

2. *Системообразующим фактором* психофизиологической речеслуховой системы является факт реализации полезного результата ее работы – речеслухового процесса;

3. *Системообразующая функция* – система функциональных или операторных зависимостей энергоинформационных параметров и характеристик РП, обеспечивающих реализацию системообразующего фактора. Параметры системообразующей функции иницируются и формируются в ЦНС посредством синтеза векторов управляющих параметров  $\vec{\mathfrak{R}}(t, \Delta t)$ . Данный вектор иницирует процесс управления физиологическими речеслуховыми исполнительными органами, которые синтезируют сигналы кинем  $k(t, \Delta t)$ . Последние управляют системным фазовым параметром  $q(t, \Delta t)$ , влияющим на РС фонем  $s(t, \Delta t)$ :

$$\begin{aligned} s(t, \Delta t) &= f_1[q(t, \Delta t)], \\ q(t, \Delta t) &= f_2[k(t, \Delta t)], \\ k(t, \Delta t) &= f_3[\vec{\mathfrak{R}}(t, \Delta t)], \\ \vec{\mathfrak{R}}(t, \Delta t) &= f_4[I(t, \Delta t)]; \end{aligned} \quad (3)$$

4. Изменение векторов управляющих параметров  $\vec{\mathfrak{R}}(t, \Delta t)$  реализации РП (3) осуществляется под непрерывным контролем ЦНС параметра эффективности  $q(t, \Delta t)$  и чувствительно к объему апостериорной информации  $I_{i-1}^{\text{ПНС}}(t, \Delta t)$ , содержащейся в накопленном опыте диктора при предыдущих реализациях РП.

$$\begin{aligned} q_i(t, \Delta t) &= Q[I_i(t, \Delta t), I_{i-1}^{\text{ПНС}}(t, \Delta t)], \\ \vec{\mathfrak{R}}(t, \Delta t) &= f_5[q_i(t, \Delta t)], \end{aligned} \quad (4)$$

где  $Q$  – системный функциональный оператор анализа информационного расстояния. Результаты контроля обладают императивными возможностями по реорганизации структуры векторов управляющих параметров физиологических исполнительных органов. Речеслуховой процесс (3) реализуется в синергетическом режиме взаимодействия, т. е. динамической взаимной адаптации речевой и слуховой подсистем по параметру влияния сигналов внешних помех  $\xi(t)$  на параметр эффективности (4) его реализации.

Обозначим через  $\vec{\mathfrak{Z}}[\mathfrak{X}_i(t, \Delta t)]$ ,  $i \in [0, n]$  пространство векторов управляющих параметров обобщенного

сигнала кинем  $i$ -й фонемы. В качестве сигнала  $k_0(t, \Delta t)$  нулевой кинемы и вектора ее управляющих параметров  $\vec{\mathfrak{R}}_0(t, \Delta t)$  примем состояние речевой подсистемы на интервале времени информационной паузы диктора (например, фаза вдоха воздуха в легкие), которое соответствует состоянию максимальной неопределенности относительно фонемы, которая будет синтезироваться в ближайшем будущем. Существует два статистических определения понятия «информация»  $I[s(t, \Delta t)]$  для импульсных сигналов  $s(t, \Delta t)$ . Первое – энтропия Больцмана:

$$\begin{aligned} I[s(t, \Delta t)] &= H[s(t, \Delta t)] = \\ &= -\sum_n w[s(t, \Delta t)] \ln w[s(t, \Delta t)], \quad n = \frac{Tp}{M\{\Delta t\}}, \end{aligned} \quad (5)$$

где  $M\{\Delta t\}$  – первый момент информационных интервалов времени  $\Delta t$ . Однако это определение не отражает динамики изменения энтропии  $H(t, \Delta t)$  РС в процессе его синтеза и, одновременно, процесса информационной самоорганизации психофизиологической системы речевого обмена.

В классической работе К. Шеннона по теории связи [7] даны два определения информации  $I(t)$ . Первое определение фактически совпадает с определением энтропии Больцмана (5). Эта информация, как и энтропия  $H(t)$ , является мерой степени неопределенности для выбранного энергетически стационарного состояния статистического описания речеслуховой системы. Такое определение информации является недостаточным для исследования биологических открытых систем речевого обмена т. к. не учитывает максимальный уровень неопределенности  $H_0(t, \Delta t)$  РС при воздействии как внешних сигналов помех ОИД  $\xi^{\text{ОИД}}(t)$ , так и внутренних физиологических шумов  $\xi^{\text{ПНС}}(t)$  речеслуховой системы.

По нашему мнению, более адекватным для открытых систем является другое, также предложенное К. Шенноном, определение информации. Суть его, в приложении к РП, состоит в следующем. Пусть имеются две функции плотности вероятностей  $w_i[s(t)]$  и  $w_{i+1}[s(t + \Delta t)]$  определенные для фонемы РС  $s(t, \Delta t)$ . Они отстоят друг от друга на минимальный интервал времени  $[\Delta t]$ , определяемый быстродействием нейронов рецептора слуха ( $\Delta t \ll Tp$ ) $_{E_{\text{eff}}(t) = \text{const}}$  и находятся в пределах интервала времени  $Tp$  при условии постоянной эффективной энергии  $E_{\text{eff}}(t)$  (2). В обоих случаях информация (5) определяется разностью безусловной и условной энтропий и связана с соответствующим снятием части энтропии диктора относительно  $i$ -й фонемы

$$\begin{aligned} I[s(t + \Delta t)] &= H[s(t)] - H[s(t + \Delta t)/s(t)], \\ &(\Delta t < Tp)_{E_{\text{eff}}(t) = \text{const}}, \end{aligned} \quad (6)$$

где условная энтропия  $H[s(t + \Delta t)/s(t)]$  определяется через соответствующую условную ФПРВ:

$$H[s(t + \Delta t)/s(t)] = - \iint w[s(t + \Delta t), s(t)] \ln w[s(t + \Delta t)/s(t)] ds(t + \Delta t) ds(t), \quad (7)$$

где  $w[s(t + \Delta t), s(t)]$  – безусловная, а  $w[s(t + \Delta t)/s(t)]$  – условная ФПРВ. В приложении к системам речевого обмена, дифференциальное по времени  $\Delta t$ , выражение (7) не учитывает зависимости энтропии диктора от императивного воздействия вектора управляющего параметра кинем  $\vec{\mathfrak{R}}(\Delta t)$  на интервале времени  $\Delta t$ . Для положительности информации (6) требуется обеспечение условия постоянства эффективной энергии

$$E_{eff}(t) = \int_t s(t) dt = \text{const}, \quad t \in [Tp]. \quad (8)$$

С учетом того, что акцептором информации  $I[s(t, \Delta t)]$  является рецептор слуха улитки внутреннего уха, постоянство эффективной энергии (8) обеспечивается системообразующей функцией среднего уха.

Учитывая факт природной идентификации фоном на интервале времени  $Tp$ , предположим, что, функция плотности вероятности  $w[s(t, \Delta t)]$ ,  $t \in [\Delta t \ll Tp]$  полностью характеризуется соответствующим набором первых моментов сигналов кинем  $k(t, \Delta t)$ , которые, в свою очередь, характеризуются соответствующим набором первых моментов векторов управляющих информационных параметров  $M\{\vec{\mathfrak{R}}(t, \Delta t)\}$

$$w[s(t, \Delta t)] = f[\vec{\mathfrak{R}}(t + \Delta t)]. \quad (9)$$

Подставим (9) в формулу (7) и выполним интегрирование по  $s(t)$ . В результате получим выражение для информации о совокупности  $s(t)$  при заданном значении вектора управляющего параметра сигнала кинем  $\vec{\mathfrak{R}}(\Delta t)$ :

$$\begin{aligned} I[s(t + \Delta t)/\vec{\mathfrak{R}}(\Delta t)] &= H[s(t)] - H[s(t + \Delta t)/\vec{\mathfrak{R}}(\Delta t)] = \\ &= H[s(t)] + \int_{s(t + \Delta t)} w[s(t + \Delta t)/\vec{\mathfrak{R}}(\Delta t)] \times \\ &\times \ln w[s(t + \Delta t)/\vec{\mathfrak{R}}(\Delta t)] ds(t + \Delta t). \end{aligned} \quad (10)$$

Необходимо учесть, что определение информации сигнала фонемы по формуле (10) не может быть использовано во всех случаях, т. к. может принимать отрицательные значения. Для выполнения условия  $I[s(t + \Delta t)/\vec{\mathfrak{R}}(\Delta t)] \geq 0$  необходимо ввести дополнительное условие, которое можно сформулировать на основании функционала Ляпунова  $\Lambda_s(t + \Delta t)$ . С уче-

том условия (8) выражение (10) можно переформулировать на языке функционала Ляпунова, который определяет разность энтропий для равновесного и неравновесного состояний потока воздуха в области действия управляющего параметра  $\vec{\mathfrak{R}}(\Delta t)$

$$\begin{aligned} \Lambda_s(t + \Delta t) &= H[s(t)] - H[s(t + \Delta t)/\vec{\mathfrak{R}}(\Delta t)] = \\ &= k \iint \left\{ \ln \frac{w[s(t)]}{w[s(t + \Delta t)]} \right\} w[s(t)] ds(t) ds(t + \Delta t) \geq 0, \end{aligned} \quad (11)$$

$$\frac{\Delta \Lambda_s(t + \Delta t)}{\Delta t} - \frac{\Delta \{H[s(t)] - H[s(t + \Delta t)/\vec{\mathfrak{R}}(\Delta t)]\}}{\Delta t} \leq 0. \quad (12)$$

При установлении неравенства (11) используется условие  $E_{eff}(t) = \text{const}$ , при установлении (12)

$$\frac{\Delta \{H[s(t)] - H[s(t + \Delta t)/\vec{\mathfrak{R}}(\Delta t)]\}}{\Delta t} > 0.$$

Вернемся к выражению (10). Пусть вектор управляющих параметров сигналов кинем  $\vec{\mathfrak{R}}(\Delta t)$  может принимать лишь положительные значения, что обеспечивается физиологической архитектоникой речевой подсистемы, а безусловная энтропия  $H[s(t, \Delta t)]$  отвечает его нулевому значению  $\vec{\mathfrak{R}}(t, \Delta t) = 0$  при синтезе нулевой фонемы  $H[s(t, \Delta t)] = H[s_0(t, \Delta t)]$ :

$$\begin{aligned} \vec{\mathfrak{R}}(\Delta t) = 0, \quad H[s_0(t, \Delta t)] &= \\ = H[s_0(t, \Delta t)/\vec{\mathfrak{R}}(t, \Delta t) = 0]. \end{aligned} \quad (13)$$

При этих условиях информация равновесного состояния равна нулю:

$$I[s_0(t)/\vec{\mathfrak{R}}(t, \Delta t)] \Big|_{\vec{\mathfrak{R}}(t, \Delta t) = 0} = 0, \quad (14)$$

а безусловная энтропия  $H[s(t, \Delta t)]$  совпадает с энтропией равновесного состояния.

При синтезе фонемы в процессе изменения вектора управляющего параметра  $\vec{\mathfrak{R}}(t, \Delta t)$  речевая система переходит из одного равновесного состояния к другому в течение информационных интервалов времени  $\Delta t \in [Tp]$  движения потока воздуха в физиологической области действия управляющего параметра  $\vec{\mathfrak{R}}(t, \Delta t)$ , при этом сумма информации и энтропии остается постоянной и равной энтропии фонемы  $H_p$ :

$$I[s(t, \Delta t)] + H[s(t, \Delta t)] = \text{const} = H_p, \quad t \in [Tp]. \quad (15)$$

При этом константа  $k$  выражения (11) определяется энтропией равновесного состояния. Информация равновесного состояния равна нулю в соответствии с выражением (12). Для формулировки критерия поло-



жительности выражения (10) необходимо перенормировать энтропию более хаотического состояния так, чтобы сопоставление информационных состояний речевой системы производилось при одинаковых значениях средней эффективной энергии  $E_{eff}(t)$ .

Проведем анализ двух реализаций РС, отвечающих следующим состояниям управляющего параметра:  $\vec{\mathfrak{R}}(t, \Delta t) = 0$  – синтез нулевой фонемы и  $\vec{\mathfrak{R}}(t, \Delta t) = \vec{\mathfrak{R}}_i(t, \Delta t)$  – некоторое стационарное, но неравновесное состояние. Будем применять соответствующие функции плотности распределения вероятностей  $w_0, w_i$  и энтропии  $H_0, H_i$ . Ренормализация к заданному значению средней энергии  $\vec{E}(t)$  сводится к замене энергии шумов неравновесного состояния ее эффективным значением  $E_{eff}(t)$ , что выполняется средним ухом на интервале времени адаптации  $\Delta t^{su}$ . Она определяется путем решения уравнения:

$$k1\vec{E}(t, \Delta t) = \int \vec{E}(t)w_0(E(t, \Delta t), \vec{\mathfrak{R}}(t, \Delta t) = 0)dE = \int E_{eff(t)}w_i(E(t, \Delta t), \vec{\mathfrak{R}} = \vec{\mathfrak{R}}_i(t, \Delta t))dE, \quad (16)$$

где  $k1$  – субъективный психофизиологический коэффициент, учитывающий чувствительность слуха аудитора. Выражение (16) служит дополнительным условием, обеспечивающим положительность информации уравнения (10). Решение этого уравнения удовлетворяет неравенству

$$E[\vec{\mathfrak{R}}(t, \Delta t)] \geq \vec{E}(t). \quad (17)$$

Знак равенства относится к случаю  $\vec{\mathfrak{R}}(t, \Delta t) = \vec{\mathfrak{R}}_0 = 0$ . Отсюда следует, что для выравнивания значений средних энергий состояния  $\vec{\mathfrak{R}}(t, \Delta t) = 0$  ЦНС «подогревает» слуховой рецептор энергией внутренних шумов  $\xi^{IHC}(t)$ :

$$E_{вн}[\xi^{IHC}(t)] = \text{const} \neq 0, \quad (18)$$

и выражение (11) для разности энтропий двух соседних состояний можно записать в виде:

$$H_i(t, \Delta t) = -k \int \left\{ \ln \frac{w_1[E_1(t, \Delta t)]}{w_2[E_2(t, \Delta t)]} \right\} w_1[E_1(t, \Delta t)]dE - H_0 \geq 0, \quad (19)$$

где  $H_0$  соответствует психофизиологической постоянной концентрации информационного внимания ЦНС на РС. Выражение (19) отображает факт улучшения разборчивости РС при его суммировании с шумовым сигналом [8], также соответствует экспериментальной формуле психофизиологического закона Вебера – Фех-

нера [1], определяющей реакцию биологической функциональной системы (энтропия  $H_i$ ) на раздражитель (энергия  $E$ ) пропорциональной его логарифму.

Относительная упорядоченность информационных состояний РС на интервалах времени адаптации слуха  $\Delta t^{su}$  определяется двумя формулами. Формула (16) подтверждает выбор равновесного состояния при значении управляющего параметра кинем  $\vec{\mathfrak{R}} = 0$  и суммы энергии внутренних и внешних шумов, т. е. системообразующей функции адаптации ЦНС к внешним помехам. Формула (19) дает количественную меру их относительной упорядоченности.

Исследуя общую формулу (10), мы можем определить информацию  $\vec{I}[E(t, \Delta t)]$  стационарного состояния речевой системы при всех значениях вектора управляющего параметра:

$$\vec{I}[E(t, \Delta t)] = H_0 - H_i(t, \Delta t) = \int \left[ \ln \frac{w_1[E_1(t, \Delta t)]}{w_0[E_0(t, \Delta t)]} \right] w_1[E_1(t, \Delta t)]dE \geq 0. \quad (20)$$

Отсюда следует, что при нулевом значении вектора управляющего параметра кинем  $\vec{\mathfrak{R}} = 0$  сообщение совпадает с равновесным и информация равна нулю. На этом же основании можно определить информацию временной реализации последовательности стационарных состояний  $\Delta t$  течения воздуха в физиологических зонах действия векторов управляющих параметров.

## ВЫВОДЫ

1. Впервые проведен информационно-энергетический анализ изменения энтропии речевого сигнала при воздействии вектора управляющих параметров ЦНС на сигналы кинем.

2. Получены выражения (10)–(12), (19), (20), анализ которых показывает, что информационная функция фонемы РС чувствительна к разнице между энтропией энергии  $i$ -го и условной энтропии  $(i+1)$ -го сигнала фонемы по воздействию вектора управляющих параметров  $\vec{\mathfrak{R}}(t, \Delta t)$  на приращение энергии  $\Delta E(t, \Delta t)$  сигнала в течение интервала времени  $\Delta t$ .

3. Получено аналитическое выражение (16), отображающее системообразующую функцию среднего уха, которая определяет процесс адаптации коэффициента передачи энергии РС по ФРПВ энергии нулевой фонемы.

## ПЕРЕЧЕНЬ ССЫЛОК

1. Чистович Л. А. Физиология речи. Восприятие речи человеком / [Чистович Л. А., Венцов А. В., Гранстрем М. П. и др.]. – Л. : «Наука», 1976. – 388 с. – (В серии «Руководство по физиологии»).

2. Николис Г. Самоорганизация в неравновесных системах / Николис Г., Пригожин И. – М. : Мир, 1979. – 512 с.
3. Журавлёв В. Н. Синергетическая концепция энергоинформационного обмена речеслуховой системы / В. Н. Журавлёв // Адаптивные системы автоматического управления. – 2007. – № 11. – С. 128–135.
4. Климонтович Ю. Л. Статистическая теория открытых систем / Климонтович Ю. Л. – М. : Янус, 1995. – 624 с.
5. Климонтович Ю. Л. Турбулентное движение и структура хаоса / Климонтович Ю. Л. – М. : Ком. книга, 2007. – 328 с.
6. Анохин П. К. Кибернетика функциональных систем / Анохин П. К. – М. : Медицина, 1998. – 400 с.
7. Шеннон К. Современные достижения теории связи. / К. Шеннон // Работы по теории информации и кибернетике. – М. : Издательство иностранной литературы, 1963. – С. 403–414.
8. Цвикер Э. Ухо как приемник информации. / Цвикер Э., Фельдкеллер Р. ; пер. с нем. под ред. Б. Г. Белкина. – М. : Связь, 1971. – 225 с.

Надійшла 16.02.2009

Після доробки 27.04.2009

*Розглянута задача, суттю якої є розкриття та аналіз імпульсного процесу психофізіологічного інформаційного керування мовним та слуховим процесами. Розроблено метод аналізу зв'язку керуючих параметрів центральної нервової системи з фізіологічними органами мови та механізмом зняття ентропії з ламінарного потоку повітря у процесі синтезу сигналів мови.*

*The task, the main point of which is the expansion and analysis of impulse character of speech-hearing process psycho-physiological informational control is under review. The method of analysis of connection between the control parameters of central nervous system under speech executive organs and the mechanism of taking down of entropy from the laminar air flow in the process of speech signal synthesis was developed.*

УДК 004.032.26

Т. В. Кіпріч, В. М. Харитонов, В. І. Дубровін, А. В. Притула

## ПОБУДОВА МЕТОДІВ ТА МОДЕЛЕЙ ДЛЯ ПОПЕРЕДЖЕННЯ ПОМПАЖНИХ ЯВИЩ У ГАЗОТУРБІННОМУ ДВИГУНІ

*Метою існуючих систем захисту газотурбінного двигуна від зривних процесів є автоматичне відновлення режиму роботи двигуна після усунення помпажу, а не запобігання виникненню газодинамічної нестійкості у турбомашинах. Запропонована модель модуля попередження помпажу ґрунтується на основі вейвлет-аналізу та карт Кохонена SOM. Приводиться порівняльний аналіз результатів виявлення зривних процесів за допомогою даного підходу та спрощених моделей.*

### ВСТУП

Існує два види зривних режимів, що призводять до аварійного стану газотурбінного двигуна (ГТД) та характеризуються значними коливаннями: обертовий зрив та явища типу помпажних. Обертовий зрив у більшості випадків пов'язаний із зривом потоку з лопаток компенсатора та розповсюдженням збурення по колу проточної частини компресора. Внаслідок цього явища може бути розвиненим помпаж ГТД, який супроводжується коливанням тиску, швидкостей та витрат газу по тракту двигуна; зменшенням частот обертання роторів; зростанням температури газів перед і за турбіною чи загасанням камери згорання [1].

Серед наведених параметрів найбільш характерним для помпажних явищ є різке падіння тиску повітря за компресором ( $P_{\text{квт}}$ ) чи по тракту двигуна, що повторюється при кожному імпульсі помпажних коливань. Тому робота більшості сигналізаторів побудова-

на саме на обробці сигналу  $P_{\text{квт}}$ , за допомогою якого при помпажі здійснюється виробка відповідного сигналу до протипомпажного та протизривного блоку (ППБ) (або системи захисту від помпажу за обертового зриву) для впливу на виконавчий орган системи автоматичного керування двигуном (САКД). Однак, метою сучасних САКД є відновлення режиму роботи двигуна після помпажу, а не попередження розвитку газодинамічної нестійкості у турбомашинах. Тому необхідна відповідна розробка нових та модифікація існуючих методів для попередження виникнення зривних процесів у турбокомпресорі та реалізація модуля попередження помпажу (МПП) у САКД.

### 1 ОГЛЯД ПОПЕРЕДЕНІХ ДОСЛІДЖЕНЬ

У сучасних САКД реалізовано наступні сигналізатори коливань із логарифмічними датчиками [1]:

- 1) штатний датчик логарифмічний типу ДОЛ із електронним перетворювачем ЕП-В;
- 2) сигналізатори помпажу ПС та ПС-2-7.

Датчик тиску ДОЛ призначений для виміру надлишкового тиску повітря із видачею сигналу змінного току пропорційного логарифму вимірювального тиску на електронний перетворювач ЕП-В [2, 3]. При втраті стійкості компресора ЕП-В замикає ланцюг керування ППБ САКД. Принцип роботи перетворювача побудований на виділенні сигналу із зада-

© Кіпріч Т. В., Харитонов В. М., Дубровін В. І., 2009

ною відносною швидкістю зміни його значення. При перевищенні деякої граничної величини (що настроюється) спрацьовує амплітудний сектор і на виході перетворювача з'являється командний сигнал:

1) ЕП-В видає до ППБ сигнал помпажу, коли відносне падіння тиску повітря за компресором  $\frac{\Delta P_2}{P_2} \geq 0,55$  та відносна швидкість падіння  $\frac{\Delta P_2}{\Delta t P_2} \geq 7$  (де  $P_2$  – зміна тиску повітря за компресором за час  $t$ ).

Частота коливань повітря при цьому сягає 5–15 Гц, час запізнення сигналу не більш ніж 0,05 с, а тривалість не більш ніж 0,02 с;

2) ЕП-В видає до ППБ сигнал обертового зриву, коли виконуються умова:  $\frac{\Delta P_2}{P_2} \geq 0,35$ .

На відміну від ДОЛ із ЕП-В для сигналізаторів помпажу ПС/ПС-2-7, помпаж визначається як різке падіння тиску повітря за компресором на величину не менш ніж  $0,4P_2$  для ПС і  $0,6P_2$  для ПС-2-7 [4].

Не дивлячись на переваги використання наведених сигналізаторів помпажу із логарифмічною характеристикою, висока швидкість розвитку помпажних явищ приводить до того, що сигнал про початок зривних процесів не може бути реалізованим у САКД настільки швидко, щоб попередити помпаж.

Одним із підходів до ранньої діагностики розвитку зривних процесів стає використання методу вейвлет-аналізу (ВА) для обробки часових та амплітудних характеристик сигналів вимірювальних систем, що визначають стійкість роботи двигуна. Отримані результати [5] свідчать про те, що на певних масштабах статистичні моменти ( $m_2$ – $m_4$ ) розподілення вейвлет-коефіцієнтів (ВК) сигналів від датчиків тиску повітря, які розташовані в компресорі, можуть помітно змінюватися при виникненні та розвитку помпажної ситуації. Однак даний підхід має наступні недоліки:

–  $m_2$ – $m_4$  відносяться до діагностичних ознак, які пов'язані із абсолютним значенням характеристик вимірювального параметру. У роботі [6] запропоновано метод на основі ВА та S-дискримінанту ( $I_d$ ), що забезпечує інваріантність отриманих оцінок та діагностування помпажних явищ;

– вибір вейвлета та рівень розкладання було обра-  
но емпіричним шляхом;

– аналіз зривних процесів виконувався тільки на основі  $P_2$ .

Тому необхідно розробити відповідну модифікацію даного методу для його ефективного застосування у ППБ САКД.

## 2 ПОСТАНОВКА ЗАДАЧІ

При побудові модуля попередження помпажу, необхідно виконати наступне:

– обрати тип вейвлета й параметри вейвлет-перетворення;

– побудувати модель газодинамічного стану двигуна, вхідними параметрами якої є дискримінантні оцінки поведінки ВК для набору сигналів вимірювальних систем, що визначають усталену роботу компресора, а вихідними – ознака виявлення помпажу й/або обертового зриву;

– визначити компоненти регулювання МПП при виникненні та й розвитку зривних процесів у ГТД.

## 3 ВИБІР ВЕЙВЛЕТУ ТА ЙОГО ПАРАМЕТРІВ

Визначення діапазону можливих частот коливань необхідно, насамперед, для розробки й налаштування засобів захисту двигунів під час помпажу та обертового зриву. Тому виникає задача вибору найбільш ефективного вейвлета, що володіє високою швидкодією й вибірковістю за частотою для пояснення фізики досліджуваного процесу.

Через здатність до швидкого алгоритму Маала у роботі були досліджені вейвети Добеші (db1–db44), Симлету (sym1–sym10), Койфлету (coif1–coif5) а також дискретний вейвлет Мейера.

Аналіз фільтруючих властивостей даних вейвлетів свідчить про те, що порядок вейвлета ( $N$ ) є еквівалентом характеристики зрізу смуги пропускання. Із збільшенням  $N$  крутість зрізу АЧХ вейвлетів збільшується і відповідно покращується якість реконструкції сигналів. Частотна характеристика вейвлетів Добеші у зрівнянні із вейветами Койфлету того ж порядку має більш низьку крутість зрізу смуги пропускання та збігається із Фур'є-образом Симлету. Дискретний вейвлет Мейера має найкращу характеристику зрізу смуги пропускання у порівнянні із db20 та coif5.

Крім того, для наочного трактування фізичної картини виникнення помпажу у турбокомпресорі ГТД до коефіцієнтів апроксимації та деталізації ВП було застосовано швидке перетворення Фур'є. У результаті частотне представлення ВК сигналу  $P_{\text{квт}}$  вейвлетів першого та другого порядків (особливо db2/sym2) на 7-му рівні дозволяє демонструвати зростання потужностей у частотних областях, що є характерними для помпажу та обертового зриву. Поступове збільшення порядку вейвлетів дає можливість спостерігати лише обертовий зрив коло частоти 100 Гц. Приведена особливість пояснюється шириною смуги пропускання Фур'є-образів вейвлетів на 7-му рівні.

Отримані результати дозволяють провести подальший аналіз зривних процесів за допомогою вейвлетів

із низьким порядком. Враховуючи високу швидкість та особливості опису виникнення помпажних явищ за допомогою швидкого перетворення Фур'є ВК сигналу  $P_{\text{квт}}$  для діагностування помпажу був обраний вейвлет db2. При цьому відслідковувати поведження ВК необхідно на 7-му рівні розкладання.

#### 4 ПОБУДОВА МОДУЛЯ ПОПЕРЕДЖЕННЯ ПОМПАЖУ

Для аналізу поточної газо-динамічної нестійкості (ГДН) двигуна та збільшення ефективності виявлення провісників помпажних явищ був обраний метод штучних нейронних мереж карт Кохонена (Self-organizing map, SOM). При виборі даного методу, насамперед, враховувалося те, що обертовий зрив не завжди переходить у помпаж і сформувавши правило такого переходу не можна. Принципи навчання SOM дозволяють аналізувати ті стани систем і процесів,

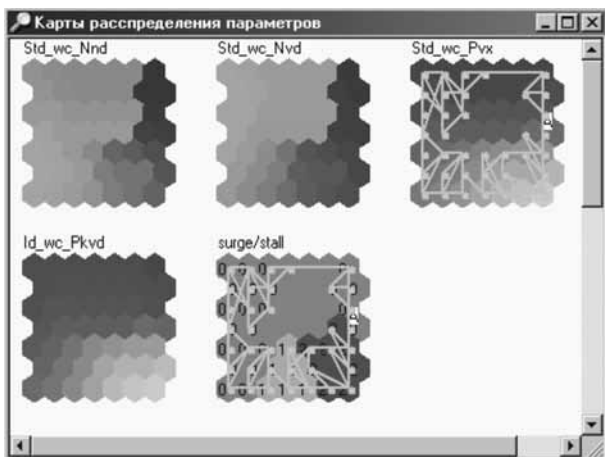


Рисунок 1 – Діагностування помпажу за допомогою слідкування за траєкторією робочої точки процесу ГДН

які важко визначити або інтерпретувати стандартними засобами математичної статистики [7].

При побудові моделі ГДН вихідними параметрами SOM були індекс перевищення порогу кліпування по дисперсії  $I_d$  для ВК сигналу  $P_{\text{квт}}$  та стандартне відхилення у (std) наступних сигналів:

- 1) частота обертання роторів низького тиску  $n_{\text{вд}}$ ;
- 2) частота обертання роторів високого тиску  $n_{\text{нд}}$ .

З метою підвищення швидкодії МПП для сигналів  $P_{\text{вх}}$ ,  $n_{\text{вд}}$ ,  $n_{\text{нд}}$  як оцінний параметр був прийнятий параметр у (std).

На виході моделі знімається інформація про газо-динамічний стан двигуна: 0 – нормальний режим роботи; 2 – обертовий зрив, 1 – помпаж.

У результаті навчання на SOM формуються зони, що відповідають перехідним процесам «нормальний режим роботи – обертовий зрив – помпаж».

На основі траєкторії руху робочої точки на SOM, можна визначити виникнення небезпечної ситуації типу помпаж при наблизенні даної точки до границі області обертового зриву (рис. 1).

Виходячи з виконаних розробок, модель МПП, можна представити в такий спосіб (рис. 2). На вхід модуля виявлення помпажу надходять сигнали від датчиків вимірювальних систем двигуна:  $P_{\text{квт}}$ ,  $P_{\text{вх}}$ ,  $n_{\text{вд}}$ ,  $n_{\text{нд}}$ . На основі вейвлет-розкладання наведених сигналів ('db2', рівні розкладання 7, 10) і аналізу S-дискримінант ВК по  $P_{\text{квт}}$  на SOM формуються вихідні дані про виявлення помпажних явищ. За допомогою SOM відбувається вироблення сигналу про газо-динамічний стан двигуна (0/1/2) на блок прийняття рішення. У даному блоці, виходячи з аналізу сигналів вимірювальних систем двигуна та сигналу – результату від SOM, формується керуючий вплив на блок протипомпажного регулювання, що включає антипомпажний клапан, насос-регулятор,  $\alpha$ -регулювання вхідного напрямного апарату. За рахунок раннього виявлення розвитку зривних процесів МПП не дозволяє розвиватися небезпечним явищам типу помпаж.

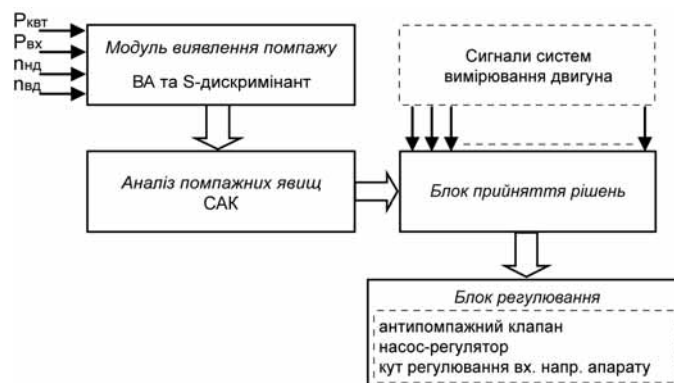


Рисунок 2 – Модуль попередження помпажу

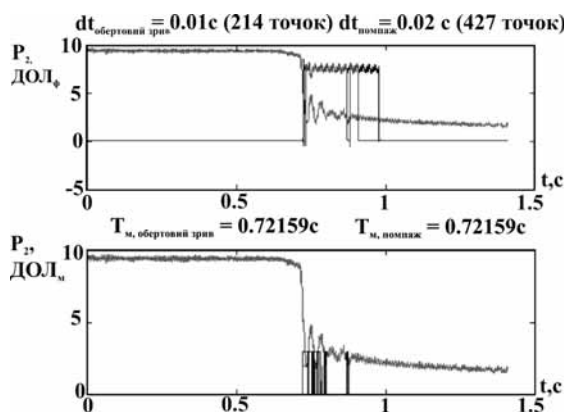


Рисунок 3 – Визначення зривних процесів за допомогою ДОЛ-32<sub>м</sub>:

червона лінія – визначення помпажу, чорна – обертового зриву;  $dt_{\text{обертовий зрив}}$ ,  $dt_{\text{помпаж}}$  – часові вікна для аналізу обертового зриву та помпажу відповідно;  $T_{\text{м, обертовий зрив}}$ ,  $T_{\text{м, помпаж}}$  – час визначення помпажу та обертового зриву за допомогою моделі

## 5 ЕКСПЕРИМЕНТАЛЬНІ РЕЗУЛЬТАТИ

У процесі дослідження було обрано дані стендових іспитів ГТД ( $P_{\text{квт}}$ ,  $P_{\text{вх}}$ ,  $n_{\text{вд}}$ ,  $n_{\text{нд}}$ ) і сигнал, що надходив від ДОЛ<sub>φ</sub> про наявність зривних процесів (рис. 3). Для навчання SOM, були обрані параметри  $I_d$  для ВК деталізації сигналу  $P_{\text{квт}}$  та у для сигналів  $P_{\text{вх}}$ ,  $n_{\text{вд}}$ ,  $n_{\text{нд}}$  (ширина тимчасового вікна 0,23 с, зсув 0,047 с).

Карта навчалась тільки на тих прикладах, для яких була характерно відсутність стану помпажу. На рис. 4 збільшення похибки апроксимації моделі (MSE) для даних по стендових випробуваннях ГТД відбувається на 205-му вимірі й поступово зростає. Зіставлення даних результатів з реальними значеннями показує, що з 210-го зразка у вибірці був присутній сигнал про виявлення помпажу.

Таблиця 1 – Результати виявлення помпажних явищ за допомогою сигналізаторів помпажу типу ДОЛ і методу на основі вейвлет-аналізу та SOM

№	Модель/метод	$T_{\text{ф}} - T_{\text{м}}$ , мс	Ширина часового вікна, с
1	ДОЛ	10	0,02
2	ПС	10	
3	ПС-2-7	9,6	
4	ВА db8	16,5	0,23
5	SOM та ВА 'db2'	230	0,23

Порівняльний аналіз роботи спрощених моделей ДОЛ<sub>м</sub>, ПС<sub>м</sub> ПС-2-7<sub>м</sub> та методу на основі ВА та SOM, реалізованих у середовищі С++, свідчить про те, що раніше за всіх розвиток помпажних явищ визначає

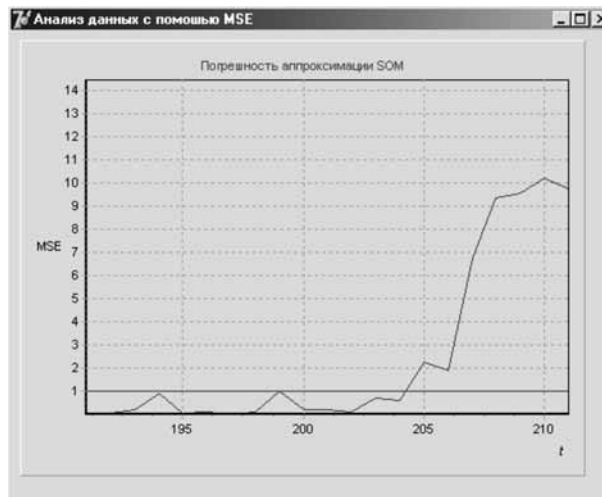


Рисунок 4 – Діагностування стану процесу ГДН за допомогою похибки навчання SOM

похибка навчання моделі ГДН – MSE (табл. 1), де  $T_{\text{ф}}$ ,  $T_{\text{м}}$  – фактичний та модельний час виявлення помпажних явищ за допомогою датчиків типу ДОЛ.

## ВИСНОВКИ

Таким чином, з метою попередження розвитку зривних процесів у турбокомпресорі запропоновано модель модуля запобігання помпажу у САК ГТД. Робота даного модуля побудована на діагностуванні помпажних явищ за допомогою ВА (із попередньою обробкою даних S-дискримінантом) та карт Кохонена SOM. Порівняльний аналіз роботи спрощених моделей ДОЛ<sub>м</sub>, ПС<sub>м</sub> ПС-2-7<sub>м</sub> та моделі газодинамічних процесів, що були отримані в результаті навчання SOM, свідчать про те, що раніше за всіх розвиток помпажних явищ визначає похибку навчання MSE.

## ПЕРЕЛІК ПОСИЛАНЬ

1. Письменный И. Л. Многочастотные нелинейные колебания в газотурбинном двигателе / Письменный И. Л. – М. : Машиностроение, 1987. – 128 с.
2. Руководство по технической эксплуатации ДОЛ РЭ (ДОЛ-16, ДОЛ-32, ДОЛ-32С), 1978. – 38 с.
3. Руководство по технической эксплуатации Преобразователя электронного ЭП-В, 1976/1979. – 45 с.
4. Руководство по технической эксплуатации ПС РЭ ПС-2-7, 1979. – 49 с.
5. Дремин И. М. Вейвлеты и их использование / Дремин И. М., Иванов О. В., Нечитайло В. А. // Успехи физических наук. – 2001. – Т. 171, № 5. – С. 493-495.
6. Киприч Т. В. Контроль помпажных явлений ГТД на основе вейвлет-анализа и дискриминантных признаков / Киприч Т. В., Дубровин В. И. // Вестник двигателестроения. – 2008. – № 1. – С. 166-169.
7. Саймон Хайкин. Нейронные сети : Полный курс : пер. с англ. / Саймон Хайкин. – 2-е издание. – М. : Вильямс, 2006. – 1104 с.

Надійшла 24.06.2008

Целью существующих систем защиты газотурбинного двигателя от срывных процессов является автома-

*тическое восстановление режима работы двигателя после устранения помпажа, а не предотвращение возникновения газодинамической неустойчивости в турбомашинах. Предложенная модель модуля предотвращения помпажа основана на методе вейвлет-анализа и картах Кохонена SOM. Приводится сравнительный анализ результатов обнаружения срывных процессов с помощью данного подхода и упрощенных моделей.*

*The aim of existing stall prevention systems in a gas-turbine engine is an automatic restoration of the engine operating conditions after a surge ceasing rather than prevention of a gas-dynamic instability in turbo machines. The model of a surge prevention modulus based on the method of wavelet-analysis and Kohonen's SOM maps is suggested. A comparative analysis for stall detection results is carried out with the help of such an approach and simplified models.*

УДК 004.056:378 (147)

М. Г. Коляда

## ИМИТАЦИОННОЕ МОДЕЛИРОВАНИЕ КАНАЛОВ СЕТЕЙ С ОЧЕРЕДЯМИ ПАКЕТОВ ЗАЩИЩЕННОЙ ИНФОРМАЦИИ

*В статье представлен механизм оптимизации каналов сетей с последовательными очередями пакетов защищенной информации на основе имитационного моделирования. Решение задачи представлено на основе вероятности и функция распределения прибывающих защищенных пакетов на один канал.*

заявки. Образовывается очередь. Задача прогноза и состоит в том, чтобы установить с возможной точностью взаимную зависимость между числом обслуживающих единиц и качеством обслуживания. Под *качеством обслуживания* понимают количество заявок получающих отказ, либо среднее время ожидания, либо среднюю длину очереди.

### ВВЕДЕНИЕ

Теория массового обслуживания предлагает несколько способов эффективного своего применения, и в первую очередь, избежание длинных очередей. Один из таких способов состоит в том, чтобы иметь только одну очередь, из которой клиенты поступают на обслуживание по принципу «первым пришел – первым обслужен». Эта система сейчас принята не только в теоретических расчетах радиотехнических и коммуникационных устройств, но и во многих ситуациях окружающей жизни: учреждениях и организациях, например, в банковских отделениях.

Исследователи В. Н. Задорожный, Е. С. Ершов, О. Н. Канева, А. А. Донец разрабатывали оптимизацию сетей передачи информации с учетом образующихся очередей [5, 6]. Но научных работ, связанных с проблемой оптимизации передачи защищенной информации по каналам связи сети с учетом образующихся очередей в настоящее время практически нет. Многие авторы, в том числе и перечисленные, используют либо свои алгоритмы оптимизации сетей с очередями, либо готовые алгоритмы, заложенные в профессиональных системах имитационного моделирования (типа GPSS World). Поэтому, специалисты, которые занимаются практическим внедрением и эксплуатацией систем с многоканальной связью, сталкиваются с проблемами несоответствия результатов моделирования известным аналитическим зависимостям.

*Имитационное моделирование* (ИМ) предоставляет универсальные возможности для исследования информационных систем с учетом разнообразных свойств и взаимосвязей их элементов. Вопросам оптимизации, с использованием имитационного моделирования в системах обработки информации уделяли внимание многие авторы, например [1, 2]. Разработку проблемы оценки эффективности защиты информации, в линиях радиосвязи разрабатывали исследователи [3, 4].

Появляются новые алгоритмы, совершенствуются средства взаимодействия пользователя с моделью, автоматизируются те этапы исследований, которые ранее были ручными, но все же имитационное исследование больше остается наукой и искусством, нежели инженерным средством для огромной армии системных аналитиков и инженеров. Имитация до сих пор остается делом профессионалов «симуляционистов», и это сдерживает ее огромные потенциальные возможности использования в реальной практической жизни. Поэтому в работе была поставлена *цель* –

В практической деятельности часто происходят ситуации, когда возникает массовый спрос на обслуживание какого-либо специального вида, причем, обслуживаемый объект, располагая лишь ограниченным числом обслуживающих единиц, не всегда способен немедленно удовлетворить все поступающие

© Коляда М. Г., 2009

сделать имитационное моделирование интуитивно понятным и легким в освоении для инженерных работников и проектировщиков сложных систем, для широкого круга специалистов (в том числе, и студентов).

В качестве инструмента выбран всем доступный табличный процессор *Excel*. Именно это приложение может быть использовано под конкретную проблему, которая должна решаться на «понятийном» уровне. Проблема должна «говорить» на языке пользователя, а инженер должен решать значимые и конкретные, а не гипотетически важные задачи. Это не значит, что в дальнейшем он не будет применять специальные интегрированные программные среды, с использованием хорошо отработанных и известных программных средств симуляторов, пакетов обработки статистики, программ мониторинга данных, и, конечно же, – системы имитационного моделирования GPSS World [7–10]. Наоборот, именно они и должны стать основной инструментальной базой, но лишь после того, как будет осмыслен механизм «ручной» прогонки построенной сложной модели.

### ПОСТАНОВКА ПРОБЛЕМЫ

Предположим, используется ведомственная линия радиосвязи, в которой для обеспечения эффективной защиты передаваемой информации от прослушивания и подмены информации используются методы криптозащиты. Для обработки поступающих заявок в такой системе необходимо некоторое время, причем оно зависит, прежде всего, от быстродействия работы технических устройств и алгоритмов, лежащих в основе криптографических методов шифрования и дешифрования информации. Кроме объективных факторов, влияющих на качество обслуживания заявок в системе, существуют и чисто «организационные» факторы, которые напрямую связаны со средней длиной очереди приходящих и обслуженных заявок.

Пусть такая сеть с очередями пакетов защищенной информации, узлами которой являются *системы массового обслуживания* (СМО) усиливается, путем ввода в ее состав  $m$  дополнительных каналов. Все каналы в сети технически идентичны, т. е. любой из вводимых каналов может быть добавлен в состав любой СМО. Различие вероятностных распределений времени обслуживания заявок в разных узлах сети связано с разной трудоемкостью соответствующих этапов обслуживания заявок. Каналы в составе одной СМО характеризуются одинаковым вероятностным распределением времени обслуживания, но разными вероятностями прибывающих защищенных пакетов к одному каналу узла сети. *Целью усиления сети является сокращение средней длины очереди прибывающих заявок.* Задача состоит в нахождении оп-

тимального распределения  $m$  дополнительных каналов по узлам сети, обеспечивающих минимум длины очереди по обслуживанию защищенных пакетов.

Были проведены предварительные тестовые замеры по отношению к прибывающим запросам в виде пакетов защищенной информации на каналы узлов сети. Известно, что потенциальные защищенные пакеты запросов приходят малыми группами на каналы узлов сети. Вероятности прибытия таких защищенных пакетов к каналам в течение суток с шестиминутным интервалом времени известны, и представлены в табл. 1. Также в таблице вычислены функции распределения (интегральные вероятности) прибытия одного, двух, трех и четырех таких пакетов.

Наша задача состоит в том, чтобы смоделировать круглосуточный прием защищенных пакетов на каналы узлов сети в течение недели с шестиминутным интервалом времени (это время, отводимое на обслуживание одного защищенного пакета).

Таблица 1 – Вероятности и функция распределения прибывающих защищенных пакетов

Количество прибывающих защищенных пакетов	Вероятность прибывающих защищенных пакетов к одному каналу, %	Функция распределения, %
0	60	0
1	6	76
2	12	82
3	4	94
4	2	98

На основе построенной модели необходимо определить интенсивность работы одного такого канала. Необходимо спрогнозировать необходимое количество таких каналов, чтобы число защищенных пакетов, стоящих в очереди у каждого из них не превышало 5 штук.

### РЕАЛИЗАЦИЯ ПОСТАВЛЕННОЙ ЗАДАЧИ

Смоделируем круглосуточный прием защищенных информационных пакетов в течение недели (7 дней по 24 часа) с 6-минутным интервалом на обслуживание одного защищенного пакета. Обозначим 1680 временных интервалов [(7 дней в неделе) × (24 часа в сутках) × (60 мин. в часе) / 6 (мин.)] с помощью диалогового окна *Прогрессия* табличного процессора *Excel* (рис. 1). Обнулیم ячейку A4, выделим ее курсором, а затем выполним команду *Правка – Заполнить – Прогрессия*.

В диапазоне J4:K8 создадим таблицу распределения вероятностей, как показано на рис. 2.

В ячейке K9 введем формулу

$$=СУММПРОИЗВ(J4:J8;K4:K8),$$

которая вычислит ожидаемое количество прибывающих защищенных пакетов к каналам узлов сети с 6-минутным интервалом на обслуживание одного такого пакета. Цифра, вычисляемая с помощью этой формулы равна в среднем половине предполагаемого защищенного пакета, то есть она на половину ниже нормы (с двойным запасом).

Для управления имитационной моделью, для каждого из 1680 временного интервала с помощью функции =СЛЧИС() сгенерируем случайные числа в столбце В (интервал В4:В1684) в пределах меняющегося интервала от 0 до 1. Введем эту формулу в ячейку В5 и скопируем ее в нижние ячейки до В1684, чтобы создать 1680 случайных величин. Для того, чтобы избежать изменяющихся значений рабочего листа при каких-либо изменениях в нем, а не только при целенаправленном пересчете рабочего листа (по клавише F9), установим переключатель *Вручную* во вкладке *Вычисления* в диалоговом окне *Параметры*, которое можно открыть с помощью команды *Сервис – Параметры* (рис. 3).

Режим ручного пересчета рабочего листа при нажатии на клавишу F9 заставляет заново вычислять все значения в формулах, которые этого требуют, о чем в строке состояния указывает слово «Готово», а затем на небольшое время появляется индикатор «Расчет ячеек» с меняющимися цифрами процентного выполнения.

Теперь создадим в диапазоне М4:Н8 таблицу функций распределения количества пришедших к каналам узлов сети защищенных пакетов, а в ячейку С5 в соответствии со значением случайного числа в столбце В впишем формулу =ВПР(В5;\$М\$4:\$Н\$8;2) (см. рис. 4), которая ищет в крайнем слева столбце массива (у нас это таблица абсолютных ячеек \$М\$4:\$Н\$8)

значение, совпадающее с заданным (у нас это элемент ячейки В5), и возвращает значение из указанного столбца массива и той строки, где найдено совпадение, в исходную ячейку (т. е. ту, где и записана эта формула). Иными словами, эта формула преобразует случайное сгенерированное десятичное число из интервала от 0 до 1 ячейки В5 в количественное целое число прибывших к серверу защищенных пакетов в ячейке С5. Затем распространим эту формулу на весь диапазон столбца С (интервал ячеек С5:С1684) (см. рис. 5).

Для одновременного просмотра верхней и нижней части рабочего листа разделим окно рабочей книги на две части, как показано на рис. 6.

Для завершения формирования таблицы значений листа введем формулы в ячейки D5:F1684, как показано на рис. 7.

Формулы в столбце D для каждого временного интервала вычисляют количество защищенных пакетов в очереди. Суммируя число вновь прибывших в текущем 6-минутном интервале и еще не обслуженных пакетов из предыдущего 6-минутного интервала, (если, конечно, таковые имеются) (=C5+F4).

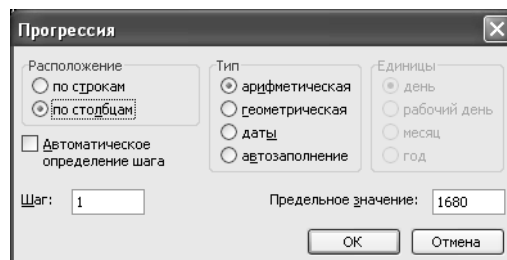


Рисунок 1 – Диалоговое окно «Прогрессия»

	А	В	С	Д	Е	Ф	Г	Н	И	Ж	К	Л	М
1	Прибытие защищенные пакетов к каналам узлов сети												
2	Количество каналов сети:												
3	Шести минутные интервалы	Сгенерированное случайное число	Количество прибывающих пакетов	Количество пакетов в очереди	Количество обслуживаемых пакетов	Количество ожидающих пакетов				Данные предварительного тестирования	Частоты прибытия		
4										Количество прибывающих пакетов	Вероятность	Функция распределения	
5										0	70%	0%	
6										1	6%	77%	
7										2	12%	83%	
8										3	4%	94%	
9										4	2%	98%	
10										Ожидаемое количество прибывающих пакетов			
11													
12													
13													
14													
15													

Рисунок 2 – Создание таблицы распределения вероятностей



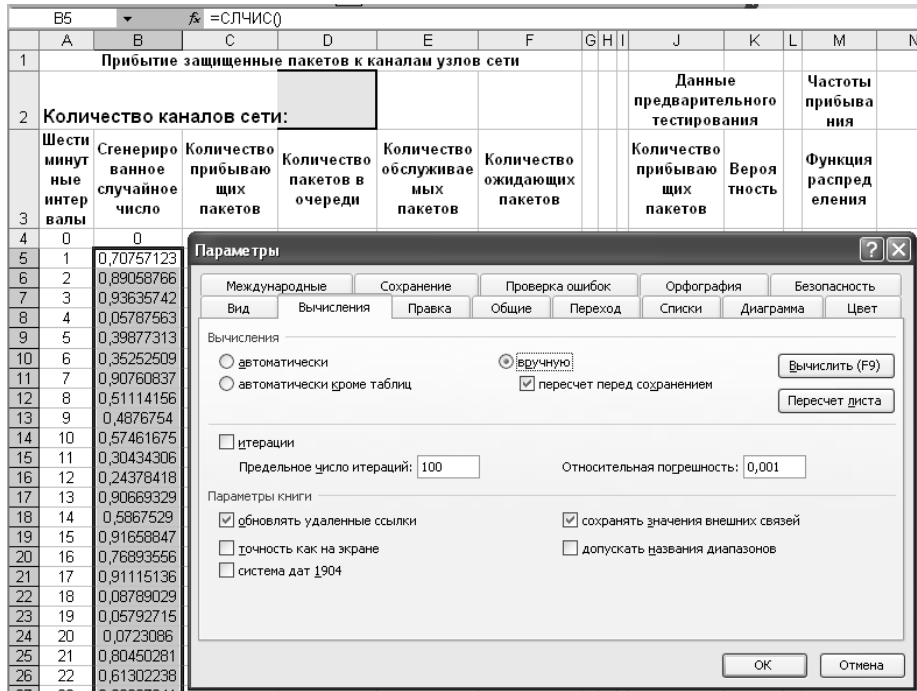


Рисунок 3 – Диалоговое окно «Параметры»

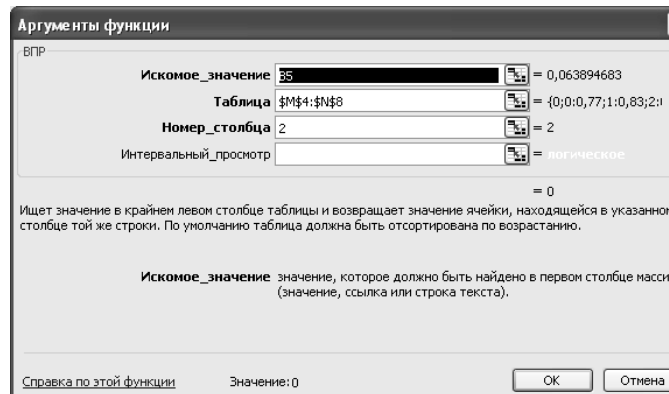


Рисунок 4 – Ввод аргументов функции ВПР

Прибытие защищенных пакетов к каналам узлов сети		Данные предварительного тестирования		Частоты прибытия пакетов		
Шести минутные интервалы	Сгенерированное случайное число	Количество прибывающих пакетов	Количество прибывающих пакетов	Вероятность	Функция распределения	Количество прибывающих пакетов
0	0	0	0	70%	0%	0
1	0,06389468	0	1	6%	77%	1
2	0,53446898	0	2	12%	83%	2
3	0,12252425	0	3	4%	94%	3
4	0,22317656	0	4	2%	98%	4
5	0,06051433	0	Ожидаемое количество прибывающих пакетов	0,5		
6	0,12744251	0				
7	0,14873327	0				
8	0,84002069	2				
9	0,18355419	0				

Рисунок 5 – Результат выполненной функции ВПР

D1685 =МАКС(D4:D1684)													
A	B	C	D	E	F	G	H	I	J	K	L	M	N
1	Прибытие защищенные пакетов к каналам узлов сети												
2	Количество каналов сети:			1					Данные предварительного тестирования		Частоты прибытия пакетов		
3	Шести минутные интервалы	Сгенерированное случайное число	Количество прибывающих пакетов	Количество пакетов в очереди	Количество обслуживаемых пакетов	Количество ожидающих пакетов			Количество прибывающих пакетов	Вероятность	Функция распределения	Количество прибывающих пакетов	
4	0	0	0	0	0	0		0	70%	0%	0		
5	1	0,038770246	0	0	0	0		1	6%	77%	1		
6	2	0,754037044	0	0	0	0		2	12%	83%	2		
7	3	0,220456294	0	0	0	0		3	4%	94%	3		
8	4	0,993302298	4	4	1	3		4	2%	98%	4		
9	5	0,104435203	0	3	1	2		Ожидаемое количество прибываю	0,5				
10	6	0,391259057	0	2	1	1							
11	7	0,96690744	3	4	1	3							
1682	1678	0,700298735	0	1	1	0							
1683	1679	0,452094641	0	0	0	0							
1684	1680	0,172782426	0	0	0	0							
1685		<b>Максимальные:</b>	4	10	1	9							
1686		<b>Всего:</b>	773	1753	773	980							
1687		<b>Среднее:</b>	0,45984533	1,042631648	0,45984533	0,582986318							

Рисунок 6 – Одновременный просмотр верхней и нижней части таблицы

C1685 =МАКС(\$C\$4:\$C\$1684)						
A	B	C	D	E	F	
1	Прибытие защищенные пакетов к каналам узлов сети					
2	Кол		1			
3	Шести минутные инт	Сгенерированное случайное число	Количество прибывающих пакетов	Количество пакетов в очереди	Количество обслуживаемых пакетов	Количество ожидающих пакетов
4	0	0	0	0	0	0
5	1	=СЛЧИС()	=ВПР(B5;\$M\$4:\$N\$8;2)	=C5+F4	=МИН(D5;\$D\$2)	=D5-E5
6	2	=СЛЧИС()	=ВПР(B6;\$M\$4:\$N\$8;2)	=C6+F5	=МИН(D6;\$D\$2)	=D6-E6
7	3	=СЛЧИС()	=ВПР(B7;\$M\$4:\$N\$8;2)	=C7+F6	=МИН(D7;\$D\$2)	=D7-E7
8	4	=СЛЧИС()	=ВПР(B8;\$M\$4:\$N\$8;2)	=C8+F7	=МИН(D8;\$D\$2)	=D8-E8
9	5	=СЛЧИС()	=ВПР(B9;\$M\$4:\$N\$8;2)	=C9+F8	=МИН(D9;\$D\$2)	=D9-E9
10	6	=СЛЧИС()	=ВПР(B10;\$M\$4:\$N\$8;2)	=C10+F9	=МИН(D10;\$D\$2)	=D10-E10
11	7	=СЛЧИС()	=ВПР(B11;\$M\$4:\$N\$8;2)	=C11+F10	=МИН(D11;\$D\$2)	=D11-E11
1683	1679	=СЛЧИС()	=ВПР(B1683;\$M\$4:\$N\$8;2)	=C1683+F1682	=МИН(D1683;\$D\$2)	=D1683-E1683
1684	1680	=СЛЧИС()	=ВПР(B1684;\$M\$4:\$N\$8;2)	=C1684+F1683	=МИН(D1684;\$D\$2)	=D1684-E1684
1685		<b>Максимальные:</b>	=МАКС(\$C\$4:\$C\$1684)	=МАКС(\$D\$4:\$D\$1684)	=МАКС(\$E\$4:\$E\$1684)	=МАКС(\$F\$4:\$F\$1684)
1686		<b>Всего:</b>	=СУММ(\$C\$4:\$C\$1684)	=СУММ(\$D\$4:\$D\$1684)	=СУММ(\$E\$4:\$E\$1684)	=СУММ(\$F\$4:\$F\$1684)
1687		<b>Среднее:</b>	=СРЗНАЧ(\$C\$4:\$C\$1684)	=СРЗНАЧ(\$D\$4:\$D\$1684)	=СРЗНАЧ(\$E\$4:\$E\$1684)	=СРЗНАЧ(\$F\$4:\$F\$1684)

Рисунок 7 – Просмотр таблицы в режиме формул

Формулы в столбце E вычисляют количество защищенных пакетов, обслуживаемых каналом узла сети, как минимум от длины очереди, которая может быть равна нулю, и числа таких каналов (задается в абсолютной ячейке \$D\$2 – у нас она равна 1): =МИН(D5;\$D\$2)).

Число пакетов, остающихся необслуженными в конце 6-минутного временного интервала, вычисляется в столбце F как разность между содержимым столбцов D и E (=D5-E5).

Формулы в диапазоне C1685:F1687 вычисляют статистические данные, необходимые для оценки эффективности работы одного канала узла сети. Эти формулы вычисляют: максимум, общее количество и среднее значение для каждого столбца.

Все представленные значения на рис. 6 являются случайными из-за вероятностной природы имитационного моделирования. Числа, которые получите вы,

будут отличаться от приведенных здесь, поскольку нажатие на кнопку F9 заполнит таблицу другими 1680 случайными числами. Повторное нажатие на клавишу F9 даст возможность проверить непостоянство этих величин.

Напомним, что в среднем только половина защищенного пакета (см. ячейку K9 на рис. 6) в каждый 6-минутный интервал прибывает к каналу сети, или в среднем каждые 12 минут прибывает один защищенный пакет на обслуживание. А чтобы обслужить этот пакет, напомним, требуется всего 6 минут. Таким образом, канал по обслуживанию будущих пакетов занят только наполовину времени, что, казалось бы, в нашей модели приводит к значительным простоям обслуживающей компьютерной техники. Но если поэкспериментировать, нажимая клавишу F9, то в течение недели возникают случаи, когда предельная загрузка канала узла сети составляет 12 и более защищен-

D1685		fx =МАКС(\$D\$4:\$D\$1684)												
	A	B	C	D	E	F	G	H	I	J	K	L	M	N
1	Прибытие защищенные пакеты к каналам узлов сети													
2	Количество каналов сети:			3	Данные предварительного тестирования									
3	Шести минутные интервалы	Сгенерированное случайное число	Количество прибывающих пакетов	Количество пакетов в очереди	Количество обслуживаемых пакетов	Количество ожидающих пакетов	Количество прибывающих пакетов		Вероятность	Функция распределения	Частоты прибывания пакетов			
4	0	0	0	0	0	0	0	0	70%	0%	0			
5	1	0,8948554	2	2	2	0	1	6%	77%	1				
6	2	0,5861994	0	0	0	0	2	12%	83%	2				
7	3	0,27561925	0	0	0	0	3	4%	94%	3				
8	4	0,06796913	0	0	0	0	4	2%	98%	4				
9	5	0,73612567	0	0	0	0	Ожидаемое количество прибывающих	0,5						
10	6	0,70156461	0	0	0									
11	7	0,27736009	0	0	0									
1683	1679	0,44194997	0	0	0	0								
1684	1680	0,60198867	0	0	0	0								
1685	Максималь		4	5	3	2								
1686	Всего:		866	905	866	39								
1687	Среднее:		0,515169542	0,53837	0,51517	0,0232								

Рисунок 8 – Для очередей не более 5 пакетов, необходимо 3 канала связи

ных пакетов в ожидании своей очереди. Поскольку обслуживание пакета занимает 6 минут, то последний пакет на обслуживании, должен ждать  $6 \times 12 = 72$  мин., что составляет 1 час и 12 мин.!

Выход из создавшейся ситуации очевиден – необходимо добавить один канал узла сети, что позволит обслуживать одновременно два защищенных пакета. Это должно привести приблизительно к двойному сокращению предполагаемых задержек, разумеется, за счет удвоения текущих расходов на обслуживание каналов.

Замените значение ячейки D2 с единицы на двойку и снова проведите эксперимент. Полученные статистические данные для модели с двумя каналами показывают, что максимальное количество защищенных пакетов, идущих на обслуживание с задержкой падает в 2–3 раза, а среднее количество таких пакетов, получающих обслуживание с задержкой, уменьшается почти в десять раз.

Напомним, что нам по условию задачи нужно было выполнить еще более жесткие требования: число защищенных пакетов, стоящих в очереди на каждом канале сети связи не должно превышать 5 штук. Эксперимент показывает, что для выполнения этого требования, количество каналов должно быть увеличено до трех (рис. 8).

В заключение необходимо отметить, что при реализации этой же задачи с помощью системы имитационного моделирования GPSS World выполнить ее было бы намного сложнее. Язык GPSS является высокоуровневым, поэтому при создании «обходного пути» надо учитывать, что любое нестандартное решение будет иметь следующие недостатки:

- усложнится логика работы модели;
- могут возникнуть проблемы с масштабируемостью модели;
- замедлится работа модели (ввиду появления дополнительных шагов).

## ВЫВОДЫ

Предлагаемый в статье метод оптимизации количества каналов связи по узлам сети с очередями защищенных пакетов можно применять для структурной оптимизации сети, содержащей сотни узлов и характеризуемой произвольными вероятностными распределениями интервалов поступления и обслуживания заявок. Он не гарантирует нахождения глобального оптимума, но, как показывают тестовые испытания, позволяет при усилении сети дополнительными каналами существенно снижать не только очереди на обработку защищенных пакетов, а и снижать время ответа и производить это быстро, без привлечения сложных программных сред. Основным достоинством метода является его скорость: он позволяет аналитически оптимизировать распределение каналов сети на доступном инструментальном уровне.

Необходимо отметить следующее:

- 1) данный подход с использованием имитационного моделирования строится на псевдодатчике случайных чисел;
- 2) чтобы получить более точные результат прогнозирования, необходимо задействовать механизмы точного генерирования случайных величин;
- 3) значения временных характеристик должны быть приближены (уменьшены) к реальным условиям (особенно время, отводимое на криптографическую обработку одного защищенного пакета).

## ПЕРЕЧЕНЬ ССЫЛОК

1. Габалин А. В. Оптимизационно-имитационный подход в задачах анализа и синтеза структуры распределенных систем обработки информации // Труды Института проблем управления РАН. Т. 2. – М.: 2005. – С. 78–85.
2. Рыжиков Ю. И. Имитационное моделирование в обосновании методик расчета многоканальных приоритетных систем // Материалы I Всероссийской конференции (ИММОД-2003), 23–24 окт. 2003 г. – С.-Пе-

- тербург: ФГУП ЦНИИТС. – Режим доступа: <http://www.gpss.ru/immod%2703/index.html>. – Название с экрана.
3. Зайцев И. В., Обухов А. Н., Демиденко В. Т., Мешкова А. Ф. Оценка эффективности защиты информации в линиях радиосвязи в условиях преднамеренных помех // Информация и безопасность. Вып. 3. – Воронеж: ВГТУ, 2002. – С. 115–118.
  4. Зайцев И. В., Молев А. А. Оценка эффективности систем радиосвязи в условиях преднамеренных помех на основе агрегативного имитационного моделирования их функционирования // Материалы II Всероссийской конференции (ИММОД-2007), 23–24 окт. 2007 г. ФГУП ЦНИИТС. – Режим доступа: <http://www.gpss.ru/immod07/doklad/20.html>. – Название с экрана.
  5. Задорожный В. Н., Донец А. А. Алгоритм структурной оптимизации сетей с очередями // Материалы II Всероссийской конференции (ИММОД-2007), 23–24 окт. 2007. – С.-Петербург: ФГУП ЦНИИТС. – Режим доступа: <http://www.gpss.ru/immod07/doklad/18.html>. – Название с экрана.
  6. Задорожный В. Н., Ершов Е. С., Канева О. Н. Использование адаптивных приближений в алгоритмах параметрической оптимизации сетей с очередями // Третья российская научно-практическая конференция по имитационному моделированию и его применению в науке и промышленности. – СПб.: ЦНИИТС, 2007. – Режим доступа: <http://www.gpss.ru/immod07/doklad/19.html>. – Название с экрана.
  7. Воробейчиков Л. А., Сосновиков Г. К. Методические указания для слушателей ФПКП по моделированию систем и сетей связи на GPSS/PC. – М.: МТУСИ, 1993. – 59 с.
  8. Бражник А. И. Имитационное моделирование: возможности GPSS WORLD. – СПб.: Реноме, 2006. – 439 с.
  9. Томашевский В., Жданова Е. Имитационное моделирование в среде GPSS. – М.: Бестселер, 2003. – 416 с.
  10. Руководство пользователя по GPSS World / Перевод с английского. – Казань: Мастер Лайн, 2002. – 384 с.

Надійшла 5.06.2008  
Після доробки 22.10.2008

*У статті представлено механізм оптимізації каналів мереж з послідовними чергами пакетів захищеної інформації на основі імітаційного моделювання. Рішення задачі представлено на основі вірогідності і функції розподілу захищених пакетів, які прибувають на один канал.*

*In the paper the gear of optimization of network's channels with sequential queues of packets of the protected information on the basis of simulation modeling is represented. The solution of a specific target is adduced in view of probability and cumulative distribution function arriving, protected packets to one channel.*

УДК 681.5.001.63: 519.711

А. А. Лисняк, С. И. Гоменюк

## ПРИМЕНЕНИЕ R-ФУНКЦИЙ ДЛЯ ГЕОМЕТРИЧЕСКОГО МОДЕЛИРОВАНИЯ ОБЪЕКТОВ СЛОЖНОЙ ФОРМЫ

*В статье рассматривается задача построения геометрических моделей инженерных конструкций нестандартной формы. Выполнен обзор основных методов представления геометрических моделей в современных САПР. Описано применение аппарата функций В. Л. Рвачева для формального описания и последующей дискретизации плоских геометрических объектов сложной формы.*

### ПОСТАНОВКА ПРОБЛЕМЫ

Построение и дискретизация адекватных геометрических моделей инженерных конструкций и сооружений является актуальной задачей, возникающей в различных приложениях численного анализа и САПР. На практике, как правило, приходится иметь дело с объектами сложной формы, что существенно затрудняет построение таких моделей и их последующую дискретизацию на конечные элементы. В то же время, от точности построенной геометрической модели будет напрямую зависеть качество полученного численного результата.

Данную проблему можно разбить на две самостоятельных задачи:

© Лисняк А. А., Гоменюк С. И., 2009

1) автоматизация построения геометрических (топологических) моделей инженерных конструкций сложной формы;

2) автоматизация разбиения топологической модели на конечные элементы заданного вида.

К настоящему времени предложены различные способы геометрического моделирования разнообразных тел двух- и трехмерной формы. Условно их можно классифицировать следующим образом:

– *параметрическое описание границы* (или поверхности), когда на специальном входном языке (например, GRUMMP [1]) или в специальном формате (например, STL [2]) описываются граничные сегменты, образующие замкнутую геометрическую фигуру;

– *твердотельное геометрическое моделирование*, когда трехмерная геометрическая модель реконструируется по чертежам трех ее проекций [3];

– *композиция топологической модели* необходимой конструкции, как некоторой совокупности базовых геометрических примитивов (например, линий, сплайнов, сфер и т. п.) и Эйлеровых операций (логическое объединение, пересечение и вычитание) над ними [4].

У каждого из этих подходов есть ряд недостатков. Параметрическое описание границы является достаточно трудоемким процессом для тел сложной формы (особенно пространственных). Твердотельное моделирование не всегда позволяет получить геометрическую модель, пригодную для последующей автоматической дискретизации. Композиция топологической модели из базовых геометрических примитивов ограничена их набором и не позволяет описывать произвольные области.

Таким образом, проблема создания универсального способа геометрического моделирования инженерных конструкций и сооружений произвольной формы по-прежнему остается открытой. Следовательно, возникает необходимость разработки альтернативных методов формального описания инженерных конструкций, особенно неклассической формы, и, соответственно, методов построения дискретных моделей представленным подобным образом. Одним из возможных вариантов описания геометрических тел является применение в этих целях аппарата теории R-функций.

В данной статье приведено описание разработанного подхода к геометрическому моделированию плоских областей произвольной формы с использованием математического аппарата функций В. Л. Рвачева, являющегося мощным аналитическим способом описания плоских и пространственных геометрических областей [5].

### МЕТОДЫ ФОРМАЛИЗАЦИИ ОПИСАНИЯ ТОПОЛОГИИ ОБЪЕКТОВ

Описание топологии геометрических объектов в инженерных задачах требует наличия формализованных правил, принципов и средств, которые будут использоваться в дальнейшем процессе построения дискретных моделей. Чаще всего для непосредственного задания геометрической области используются различные способы граничных представлений и конструктивная геометрия.

*Граничное представление* (BREP – Boundary Representation) является одним из наиболее простых и широко применяемых форматов описания топологии двух- и трехмерных геометрических объектов. Простота данного метода заключается в том, что граница (поверхность) объекта заменяется набором соответствующих отрезков (полигональных плоскостей), в той или иной мере приближающих границу исходной области. Примером такого подхода является универсальный формат IGES (Initial Graphics Exchange Specification), который используется для обмена данными между различными САПР и графическими приложениями. IGES является стандартом ANSI для передачи рисунков и трехмерных графических моделей. Однако, процесс построения описания топологии

объекта в формате IGES (особенно в трехмерном случае) является весьма трудоемким. Для его упрощения предлагается использовать R-функции.

Аппарат теории функций В. Л. Рвачева позволяет аналитически описать границу произвольного геометрического объекта. Процесс ее описания сводится к заданию некоторой функции с координатными аргументами, принимающей нулевые значения на границе области, положительные значения внутри области и отрицательные – за ее пределами. R-функция строится по определенным правилам. Сначала задаются «опорные функции», которые определяют необходимые базовые геометрические примитивы на плоскости или в пространстве, а затем при помощи R-операций строится их необходимая комбинация, образующая требуемую область (рис. 1).

Базовые R-операции объединения, пересечения и отрицания для образующих функций описываются так, как это показано в выражении

$$\{R\} \equiv \begin{cases} x \wedge y = x + y - \sqrt{x^2 + y^2}, \\ x \vee y = x + y + \sqrt{x^2 + y^2}, \\ \bar{x} = -x. \end{cases}$$

Для формального описания геометрических моделей с помощью R-функций в САПР необходимо наличие специализированного языка. В разработанной системе геометрического моделирования для этих целей используется язык FORTU-3, являющийся входным в системе FORTU-FEM [6]. Пример описания R-функции на языке FORTU-3 приведен на рис. 2.

$$\begin{aligned} F_1 &= 16 - x^2, \\ F_2 &= 16 - y^2, \\ F_3 &= 20 - x^2 - y^2, \\ F_4 &= y - e^{-x^2}, \\ F &= (F_1 \wedge F_2 \wedge F_3) \wedge (-F_4) \end{aligned}$$

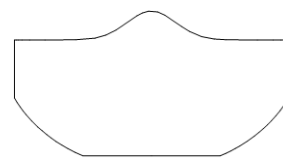


Рисунок 1 – Пример геометрической области и соответствующей ей R-функции

*!Волнистая лента*  
**Argument** x,y  
**Function** F, F1, F2, F3, F4, F5  
 $F=25-x*x-y*y$   
 $F1=y-\sin(x+0.5)$   
 $F3=y-2*\sin(x)$   
 $F4=F1-F3-(F1*F1+F3*F3)*0.5$   
 $F5=F+F4-(F*F+F4*F4)*0.5$



Рисунок 2 – Пример описания геометрической области с помощью языка FORTU-3

### АЛГОРИТМ ПОСТРОЕНИЯ КОНЕЧНО-ЭЛЕМЕНТНОЙ СЕТКИ

В данной статье рассматривается применение R-функций для компьютерного моделирования и дискретизации плоских геометрических областей. Автоматизация использования аппарата R-функций сопряжена с рядом проблем, связанных с исходным аналитическим способом описания области моделирования. Большинство существующих алгоритмов строят триангуляцию на базе заданного множества опорных точек. Получение набора опорных точек для аналитически заданной с помощью R-функции области является наиболее сложной проблемой, т. к. определение точек, принадлежащих нулевой границе области является весьма трудоемким процессом.

Дискретизация плоской области на треугольные конечные элементы сводится к построению триангуляции для заранее заданного набора точек на границе объекта с последующим добавлением новых узлов в процессе уточнения границ области, а также оптимизации полученной триангуляции. Таким образом, исходную проблему можно разбить на следующие задачи:

- формирование опорного набора точек;
- выбор алгоритма дискретизации;
- определение точности отображения границы;
- нахождение острых углов на границе области;
- оптимизация полученного разбиения.

*Опорный набор точек* – это узлы на границе области, описанной при помощи R-функции. Они необходимы для построения начальной (каркасной) триангуляции области моделирования. Нахождение опорного набора точек можно формально описать в виде следующей последовательности шагов:

- 1) выбор подходящей системы координат (например, декартовой или полярной) для построения регулярной сетки;
- 2) выбор заведомо покрывающего исходную моделируемую область начального бокса (форма которого зависит от выбранной системы координат);

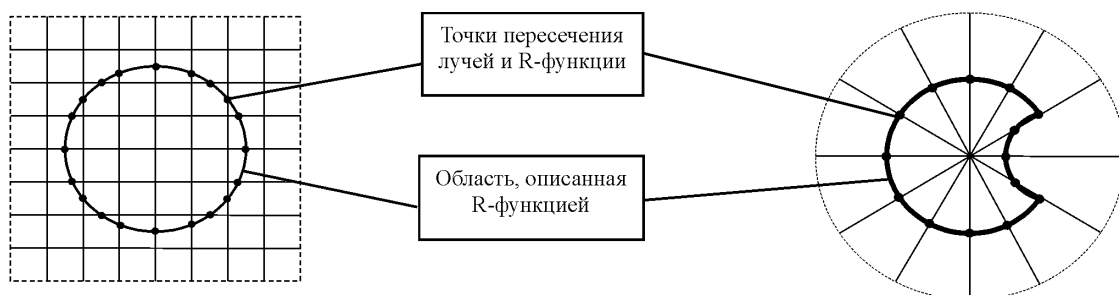


Рисунок 3 – Наложение регулярной сетки и определение множества опорных точек

3) формирование начального множества опорных точек путем наложения на бокс регулярной сетки и определения точек ее пересечения с границей исходной области (рис. 3).

Нахождение точек пересечения регулярной сетки с границей области моделирования является одной из самых трудоемких процедур работы алгоритма. Для любой точки пространства можно однозначно сказать, принадлежит ли она области моделирования или нет. Следовательно, применяя метод половинного деления отрезка с крайними точками, имеющими противоположные по знаку значения R-функции можно с некоторой заданной точностью определить точку, которая является одним из множества решений уравнения  $F(x, y) = 0$ .

На базе полученного набора опорных граничных точек можно построить первоначальную триангуляцию. Триангуляция Делоне для заданного набора точек на плоскости может быть выполнена достаточно большим количеством алгоритмов [7]. Наиболее простым с точки зрения программной реализации является семейство *итеративных алгоритмов*, кроме того, они обладают приемлемой вычислительной трудоемкостью. Работа алгоритма начинается с определения некоторой суперструктуры, которая заведомо содержит множество опорных точек триангуляции и может совпадать с первоначальным боксом. Основная идея алгоритмов данного класса заключается в добавлении вершин (узлов) в уже существующую триангуляцию. После чего осуществляется поиск треугольника, в который попадает данный узел, а также выполняется локальная проверка условия Делоне, часто приводящая к необходимости перестройки триангуляции. В зависимости от выбора способа ускорения поиска нужного треугольника получают различные итеративные алгоритмы.

Важным моментом при построении адекватной модели является точность, с которой будет отображаться граница описанной при помощи R-функции области. Пусть, например, необходимо с некоторой точностью отобразить границу  $\Gamma$  области  $\Omega$  (рис. 4).

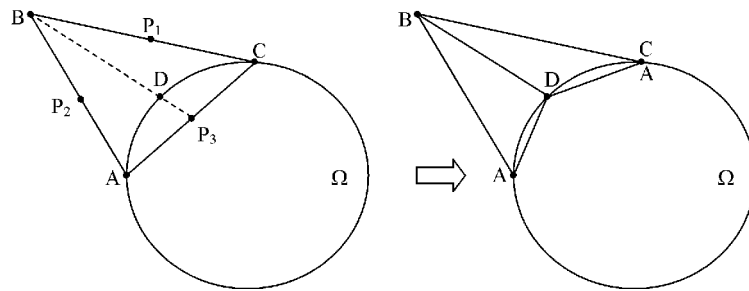


Рисунок 4 – Уточнение границы

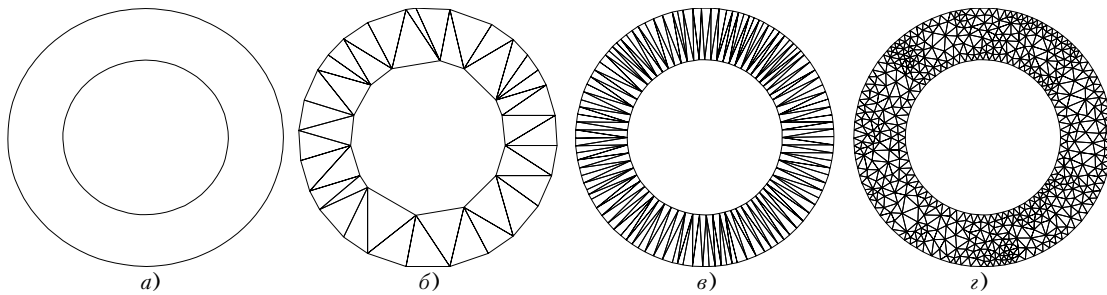


Рисунок 5 – Этапы работы алгоритма:

*a* – исходная область; *б* – опорные точки, полученные путем наложения регулярной сетки; *в* – уточненная граница; *г* – оптимизированная дискретизация

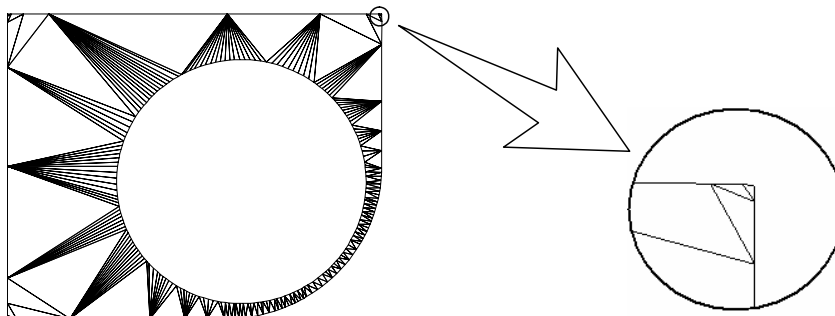


Рисунок 6 – Проблема нахождения и уточнения углов

**Argument**  $x, y$   
**BaseFunction**  $f, g$   
**Function**  $r$

$$f = F(x, y)$$

$$g = G(x, y)$$

$$r = R(f, g)$$

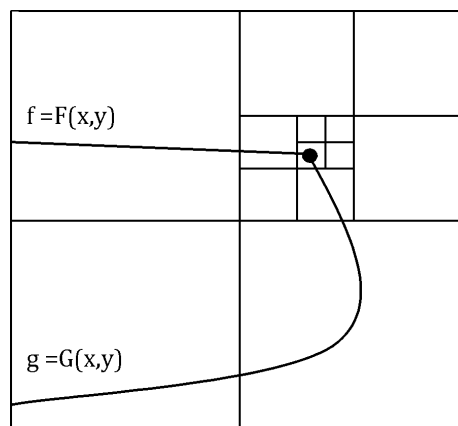


Рисунок 7 – Формальное описание некоторой области на языке FORTU-3 и алгоритм нахождения точек пересечения образующих (опорных) функций

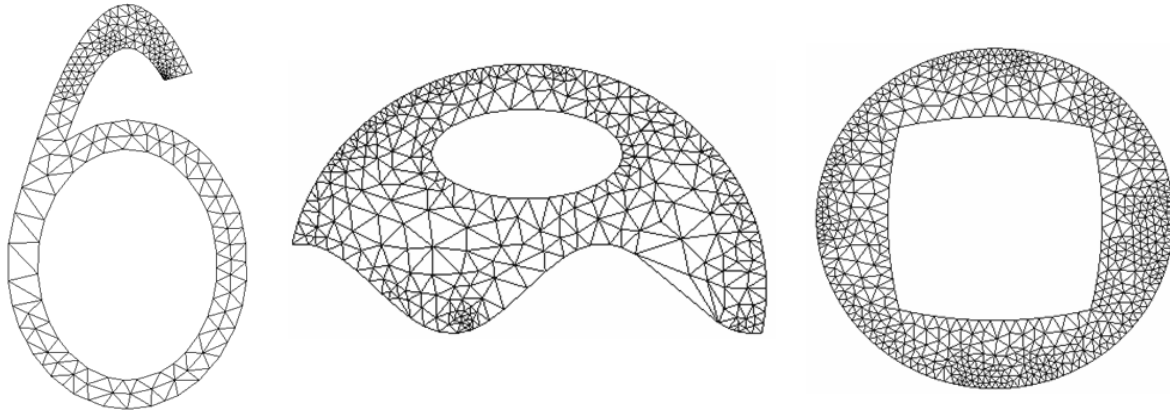


Рисунок 8 – Примеры дискретных моделей, описанных с помощью R-функций

Весь процесс нахождения точек на границе и внутри моделируемой области происходит параллельно с их вставкой в триангуляцию. Если граница  $G$  области  $\Omega$  является криволинейной, то можно утверждать, что в триангуляции всегда существует треугольник, серединам сторон которого соответствуют значения функции  $F$  разных знаков (например,  $ABC$ ). Обозначим через  $P_1$ ,  $P_2$ , и  $P_3$  – середины его сторон. Функция  $F$  принимает положительное значение в точке  $P_3$  и отрицательное в точках  $P_1$  и  $P_2$ . Отрезок  $AC$  в данном случае является граничным сегментом области  $\Omega$ . Тогда, проведя отрезок от точки  $P_3$  до противоположной ей вершины, можно найти точку его пересечения с границей  $G$  и, следовательно, заменить граничный сегмент  $AC$  на сегменты  $AD$  и  $CD$ . Аналогичные рассуждения, только с обратными знаками, справедливы для случая с невыпуклой границей. Процесс необходимо повторять до тех пор, пока абсолютное значение функции  $F$  для середин всех граничных сегментов не будет превышать некоторого наперед заданного значения.

Такая методика уточнения границ является эффективной для областей с гладкой границей. На рис. 5 приведены этапы построения дискретной модели геометрической подобной геометрической области. Как видно из рисунка происходит довольно точная аппроксимация криволинейной границы, но при использовании приведенного выше алгоритма уточнения границы существует проблема выявления острых углов, т. е., в общем случае все углы при увеличении масштаба не имеют остроугольной вершины (рис. 6). Эту проблему можно решить, если принять утверждение, что все углы получаются путем пересечения границ образующих функций.

Для нахождения точки пересечения двух образующих функций можно применить рекурсивный алгоритм разбиения области, но не для всей R-функции, а для всего множества возможных пар образующих функций. Для реализации данного подхода в язык

FORTU-3 был введен дополнительный оператор декларации образующих функций – **BaseFunction**. С его помощью при описании области моделирования пользователь может изначально описать все образующие функции, а затем, используя функции, объявленные идентификатором **Function**, определить различные комбинации (объединение, пересечение, вычитание) с образующими функциями для получения требуемого результата. Пример описания некоторой области в общем виде и несколько шагов работы алгоритма поиска точек пересечения образующих функций показаны на рис. 7.

Предложенный алгоритм геометрического моделирования и дискретизации плоских областей с помощью аппарата функций В. Л. Рвачева реализован программно. Некоторые примеры его работы приведены на рис. 8.

## ЗАКЛЮЧЕНИЕ

Предложенный подход к использованию R-функций для геометрического моделирования плоских областей в отличие от других существующих методов и алгоритмов позволяет достаточно легко и эффективно описывать геометрические модели плоских областей произвольной формы с их последующей триангуляцией. Кроме того, итоговая конечно-элементная модель за счет нерегулярной структуры имеет меньшую размерность: сетка автоматически сгущается в окрестностях особенностей (сингулярностей) топологии области, таких, например, как острые углы или отверстия, а в остальных местах содержит меньшее количество элементов, максимальный размер которых задается пользователем.

Развитие предложенного подхода на трехмерный случай позволит существенно повысить качество геометрических моделей реальных инженерных конструкций и сооружений.



### ПЕРЕЧЕНЬ ССЫЛОК

1. GRUMMP – Generation and Refinement of Unstructured, Mixed-Element Meshes in Parallel [Электронный ресурс] – Режим доступа : <http://tetra.mech.ubc.ca/GRUMMP/>.
2. *Останин А. Н.* Применение математических методов и ЭВМ. Вычислительные методы проектирования оптимальных конструкций: Учеб. пособие для вузов / Под общ. ред. А. Н. Останина. – Мн.: Выш. шк., 1989. – 279 с.
3. *Куприков М. Ю.* Твердотельное моделирование – новый подход к вопросам проектирования и подготовки технической документации [Электронный ресурс] / М. Ю. Куприков. – Режим доступа : <http://nit.miem.edu.ru/cgi-bin/article?id=76>.
4. *Голованов Н. Н.* Геометрическое моделирование / Н. Н. Голованов. – М.: Издательство физико-математической литературы, 2002. – 472 с.
5. *Рвачев В. А.* Теория R-функций и некоторые ее приложения / В. А. Рвачев. – К. : Наук. думка, 1982. – 552 с.
6. *Толок В. А.* Метод конечных элементов: теория, алгоритмы, реализация / В. А. Толок, В. В. Киричевский, С. И. Гоменюк, С. Н. Гребенюк, Д. П. Бувайло. – К. : Наук. думка, 2003. – 316 с.

7. *Скворцов А. В.* Эффективные алгоритмы построения триангуляции Делоне / А. В. Скворцов, Ю. А. Костюк // Геоинформатика. Теория и практика. – Томск : Изд-во Томского ун-та, 1998. – Вып. 1. – С. 22–47.

Надійшла 4.08.2008

*В статті розглянуто задачу побудови геометричних моделей інженерних конструкцій нестандартної форми. Виконано огляд основних методів представлення геометричних моделей у сучасних САПР. Описано застосування апарату функцій В. Л. Рвачеві для формального опису та наступної дискретизації плоских геометричних об'єктів складної форми.*

*The problem of plotting geometrical models for engineering structures of non-standard form is considered in the paper. The review of the basic methods to describe geometrical models in modern CAD is done. Application of V. I. Rvachev's function apparatus for formal description and subsequent discretization of plane geometrical objects of complex form is discussed.*

УДК 004.4'24

М. Ю. Терновой

## ПІДХІД ДО ПРЕДСТАВЛЕННЯ ІНФОРМАЦІЇ В ТЕРМІНАХ ПРЕДМЕТНОЇ ОБЛАСТІ

*Запропоновано підхід до отримання інформації з бази даних та її представлення в термінах предметної області. До складу цього підходу входять метод прив'язки до предметної області та метод побудови запиту до бази даних в термінах предметної області.*

### ВСТУП

Для ефективного управління будь-якою організаційною структурою керівнику необхідна різноманітна інформація, причому сутність цієї інформації та форма її подання будуть змінюватися в залежності від задачі, що вирішується. Від оперативності отримання та зручності подання необхідної інформації залежить своєчасність управлінських рішень [1]. На сьогоднішній день інформаційні технології складають основу інформаційно-аналітичної діяльності будь-якої організації. В сучасних інформаційних системах (ІС), що функціонують в організаціях, дані зберігаються в базах даних (БД), які найчастіше є реляційними [2]. Тому описана вище задача подання інформації зводиться до отримання даних з бази даних та їх представлення у формі, що вимагається.

Як правило, у впроваджених в організаціях ІС закладена можливість отримання звітів на основі поточних даних, але кількість різних за формою звітів обмежена. Створення нових звітів вимагає вдосконалення або переробки ІС за участі розробників.

З іншого боку на ринку програмного забезпечення існує багато систем, які надають можливість формувати різноманітні звіти, так звані генератори звітів (ГЗ). Генератори звітів існують як у вигляді самостійних програм, так і у вигляді генераторів звітів, що вбудовані в системи керування базами даних (СКБД) або середовища розробки. Серед останніх можна виділити [3–6]: MS SQL Server, MS Access, 1С. Аналогічні інструменти є в Delphi та інших середовищах розробки. Однак, користуватися вбудованими ГЗ можна лише за умови, що і вся система побудована на цій СКБД, або за допомогою цього середовища розробки.

Серед ГЗ, які є самостійними програмами, можна виділити [7–10]: MS SQL Server Reporting Services, Seagate Crystal Reports, VSREPORT, Report Sharp-Shooter та інші.

До загальних недоліків використання існуючих ГЗ можна віднести те, що для їх використання необхідні спеціальні знання в області інформаційних технологій та теорії баз даних. І хоча людина, що приймає рішення, є достатньо високо кваліфікованою в своїй предметній області, в області роботи з базами даних вона може не мати потрібних знань. Необхідність таких знань пов'язана з тим, що для отримання інформації з бази даних необхідно побудувати SQL-запит на вибірку для чого потрібно знати не тільки

саму мову SQL, а також необхідно знати структуру бази даних. Тому за відсутності таких знань очевидною стає необхідність звертатися до IT-спеціалістів, що підвищує вартість та знижує оперативність.

Виходячи з вищенаведеного актуальною є задача розробки ГЗ, який би дозволив працювати з даними з БД в термінах предметної області. Для цього, перш за все, необхідно розробити теоретичні основи такого ГЗ. Тому для вирішення цієї задачі нижче запропоновано підхід до отримання інформації з бази даних та її представлення, використовуючи лише терміни предметної області. Під терміном предметної області розуміється словесний опис тієї інформації, яка зберігається в базі даних та може бути отримана як результат виконання SQL-запиту.

### 1 ПРОБЛЕМА ПОБУДОВИ SQL-ЗАПИТУ НА ВИБІРКУ

В мові SQL запит на вибірку визначається оператором SELECT. Тоді можна сказати, що проблема полягає в автоматичній побудові SQL-запитів SELECT, при умові що користувач задає лише терміни предметної області та відповідні обмеження на них. Рішення цієї проблеми полягає в розробці підходу до формування такої службової інформації, та способу побудови запиту на основі цієї інформації, використовуючи які ГЗ зможе автоматично будувати запити на вибірку.

Запишемо загальний формат SQL-запиту на вибірку [2], децю спростивши його та отримаємо SQL-запит наступного вигляду (1):

```
SELECT column_expression
FROM table_name
[WHERE condition [AND column_condition]] (1)
```

При побудові багатотабличного запиту необхідно в операторі FROM вказати всі таблиці, що використовуються в запиті, а також проміжні таблиці, через які проходить зв'язок між таблицями, що не пов'язані напряму, а в операторі WHERE необхідно вказати як саме зв'язані таблиці *condition*, а також умови на значення полів *column\_condition*, які входять в *column\_expression*. При виникненні помилок в побудові запиту можуть бути отримані помилкові дані. Наприклад, у випадку необхідності виведення даних з полів двох незв'язаних таблиць буде отримано декартовий добуток всіх значень поля з однієї таблиці на всі значення поля другої таблиці.

### 2 ПРИВ'ЯЗКА ДО БАЗИ ДАНИХ

Для надання можливості використання термінів предметної області при проведенні обробки даних

пропонується робити одноразову прив'язку ГЗ до бази даних. Для цього заповнюються таблиця прив'язки (ТПр) термінів предметної області (ПО) до БД (табл. 1) та таблиця зв'язків (ТЗв).

Таблиця 1 – Таблиця прив'язки термінів ПО до бази даних

Термін предметної області	Таблиця, в якій зберігаються дані	Назва поля в таблиці, яке відповідає критерію

В таблиці ТЗв перший стовпчик та перший рядок – це назви таблиць, що є в базі даних та будуть використовуватися в запитах, а на перетині стовпчика з рядком, у випадку, коли є прямиий зв'язок між таблицями, записуються поля таблиці, що знаходиться у верхньому рядку, та які пов'язані з відповідними полями таблиці, що знаходиться в першому стовпчику. Поля таблиць організуються у вигляді масивів. У випадку коли зв'язку немає, поле залишається порожнім. На діагоналі знаходяться порожні поля.

Під прив'язкою ГЗ до БД буде розумітись заповнення ТПр та ТЗв. Повною прив'язкою назвемо таке заповнення таблиці ТПр та ТЗв, виходячи з якого можна побудувати вірний SQL-запит на вибірку.

Множиною імен таблиць буде називатись впорядкована множина потужністю  $N$ , елементами якої є назви таблиць в базі даних  $TN = \{TableName_1, TableName_2, \dots, TableName_N\}$ .

Під відображенням зв'язності буде розумітись  $C:TN \times TN \rightarrow \{0, 1\}$ , де  $C(TableName_i, TableName_j) = 1$  у випадку коли таблиця  $i$  зв'язана з таблицею  $j$ , та  $C(TableName_i, TableName_j) = 0$  у протилежному випадку.

Матрицею зв'язності  $\hat{C}$  буде називатись  $N \times N$  розмірна матриця, яка відповідає відображенню  $C$ , та кожен елемент якої вказує на те, чи існує прямиий зв'язок між таблицями. Ця матриця будується на основі таблиці зв'язків заміною непустиого поля на одиницю, а пустиого на нуль.

Головною вимогою до прив'язки є її повнота, тобто прив'язка повинна гарантувати, що буде отримано вірний запит.

Схему даних БД можна представити у вигляді зваженого графу  $G = (V, E)$ , де замість таблиць слід поставити вершини, а замість зв'язків між таблицями – ребра з одиничною вагою. Множина вершин  $V$  графу  $G$  буде дорівнювати множині назв таблиць  $TN$ , а матриця суміжності  $A$  графу  $G$  буде відповідати матриці зв'язків  $C$  (2).

$$V = TN, A = C. \quad (2)$$

Тоді пошук зв'язку між таблицями БД можна інтерпретувати як знаходження мінімального підграфу  $H$  графу  $G$ , який буде містити всі вершини, що відповідають таблицям, які входять до запити. Така задача носить назву задачі про мінімальне зв'язування [11].

Для отримання вірного та однозначного запиту необхідно існування однозначного зв'язку між таблицями. Адже, як було описано вище результатом виконання SQL-запиту, який побудовано по таблицях, що не зв'язані напряму, та між якими не вказано зв'язок через інші таблиці, буде декартовий добуток значень одного поля на значення другого. До того ж, якщо зв'язок є неоднозначним, тобто може реалізуватися через різні таблиці, то з теорії побудови реляційних баз даних та реляційної алгебри буде слідувати можливість побудови різних запитів, використовуючи які будуть отримані різні результуючі дані з БД. Якщо розглядати відповідний граф  $G$ , то це буде означати існування всього одного ланцюгу між кожними двома вершинами та, як наслідок, одиничність підграфу  $H$ .

З теорії графів відомо, що необхідною та достатньою умовою для існування між кожними двома вершинами тільки одного ланцюгу є зв'язність відповідного графу, а також відсутність циклів, тобто відповідний граф повинен бути деревом [11, 12].

Виходячи з вищесказаного, слід відмітити, що для доведення повноти прив'язки слід визначити, чи є граф  $G$ , який відповідає структурі БД та визначається матрицею суміжності  $A = C$ , деревом. Перевірка цього факту проводиться описаним нижче методом.

Метод перевірки того, що прив'язка є повною, можна розділити на три частини.

По-перше, виходячи з ТПр (табл. 1) та ТЗв, слід побудувати матрицю зв'язності  $\hat{C}$ , метод побудови якої описано вище, та визначити граф  $G$ , який буде відповідати структурі даних БД предметної області. Як було показано раніше, матриця суміжності  $A$  буде дорівнювати матриці зв'язності  $A = C$ .

По-друге, перевірити, чи є граф  $G$  зв'язним. Перевірка зв'язності графу  $G$  проводиться виходячи з теореми [12], яка говорить, що граф  $G$ , який визначається матрицею суміжності  $A$ , зв'язний тоді і тільки тоді, коли  $\hat{A}_{ij}^I = 1$  для всіх його вершин  $i$  і  $j$ . Де  $\hat{A}^I = I \vee A$ ,  $I$  – мультиплікативна одинична матриця.

$$A = A \vee A^{\otimes 2} \vee \dots \vee A^{\otimes n}, \quad A^{\otimes k} = A^{\otimes k-1} \otimes A, \\ A^{\otimes 1} = A, \quad k = \overline{1, n}.$$

$B = C \vee D$  означає  $B_{ij} = C_{ij} \vee D_{ij}$ , де  $B$ ,  $C$ ,  $D$  – матриці розмірністю  $m \times s$ .

$F = C \otimes E$  означає  $F_{ij} = (C_{i1} \wedge E_{1j}) \vee (C_{i2} \wedge E_{2j}) \vee \dots \vee (C_{is} \wedge E_{sj})$ ,  $i = \overline{1, m}$ ,  $j = \overline{1, p}$ , де  $F$ ,  $C$ ,  $E$  – матриці, які мають розмірність  $m \times p$ ,  $m \times s$ ,  $s \times p$  відповідно.

Тоді, перевірка зв'язності буде здійснюватись таким чином:

1. Знайти матрицю  $A$ , використавши алгоритм Уоршола [12];

2. Визначити матрицю  $\hat{A}^I = I \vee \hat{A}$ ;

3. Перевірити необхідну та достатню умову зв'язності  $\hat{A}_{ij}^I = 1$ ,  $i = \overline{1, n}$ ,  $j = \overline{1, n}$ .

По-третє, у випадку зв'язності графу перевірити, чи є граф  $G$  деревом. Для цього слід скористатися теоремою, яка говорить про те, що зв'язний граф є деревом у випадку, коли кількість його вершин більше на одиницю, ніж кількість його ребер, тобто  $|V| = |E| + 1$ .

Та якщо граф  $G$  є деревом, то робиться висновок про те, що прив'язка є повною.

В графі  $G$  напряму не задані ребра, зв'язки між вершинами графу визначаються через матрицю зв'язності  $A$ . Для того, щоб скористатися згаданою вище теоремою, слід визначити кількість ребер. Вона буде дорівнювати сумі елементів матриці  $A$ , які знаходяться вище головної діагоналі:

$$|E| = \sum_{i=2}^{N-1} \sum_{j=i+1}^N A_{ij}.$$

### 3 ПОБУДОВА ЗАПИТУ ДО БАЗИ ДАНИХ

Як було зазначено вище, основною проблемою при побудові SQL-запиту є знаходження зв'язку між таблицями бази даних, що не зв'язані між собою напряму, та, виходячи із знайденого зв'язку, формування аргументів операторів FROM та WHERE.

Постановка цієї задачі може бути записана наступним чином. Базуючись на ТПр, ТЗв,  $\hat{C}$  та  $TN = (TableName1, TableName2, \dots, TableNameN)$  та виходячи з вхідних даних  $\hat{D}_{I_F} = \{(T_{ik_l}, cond_{ik_l}) | l = \overline{1, n_F}\}$ , якими є обрані користувачем терміни предметної області  $T_{ik_l}$  та задані обмеження на них  $cond_{ik_l}$ , побудувати SQL-запит на вибірку.

Метод рішення задачі динамічної побудови запиту полягає в послідовному розв'язку двох більш дрібних задач. По-перше, визначення полів таблиць, що відповідають термінам предметної області, та формування *column\_expression* і завдання умов на значення полів *column\_condition* в (1). По-друге, визначення зв'язку між таблицями та формування *table\_name* і *condition* в (1). Далі розглянуто більш докладно постановки та методи розв'язання цих задач.

Постановка першої задачі полягає в тому, що виходячи з ТПр та  $\hat{D}_{I_F} = \{(T_{ik_l}, \text{cond}_{ik_l}) | l = 1, n_F\}$ , необхідно визначити *column\_expression* і *column\_condition*.

Розв'язок цієї задачі полягає в співставленні кожному терміну ПО  $T_{ik_l}$  з  $\hat{D}_{I_F}$  відповідної назви поля *ColumnName* <sub>$t_l, r_l$</sub>  та таблиці *TableName* <sub>$t_l$</sub>  з БД, які знаходяться в ТПр. Після цього *column\_expression* формується простим переліком через кому назв відповідних полів, а *column\_condition* переліком умов на значення визначених раніше полів, які записуються визначеним в мові SQL способом (3).

$$\begin{aligned} & \text{column\_expression} = \\ & = \text{TableName}_{t_1}.\text{ColumnName}_{t_1, r_1}, \\ & \text{TableName}_{t_2}.\text{ColumnName}_{t_2, r_2}, \dots, \\ & \text{TableName}_{t_{n_F}}.\text{ColumnName}_{t_{n_F}, r_{n_F}}; \\ & \text{column\_condition} = \\ & = \text{cond}_{ik_1}(\text{TableName}_{t_1}.\text{ColumnName}_{t_1, r_1})\text{AND} \\ & \text{cond}_{ik_2}(\text{TableName}_{t_2}.\text{ColumnName}_{t_2, r_2})\text{AND} \dots \\ & \text{AND cond}_{ik_{n_F}}(\text{TableName}_{t_{n_F}}.\text{ColumnName}_{t_{n_F}, r_{n_F}}). \end{aligned} \quad (3)$$

Наступним кроком формується множина назв таблиць  $\text{Table}_N = \{\text{TableName}_{t_l} | l = 1, n_F\}$ , поля яких відповідають обраним користувачем термінам ПО.  $\text{Table}_N$  буде використовуватись при розв'язку наступної задачі.

Постановку другої задачі можна записати наступним чином. Виходячи з  $\text{Table}_N$ ,  $C$  та  $TN$  визначити зв'язок між таблицями та сформувані *table\_name* і *condition* в (1).

Як відмічалось раніше, пошук зв'язку між таблицями БД можна інтерпретувати як розв'язок задачі про мінімальне зв'язування в теорії графів. В цьому випадку постановку другої задачі можна переписати наступним чином: для графу  $G = (V, E)$  необхідно визначити частковий підграф  $H = (V', E')$ , який буде містити вказані в деякій множині  $Q = \text{Table}_N$  вершини з найменшою сумарною вагою ребер. Та, використовуючи підграф  $H$ ,  $TN$  та ТЗв, визначити *table\_name* і *condition* в (1). Слід відмітити, що  $|Q| > 1$ .

В загальному випадку першу частину цієї задачі пропонується розв'язувати шляхом вирішення задачі про найкоротший ланцюг для вершин множини  $Q$ , вказуючи як джерело по черзі кожному вершину  $v_Q \in Q$  [11]. Тоді для кожної вершини  $v_V \in V$  будуть обрані  $|Q|$  чисел – відстаней від  $v_Q \in Q$ . Просумувавши їх для всіх вершин, визначається вершина з найменшою сумою. Через неї вершини з  $Q$  з'єднуються найкоротшими шляхами, що й дає розв'язок поставленої задачі.

Однак, оскільки граф  $G$  є деревом можна запропонувати простіший підхід для визначення підграфу  $H$ . Ідея такого підходу полягає в обранні довільної вершини  $v_{i_1} \in Q$  в якості кореня, та введення відображення нумерації *pred* вершин графу  $G$  таким чином, що кожній не кореневій вершині ставиться у відповідність інформація про номер вершини, яка знаходиться вище, та з якою вона пов'язана (4). Таку вершину будемо називати батьківською. Оскільки  $G$  дерево, то задане таким чином відображення *pred* однозначне:

$$\begin{aligned} & \text{pred}: V \setminus v_{i_1} \rightarrow [1; N], \quad \forall v_j \in V \setminus v_{i_1}, \\ & \exists! \text{pred}(v_j) = i: (v_i, v_j) \in E. \end{aligned} \quad (4)$$

Процес перенумерації, тобто визначення відображення *pred*, проводиться по вершинах графу  $G$  до тих пір, доки не перенумеровані всі вершини з множини  $Q$ . Тобто множина вершин, для яких задана нумерація  $V_{\text{num}}$ , та на яких буде визначено відображення *pred*, в загальному випадку є підмножиною множини вершин  $V$  графу  $G$  (5). Множину номерів вершин, які складають множину припустимих значень відображення *pred* позначимо через  $U$ :

$$\begin{aligned} & V_{\text{num}} \subseteq V, \quad \text{pred}: V_{\text{num}} \rightarrow U, \quad \forall v_j \in V_{\text{num}}, \\ & \exists! \text{pred}(v_j) = i: (v_i, v_j) \in E. \end{aligned} \quad (5)$$

Далі вводиться відображення розмітки вершин *mark*, що визначено на множині вершин  $V_{\text{num}} \subseteq V$  та дає одиницю, якщо вершина  $v_s \in V_{\text{num}}$  є частиною ланцюгу  $P_{i_k i_1}$ , який з'єднує вершину  $v_{i_k} \in Q$  та кореневу вершину  $v_{i_1}$ , та нуль в протилежному випадку (6). У випадку коли  $\text{mark}(v_s) = 1$  будемо казати, що вершина  $v_s \in V_{\text{num}}$  позначена, в протилежному випадку – не позначена.

$$\text{mark}: V_{\text{num}} \rightarrow \{0; 1\}, \quad \begin{cases} \text{mark}(v_s) = 1, & v_s \in P_{i_k i_1}; \\ \text{mark}(v_s) = 0, & v_s \notin P_{i_k i_1}. \end{cases} \quad (6)$$

Процес визначення відображення розмітки проводиться по всіх вершинах  $v_{i_k} \in Q$ , які ще не були позначені, та проводиться вгору до першої на шляху позначеної вершини. Всі інші вершини вважаються не позначеними. Тоді граф, який складається з всіх позначених вершин графу  $G$ , і буде шуканим підграфом  $H$ .

Описаний спосіб реалізується двома послідовно виконуваними алгоритмами: алгоритмом нумерації графу  $G$  (рис. 1) та алгоритмом розмітки графу  $G$  і знаходження  $H$ . Після цього, виходячи з  $H$ ,  $TN$  та ТЗв можна сформувані *table\_name* і *condition*. Алгоритм формування *table\_name* і *condition* буде містити кроки, які аналогічні тим, що використовуються

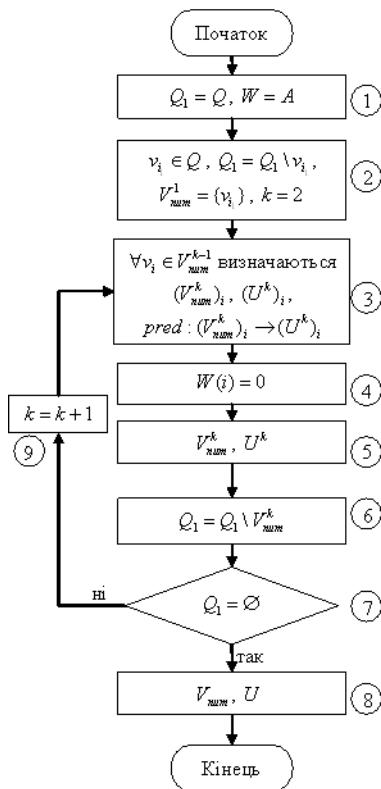


Рисунок 1 – Блок-схема алгоритму нумерації

в алгоритмі розмітки, тому доречним буде об'єднати ці алгоритм в один. Цей алгоритм буде мати назву алгоритм розмітки та формування *table\_name* і *condition* (рис. 2). Слід зазначити, що оскільки граф  $G$  відповідає всій схемі даних БД, то підграф  $H$  відповідає частині схеми даних БД. Це означає, що оскільки  $TN = V$ , а  $V' \subseteq V$ , то  $V' \subseteq TN$ . Позначимо  $TN' = V'$  множини імен таблиць БД, через які необхідно побудувати зв'язок в SQL-запиті, для його вірної реалізації. Множина ребер  $E'$  буде відповідати множині зв'язків між таблицями з  $TN'$ .

Додатково до введених позначень в цих алгоритмах будуть використовуватись наступні. Матриця  $W$ , яка перед початком роботи алгоритму дорівнює матриці суміжності  $A$  графу  $G$ ,  $W(i)$  буде означати  $i$ -ту строку матриці  $W$ , а  $w_{ij}$  – елемент матриці  $W$ , який знаходиться в  $i$ -му рядку та  $j$ -му стовпчику. Множина  $Q_1$ , яка буде складатись з тих вершин множини  $Q$ , які на даному кроці проведення перенумерації та розмітки не були помічені. Змінні  $temp$  та  $temp1$ , які будуть використовуватись для тимчасового зберігання номеру вершини. Множина вершин  $V_{num}^k$  з області визначення відображення  $pred$ , які були визначені на  $k$ -й ітерації, а множина припустимих значень –  $U^k$ . Тоді область визначення та множина припустимих

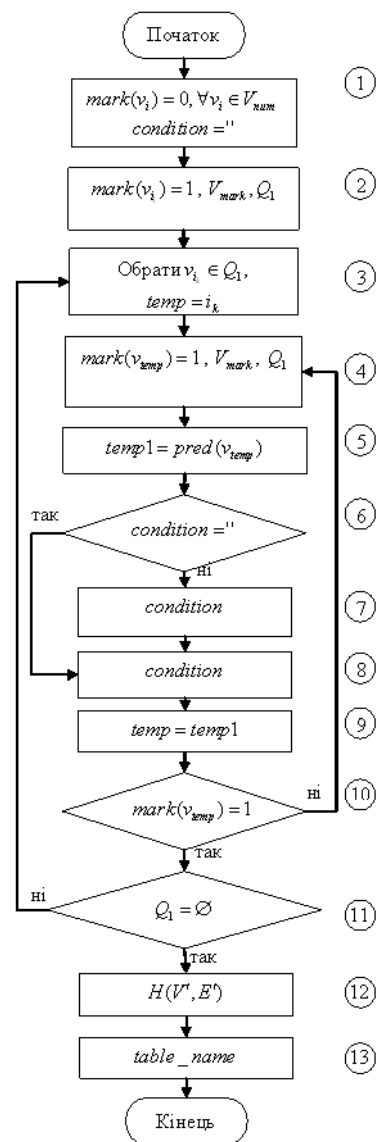


Рисунок 2 – Блок-схема алгоритму розмітки та формування *table\_name* і *condition*

значень відображення  $pred$  визначається через  $V_{num} = \bigcup_k V_{num}^k, U = \bigcup_k U^k$ .

Використавши отримані в (3) *column\_expression* і *column\_condition*, а також сформовані в результаті виконання алгоритму *table\_name* і *condition* можемо записати SQL-запит у вигляді (1).

### ВИСНОВКИ

Запропонований підхід до отримання інформації з бази даних та її представлення в термінах предметної області має наступні переваги:

1. При коректному формуванні таблиць прив'язки та зв'язків можна обійти недоліки, які були зроблені при проектуванні бази даних;

2. Даний підхід гарантує вірність побудови складних запитів та підвищує оперативність отримання необхідної інформації;

3. Виключає необхідність для користувача володіння спеціальними інформаційно-технічними знаннями та дозволяє зосередитись на вирішенні задач управління.

### ПЕРЕЛІК ПОСИЛАНЬ

1. Гольшев А. К. Сложные системы с развитой функцией информационно-аналитической поддержки управления. Элементы теории, методологии, практики / А. К. Гольшев. – К. : 2001. – 253 с.
2. Коннолли Т. Базы данных: проектирование, реализация и сопровождение. Теория и практика / Т. Коннолли, К. Бегг, А. Страчан. – М. : «Вильямс», 2000. – 1120 с.
3. Microsoft SQL Server 2008 [Електронний ресурс]. Електрон. текстові дані. – Режим доступу: [www.microsoft.com/SQL/default.aspx](http://www.microsoft.com/SQL/default.aspx). – Wednesday, 15 April, 2009.
4. Microsoft Office Access 2007 [Електронний ресурс]. Електрон. текстові дані. – Режим доступу: <http://office.microsoft.com/ru-ru/access/default.aspx>. – Wednesday, 15 April, 2009.
5. Праг К. Н. Microsoft Access 2000. Библия пользователя / К. Н. Праг, М. Р. Ирвин. – М. : Диалектика, 2001. – 1039 с.
6. 1С: Предприятие [Електронний ресурс]. Електрон. текстові дані. – Режим доступу: <http://www.1c.ru/>. – Wednesday, 15 April, 2009.
7. Microsoft SQL Server 2008: Reporting Services [Електронний ресурс]. Електрон. текстові дані. – Режим доступу: <http://www.microsoft.com/sqlserver/2008/en/us/reporting.aspx>. – Wednesday, 15 April, 2009.
8. SAP BusinessObjects. Crystal Reports [Електронний ресурс]. Електрон. текстові дані. – Режим доступу: <http://www.sap.com/solutions/sapbusinessobjects/sme/reporting/crystalreports/index.epx>. – Wednesday, 15 April, 2009.
9. The VSReport Designer. [Електронний ресурс]. Електрон. текстові дані. – Режим доступу: [http://help.globalscape.com/help/eft5/arm/the\\_vsreport\\_designer.htm](http://help.globalscape.com/help/eft5/arm/the_vsreport_designer.htm). – Wednesday, 15 April, 2009.
10. Flexible reporting to leverage your Business Intelligence capabilities [Електронний ресурс]. Електрон. текстові дані. – Режим доступу: <http://www.perpetuumsoft.com/Product.aspx?lang=en&pid=21>. – Wednesday, 15 April, 2009.
11. Касьянов В. Н. Графы в программировании: обработка, визуализация и применение / В. Н. Касьянов, В. А. Евстигнеев. – СПб. : БХВ-Петербург, 2003. – 1104 с.
12. Андерсон Дж. А. Дискретная математика и комбинаторика: Пер. с англ. / Дж. А. Андерсон – М. : Вильямс, 2003. – 960 с.

Надійшла 23.02.2009  
Після доробки 11.05.2009

*Предложен подход к получению информации из базы данных и ее представлению в терминах предметной области. В состав этого подхода входят метод привязки к предметной области и метод построения запроса к базе данных в терминах предметной области.*

*The approach to information obtaining from database and its representation in the terms of data domain was proposed. This approach is represented by the method of data domain binding and the method of query construction in data domain terms.*

УДК 681.326

А. С. Шкиль, Е. Е. Сыревич, Д. Е. Кучеренко, Самер Альмадхоун

## МЕТОД ОБРАТНОГО ПРОСЛЕЖИВАНИЯ ДЛЯ ПОИСКА ОШИБОК ПРОЕКТИРОВАНИЯ В HDL-КОДЕ

*В данной статье были рассмотрены методы поиска ошибок проектирования в неструктурированном HDL-коде. Разработан метод обратного прослеживания. Проведен эксперимент над HDL-моделью цифрового устройства с использованием данного метода.*

### 1 ПОСТАНОВКА ЗАДАЧИ

Верификация цифровых проектов, то есть аппаратных или встроенных аппаратно-программных систем, описанных на языке описания аппаратуры (Hardware Description Language – HDL), является важной задачей в процессе проектирования цифровых устройств. Часто более 70 % времени разработки затрачивается на поиск и исправление ошибок в проекте. Целью данной работы является разработка методов поиска дефектов/ошибок проектирования в

неструктурированном HDL-коде, позволяющих сократить время проведения диагностического эксперимента и уменьшить длину диагноза. Исходя из вышесказанного, необходимо решить задачу адаптации метода обратного прослеживания при верификации HDL-модели и провести диагностический эксперимент по поиску ошибок проектирования в рамках верификации [1].

### 2 АНАЛИЗ МЕТОДОВ ПОИСКА ДЕФЕКТОВ ПРИМЕНИТЕЛЬНО К МОДЕЛЯМ ЦИФРОВЫХ УСТРОЙСТВ

В данной работе аналогично понятию дефект используется понятие ошибки проектирования. Диагно-

© Шкиль А. С., Сыревич Е. Е., Кучеренко Д. Е., Самер Альмадхоун, 2009

стический эксперимент (ДЭ) над HDL-кодом осуществляется в два этапа [2]. На первом этапе проводится безусловный эксперимент. Если результат хотя бы на одном внешнем выходе не совпадает с эталоном, выполняется второй этап ДЭ. Проблема диагностирования технического состояния модели цифрового устройства заключается в минимизации области существования ошибок проектирования относительно наперед заданного класса, которая называется множеством эквивалентных дефектов, при допустимых параметрах длины теста и числа наблюдаемых выходов. Чем больше длина теста и количество наблюдаемых выходов, тем выше разрешающая способность алгоритма диагностирования, которая должна обеспечивать определение неисправного элемента только по фактическим реакциям без доступа к контрольным точкам (КТ). При фиксации параметров длины теста и числа наблюдаемых выходов, что имеет место в алгоритмах поиска дефектов, основанных на анализе таблиц неисправностей, можно повысить глубину диагностирования путем использования структуры связей входо-выходных элементов (операндов или операторов) модели цифрового устройства. Максимально адекватной реакцией модели на тест-вектор при наличии дефекта является вектор-строка значений наблюдаемых элементов модели. В качестве примера можно выделить метод тестового диагностирования ошибок проектирования в модели ЦУ на основе использования многозначных таблиц неисправностей (МТН) и анализе структуры объекта.

В начальную область подозреваемых ошибок проектирования входят все функциональные элементы (ФЭ) либо операндовые вершины. ДЭ проводится исходя из предположения, что в коде имеется одиночная ошибка проектирования, результирующая в неисправное значение на некоторой операндовой вершине, которая является либо контрольной точкой, либо внешним выходом. Основной принцип, лежащий в основе структурных методов поиска ошибок проектирования, следующий. Если в очередной контрольной точке результат элементарной проверки отрицателен, то в область подозреваемых ошибок на очередном шаге алгоритма входят сам ФЭ и все его предшественники. Если результат проверки положителен, то все предшественники предполагаются исправными, а подозреваемая ошибка произошла среди остальных ФЭ области подозреваемых ошибок пре-

дыдущего шага алгоритма. Если транспортирование ошибки проектирования на внешний выход невозможно, то граф необходимо разбить на подграфы по принципу существования активизации. Причиной невозможности активизации может быть либо так построенный код в силу ряда причин, либо в коде имеется ошибка (предполагаемое место ошибки – подграф его предшественников).

Стандартные проверяющие тесты неадекватны по нескольким причинам [3]. Основной из них является то, что в основе построения проверяющих тестов лежат понятия транспортирования и активизации пути, по которому неисправность транспортируется на внешний выход схемы. А это не всегда представляется возможным сделать, так как в HDL-коде функциональные неисправности могут быть замаскированы дальнейшими вычислениями таким образом, что ошибка не будет наблюдаться на внешнем выходе. Например, ошибка на выходе любого функционального элемента не транспортируется на внешний выход, если среди его преемников есть оператор логического ИЛИ с константной 1 на одном из входов. Для декомпозиции исходного графа используются контрольные точки (аналогично контрольным точкам при генерации тестов, которые позволяли «разбить» путь активизации и определить границы подграфов). Данные контрольные точки будут выходами каждого из подграфов. Таким образом, ошибка проектирования транспортируется на внешний выход каждого подграфа. Существуют два типа контрольных точек, используемых при поиске места ошибки. Контрольные точки первого рода – сигнал (переменная) модели, эталонные значения которых известны из спецификации. Контрольные точки второго рода – сигнал (переменная) модели, значения которых наблюдаемы, но до начала эксперимента неизвестны. Структуризация исходной модели на HDL выполняется путем выделения многовыходового подграфа, который порождается (активизируется) в результате подачи на модели ЦУ специализированного теста. Входами подграфа являются внешние входы (операндовые вершины, не только физические), а выходами – КТ первого рода. По результатам проведения первого этапа диагностического эксперимента (подачи теста для определения наличия ошибки проектирования в описании) формируется вектор результатов экспериментальной проверки, который фиксирует равенство эталонной и

```
Entity SCH is Port (X1,X2,X3,X4,X5,X6,X7, X8: in bit; O15: out bit); End;  
Architecture BEH of SCH is signal S9, S10, S11, s12, s13, S14: bit; begin  
S9<=X1 and x2; S10<=X3 and x4; S11<=X5 and x6; S12<=X7 and X8;  
S13<=S9 and S10; S14<= S11 and S12; O15<=S13 nand S14; End;
```

Рисунок 1 – Пример HDL-модели цифрового устройства

экспериментальной реакции. Другими словами, определяется, на какие из множества выходов оказывает влияние ошибка проектирования.

**3 МЕТОД ОБРАТНОГО ПРОСЛЕЖИВАНИЯ**

В спецификациях HDL-моделей реальных цифровых устройств КТ первого рода, в которых известны значения эталонных сигналов до начала написания кода, как правило, достаточно немного. Поэтому локализация до подграфа с выходной КТ первого рода является недостаточной. Желательно локализацию ошибочных операторов проводить до КТ второго рода, но эталонов нет и взять их негде. Выходом из этого противоречия является применение в подграфе, где выходом является КТ первого рода, метода обратного прослеживания, который первоначально был ориентирован поиск константных неисправностей в цифровых схемах в условиях отсутствия эталонов во внутренних точках. Классический метод обратного прослеживания, который схемотехники иногда называли методом «доискивания», основывался на следующих положениях: есть устройства с доступными внутренними линиями и его структурно-функциональная модель; есть условие существенности для каждого элемента; есть тест, который активизирует пути в модели; есть эталонные реакции на внешнем выходе модели [4].

Для использования этого метода при «доискивании» в классе эквивалентных ошибок при диагностике HDL-кода, необходима модификация метода. Принимаются следующие положения:

1. Класс неисправности – замена функционального элемента;
2. В качестве тестов используются различающая последовательность (РПС) для каждого ФЭ;
3. Неисправным (ошибочным) признается ФЭ, на входах которого наблюдалось последнее несовпадение с эталоном.

Алгоритм выполнения «доискивания»:

1. Выполняется моделирование тестов на реальной модели до внешнего выхода (КТ);
2. Если на выходе обнаружено несовпадение с эталоном, то выполняется обратное прослеживание;
3. Обратное прослеживание идет до тех пор, пока результат не равен ∅, путем пересечения текущего вектора с условиями существенности функционального элемента;
4. Из непустых пересечений формируется подмножество подозреваемых элементов.

Пусть имеется некоторый ошибочный HDL-код (рис. 1), реализующий схему, представленную на рис. 2. Ошибка в коде: в операторе S12 функция **or** заменена на **and**.

Для этого фрагмента кода построен тест 11011100 и эталонное значение 0 на внешнем выходе (сигнал O15). В результате подачи теста на HDL-модель на ее выходе получена 1, что не совпадает с эталоном. Выполним обратное прослеживание, отмечая его результаты для наглядности на фрагменте цифровой схемы (рис. 2), эквивалентной приведенному коду.

Наиболее эффективной формой для реализации обратной импликации является кубическая форма представления моделей примитивов, поэтому для наглядности рассмотрим применение метода «доискивания» в кубической форме. Если рассматривать условие существенности в кубической форме, то они подобны D-покрытиям, но не учитывают направление перехода (инверсию). Покрытия существенности для элементов (операторов) AND и OR приведены на рис. 3.

Процедура выполнения алгоритма обратного прослеживания для фрагмента HDL-кода с predetermined операторами параллельного назначения сигналов (ПНС) подобна процедуре обратной импликации для комбинационной схемы, при условии, что в качестве примитивных элементов выступают покрытия существенности соответствующих ПНС. Первоначально выполняется моделирование теста на реальном (неисправном) коде и результат моделирования записывается в первую строку [5]. Результат из спецификации не совпадает с эталоном, поэтому выполняется обратное прослеживание путем пересечения

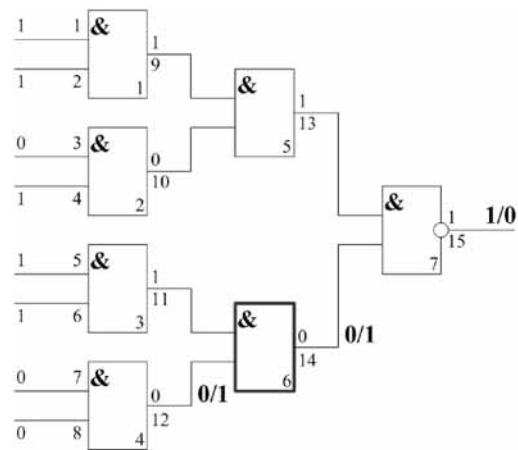


Рисунок 2 – «Доискивание» на примере структурно-функциональной схемы

X1	X2	X3	Y	X1	X2	X3	Y
Z	1	1	Z	Z	0	0	Z
1	Z	1	Z	0	Z	0	Z
1	1	Z	Z	0	0	Z	Z

Рисунок 3 – Покрытия существенности элементов AND и OR



$$C = X1 \cdot X2 \vee X3 \cdot X4$$

X1	X2	X3	X4	C
0	X	0	X	0
0	X	X	0	0
X	0	0	X	0
X	0	X	0	0
1	1	X	X	1
X	X	1	1	1

Рисунок 4 – Решение равенства по логическому условию

X1	X2	X3	X4	D1	D2	Y
0	X	0	X	X	Z	Z
0	X	X	0	X	Z	Z
X	0	0	X	X	Z	Z
X	0	X	0	X	Z	Z
1	1	X	X	Z	X	Z
X	X	1	1	Z	X	Z

Рисунок 5 – Покрытие существенности оператора IF()

полученного вектора с покрытиями существенности элементов модели. Пересечения выполняются до тех пор, пока результат не будет равен  $\emptyset$ . Процедура выполнения обратного прослеживания приведена в табл. 1.

Таким образом, неисправным является элемент, у которого входом является линия 12, т. е. элемент 6, который был ИЛИ, а стал И, что и требовалось доказать. Каждый специализированный тест порождает многовыходной подграф, входами которого являются внешние входы (либо КТ первого рода), а выходами внешними выходами (либо КТ первого рода). Кроме операторов ПНС в HDL-коде присутствуют последовательные операторы. Особый интерес при выполнении обратного прослеживания представляют условные операторы (*IF*, *CASE*), т. к. они находятся на стыке данных и управления. Рассмотрим оператор

*IF*: *IF* ((*x1* and *x2*) or (*x3* and *x4*)) *THEN* *y* <= *d1* *ELSE* *y* <= *d2*. Оператор *IF*( ), реализующий двухвыходовую условную вершину, состоит из логического условия *C* и двух выполняемых операторов ПНС или присваивания значений переменных. Логическое условие *C* в операторе *IF* (*C*) – это уравнение, решение в кубической форме представлено на рис. 4.

Таким образом, оператор *IF*( ) может быть представлен моделью мультиплексора с двумя информационными входами, группой управляющих входов и одним (двумя) выходами (два выхода будет в случае, если назначение выполняется для разных сигналов). Покрытие существенности оператора *IF*( ) представлено на рис. 5.

Кроме оператора *IF*( ) к условным операторам можно отнести и оператор *CASE*( ), который реализует многовыходовую условную вершину. Фрагмент HDL-кода с оператором *CASE*( ) представлен на рис. 6.

Модель оператора *CASE*( ) в форме мультиплексора и его покрытие существенности представлено на рис. 7.

К операторам ПНС относятся также арифметические операторы, выполняемые, как правило, над многоразрядными операторами. Если рассмотреть операцию сложения одноразрядных операндов  $S = A + B$ , то схемной реализацией ее является одноразрядный двухходовой сумматор. Его закон функционирования и фрагмент покрытия существенности приведен на рис. 8. Следует отметить, что покрытие существенности строится динамически на основании закона функционирования сумматора. Имея покрытия существенности для одноразрядного сумматора с учетом процедур обратной импликации через арифметические операнды, легко строятся процедуры обратного продвижения существенности через многоразрядные арифметические операции. Отметим, что для арифметических операций корректными могут быть только одномерные по входам покрытия существенности.

Таблица 1 – Процедура обратного прослеживания

1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	
1	1	0	1	1	1	0	0	1	0	1	0	1	0	1	результат из спецификации 0 линия 15 сущ. z $\cap \neq \emptyset$ $\rightarrow \emptyset$
												1	z	z	
												z	1	z	
1	1	0	1	1	1	0	0	1	0	1	0	1	z	z	линия 14 сущ. $\rightarrow \emptyset$ $\cap \neq \emptyset$
												z	1	z	
												1	z	z	
1	1	0	1	1	1	0	0	1	0	1	z	1	z	z	линия 12 сущ. $\rightarrow \emptyset$ $\rightarrow \emptyset$

```
variable xxx : in std_logic_vector(2 downto 0);
case xxx is
  when «000»=> Y<=D1;
  when «010»=> Y<=D3;
  when «100»=> Y<=D5;
  when «110»=> Y<=D7;
  when «001»=> Y<=D2;
  when «011»=> Y<=D4;
  when «101»=> Y<=D6;
  when «111»=> Y<=D8;
end case;
```

Рисунок 6 – Фрагмент HDL-кода с оператором CASE( )

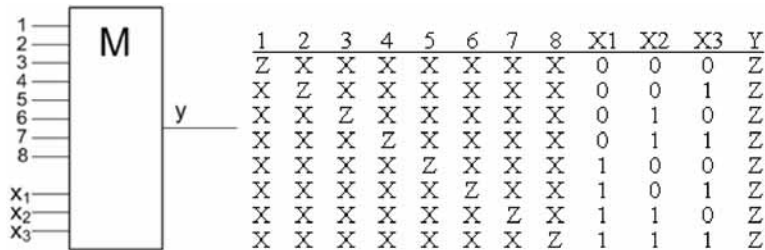


Рисунок 7 – Модель CASE() в форме мультиплексора и его покрытие существенности

A <sub>i</sub>	B <sub>i</sub>	C <sub>i-1</sub>	S	C <sub>i</sub>	A <sub>i</sub>	B <sub>i</sub>	C <sub>i-1</sub>	S	C <sub>i</sub>
0	0	0	0	0	Z	0	0	Z	0
0	1	0	1	0	0	Z	0	Z	0
1	0	0	1	0	0	0	Z	Z	0
1	1	0	0	1	Z	1	1	Z	1
0	0	1	1	0	1	Z	1	Z	1
0	1	1	0	1	1	1	Z	Z	1
1	0	1	0	1	Z	0	1	Z	Z
1	1	1	1	1	0	Z	1	Z	Z

Рисунок 8 – Закон функционирования и фрагмент покрытия существенности одноразрядного сумматора

В качестве примера рассмотрим операцию сложения двух четырехразрядных операторов  $S(4) = A(4) + B(4)$  на примере  $A = 5, B = 6, S = 5 + 6 = 11$ . В двоичном виде  $S = A + B = 0101 + 0110 = 1011$ . Если  $S = 10Z1$  (символ существенности находится в третьем выходном разряде), то существует только одно корректное решение при  $A = 0101$  и  $B = 01Z0$ . В качестве проверки рассмотрим  $S = A + B = 0101 + 01Z0 = 10Z1$ , при условии, что  $Z = \{0, 1\}$ . Приведенный пример подтверждает возможность выполнения обратной импликации символа  $Z$  через многоразрядные арифметические операторы. Таким образом, имея покрытия существенности для операторов ПНС и условных операторов можно выполнять обратное прослеживание для любых фрагментов HDL-кода.

**ЗАКЛЮЧЕНИЕ**

Методы поиска дефектов (ошибок проектирования) для верификации HDL-моделей позволяют не

только говорить о наличии ошибки проектирования, но и точно определить место ее возникновения (локализовать дефект). В работе были усовершенствован метод обратного прослеживания, который применим для «доискивания» в подграфах, не имеющих внутри себя контрольных точек первого рода.

**ПЕРЕЧЕНЬ ССЫЛОК**

1. Сыревич Е. Е. Верификация моделей цифровых устройств, представленных на языках описания аппаратуры: дисс. канд. техн. наук: 05.13.12 / Сыревич Евгения Ефимовна. – Харьков, 2007. – 176 с. – Библиогр.: с. 82–130.
2. Структурный метод поиска ошибок проектирования в моделях радиоэлектронных устройств на HDL: тез. докл. Третьего Международного радиоэлектронного форума «Прикладная радиоэлектроника. Состояние и перспективы развития» (октябрь, 2008) / отв. ред. Бондаренко. – Харьков : 2008. – 48–51 с.
3. Шкиль А. С. Реализация процедур импликации на графовой структуре / Шкиль А. С., Чегликов Д. И., Зинченко Д. Е. // Журн. радиоэлектронные компьютерные системы. – 2006. – Вып. 6. – С. 172–177.
4. Abramovici M. Digital System Testing and Testable Design / Abramovici M., Breuer M., Friedman A. – New York : Computer Science Press, 1998. – 652 p.
5. Test Generation for VHDL Descriptions Verification: Proceedings of IEEE East-West Design & Test Workshop (September, 2005). – Odessa : 2008. – Pp. 191–195 p.

Надійшла 16.03.2009  
Після доробки 18.05.2009

*У даній статті були розглянуті методи пошуку помилок проектування в неструктурованому HDL-кодi. Був розроблений метод зворотнього простежування. Був проведений експеримент над HDL-моделлю цифрового пристрою з використанням даного методу.*

*In this paper the method of design error searching in unstructured HDL-code was considered. The backtracing method was developed. An experiment on HDL-model of digital device that uses this method was carried out.*

# НЕЙРОІНФОРМАТИКА ТА ІНТЕЛЕКТУАЛЬНІ СИСТЕМИ

## НЕЙРОИНФОРМАТИКА И ИНТЕЛЕКТУАЛЬНЫЕ СИСТЕМЫ

### NEUROINFORMATICS AND INTELLIGENT SYSTEMS

---

---

УДК 004.023

Е. В. Бодянский, В. В. Волкова, К. В. Коваль

#### **АВТОМАТИЧЕСКАЯ КЛАСТЕРИЗАЦИЯ ТЕКСТОВЫХ ДОКУМЕНТОВ НА ОСНОВЕ ГЕНЕТИЧЕСКОГО АЛГОРИТМА С ИСКУССТВЕННЫМ ОТБОРОМ**

---

*Предложен новый генетический алгоритм с искусственным отбором, в основе которого лежит синтез обычного эволюционного генетического подхода с идеями последовательного комплекс-метода отыскания экстремума произвольных функций многих переменных. Алгоритм используется для кластеризации больших объемов текстовых документов в режиме последовательной обработки.*

#### **ВВЕДЕНИЕ**

В общей проблеме интеллектуального анализа данных – Data Mining, Exploratory Data Analysis, Text Mining и, особенно, Web Mining важное место занимает задача поиска и классификации информации, содержащейся в текстовых документах, количество которых в Web практически неограниченно и постоянно увеличивается, достигая астрономических значений [1]. Фактически речь идет об очень больших и непрерывно растущих во времени базах данных и знаний, образованных зачастую не связанными между собой текстами самого различного содержания и происхождения, поиск и анализ в которых должен производиться в on-line режиме, обеспечивая эффек-

тивное автоматическое разбиение обрабатываемых документов на обычно заранее неизвестное число классов [2].

К настоящему времени сложилось достаточно много подходов к решению этой задачи, однако, большинство из них связано с использованием человеческого интеллекта и труда, которые весьма дороги. В связи с этим сегодня упор делается на средства искусственного интеллекта, позволяющие решать задачу в автоматическом режиме при минимальном участии человека. Среди таких средств достаточно высокую эффективность продемонстрировали как ставшие уже традиционными методы кластеризации, так и более современные методы вычислительного интеллекта такие, как искусственные нейронные сети (самоорганизующиеся карты Кохонена, BSB-нейромодели, сети теории адаптивного резонанса и т. д.) и нечеткие системы (fuzzy C-means Бездека, алгоритмы Густафсона – Кесселя, Ягера – Филева, Хёппнера – Клавонна – Крузе и т. д.).

Применение математических методов кластеризации требует предварительного преобразования анали-

зируемых документов в приемлемые для дальнейшей обработки компактные формы такие, как концепты, число которых также может быть велико [1], либо векторы, образованные частотами появления отдельных термов в тексте, размерность которых может быть неприемлемой для эффективного использования как нейронных сетей, так и четких, и нечетких методов кластеризации. Так или иначе, для решения задачи классификации документ должен быть сначала преобразован в векторную форму с числовыми компонентами.

Дальнейшее решение связано с проблемой оптимизации (критерий самообучения нейронной сети, целевая функция кластеризации) функции многих аргументов, число которых очень велико, а сама эта функция, описывающая либо корреляцию между концептами [3–4], либо меру семантического подобия текстов [3], как правило, многоэкстремальна [2] и может иметь сколь угодно сложную форму [1].

Именно это обстоятельство привело в последние годы к использованию в проблеме кластеризации текстов методов третьего основного направления вычислительного интеллекта – эволюционных вычислений и, прежде всего, генетических алгоритмов [1–7] и дало начало направлению, получившему название Genetic Mining [4]. Генетические алгоритмы в общем случае представляют собой методы эвристической оптимизации, чьи механизмы подобны биологической эволюции на основах принципов естественного отбора [8–12]. Такие их свойства, как адаптивность, робастность, возможность распараллеливания вычислений и отыскание глобального экстремума принятой функции приспособленности (fitness function), обеспечили их эффективное использование для решения задач кластеризации в пространствах высокой размерности [3]. При этом каждый документ, подлежащий обработке, кодируется в форме хромосомы (string) с бинарными компонентами, множество документов образует популяцию с изменяющимся числом особей, а в качестве функции приспособленности используется, как правило, та или иная мера семантического подобия анализируемых текстов.

Вместе с тем, холландовские алгоритмы, основанные на идеях естественного отбора, характеризуются низкой скоростью сходимости, не позволяющей им отыскивать решение за приемлемое время. В связи с этим представляется целесообразным ввести в процедуры генетической оптимизации элементы искусственного отбора, отличного от общепринятых стратегий элитизма (пчелиная семья, модель островов и т. д.) и имеющие под собой более строгое математическое обоснование.

### ГЕНЕТИЧЕСКИЙ АЛГОРИТМ С ИСКУССТВЕННЫМ ОТБОРОМ

В основе предлагаемого алгоритма лежит синтез обычного эволюционного генетического подхода с идеями адаптивной оптимизации [13] и, прежде всего, последовательного комплекс-метода отыскания экстремума функций многих переменных [13–15]. При этом в каждый момент времени текущая популяция отождествляется с «облаком» – комплексом точек в пространстве переменных-факторов, а кроме традиционных генетических операторов мутации, кроссовера и инверсии дополнительно вводятся операторы комплекс-поиска такие, как отражение, растяжение и сжатие. При этом в отличие от традиционного комплекс-метода нами предлагается отражать не одну наихудшую вершину комплекса, а целое множество наихудших особей популяции.

В общем случае процедура оптимизации на основе обычного последовательного комплекс-метода выглядит следующим образом: требуется отыскать минимум некоторой функции

$$E(x) \rightarrow \min_{x \in R^n}$$

достаточно общего вида, при этом о характере этой функции не делается практически никаких априорных предположений. Работа алгоритма начинается с формирования начального комплекса

$$x_i(0) = (x_{i1}(0), x_{i2}(0), \dots, x_{ij}(0), \dots, x_{in}(0))^T, \\ i = 1, 2, \dots, N \geq n + 1,$$

представляющего собой «облако» (популяцию) точек (векторов), достаточно произвольно расположенных в  $n$ -мерном пространстве факторов. Среди множества этих точек находится «наихудшая»  $x_i(1)$ , в которой значение функции  $E(x_H(0))$  максимально, после чего эта точка отражается через центр тяжести всех остальных вершин-точек, формируя новый комплекс  $x_i(1)$ ,  $i = 1, 2, \dots, N$ . Такое отражение вместе с растяжением и сжатием обеспечивают движение комплекса к экстремуму функции  $E(x)$ , при этом, благодаря достаточно случайному распределению точек «облака», поиск имеет глобальный характер.

С формальной точки зрения рассмотрим процесс оптимизации на  $k$ -й итерации поиска, когда сформирован комплекс  $x_i(k)$ ,  $i = 1, 2, \dots, N$ . Среди множества точек  $x_i(k)$  находится «наихудшая» такая, что

$$E(x_H(k)) = \max_i \{E(x_1(k)), \dots, E(x_N(k))\},$$

после чего определяется центр тяжести «облака» без наихудшей точки:

$$x_C(k) = \frac{1}{N-1} \left( \sum_{i=1}^N x_i(k) - x_H(k) \right).$$

Далее  $x_H(k)$  отражается через центр тяжести  $x_C(k)$ , формируя новую вершину комплекса  $x_R(k)$ , которая теоретически расположена ближе к экстремуму чем  $x_H(k)$  и  $x_C(k)$ , т. е.

$$E(x_R(k)) < E(x_C(k)) < E(x_H(k)).$$

Операция отражения формально имеет следующий вид:

$$\begin{aligned} x_R(k) &= x_C(k) + \eta_R(x_C(k) - x_H(k)) = \\ &= \frac{1}{N-1}x_1(k) + \dots + \frac{1}{N-1}x_{N-1}(k) + \frac{\eta_R}{N-1}x_1(k) + \dots + \\ &+ \frac{\eta_R}{N-1}x_{N-1}(k) - \eta_R x_H(k) = X(k)R, \end{aligned}$$

где  $\eta_R$  – параметр шага отражения, часто полагаемый равным единице,  $X(k) = (x_H(k), x_1(k), \dots, x_{N-1}(k)) - (n \times N)$ -матрица координат вершин комплекса,  $R = \left( -\eta_R, \frac{1+\eta_R}{N-1}, \dots, \frac{1+\eta_R}{N-1} \right)^T - (N \times 1)$ -вектор.

В случае, если отраженная вершина  $x_R(k)$  окажется «наилучшей» среди всех остальных точек комплекса, т. е.

$$E(x_R(k)) < E(x_i(k)) < E(x_H(k)), \quad i = 1, 2, \dots, N-1,$$

производится операция растяжения комплекса в направлении от центра тяжести  $x_C(k)$  до  $x_R(k)$  согласно выражению

$$x_E(k) = x_C(k) + \eta_E(x_R(k) - x_C(k)) = X(k)E,$$

где  $\eta_E$  – параметр шага растяжения, часто полагаемый равным двум,

$$E = \left( -\eta_E \eta_R, \frac{1 - \eta_E(1 - \eta_R)}{N-1}, \dots, \frac{1 - \eta_E(1 - \eta_R)}{N-1} \right)^T.$$

Если же  $x_R(k)$  окажется наихудшей среди всех  $x_i(k)$ , комплекс сжимается согласно соотношению

$$x_S(k) = x_C(k) + \eta_S(x_R(k) - x_C(k)) = X(k)S,$$

где  $\eta_S$  – параметр шага сжатия, обычно полагаемый равным 0,5,

$$S = \left( -\eta_S \eta_R, \frac{1 - \eta_S(1 - \eta_R)}{N-1}, \dots, \frac{1 - \eta_S(1 - \eta_R)}{N-1} \right)^T.$$

При  $\eta_S = 1$ ,  $\eta_E = 2$ ,  $\eta_S = 0,5$  приходим к простым выражениям

$$\begin{aligned} R &= \left( -1, \frac{2}{N-1}, \dots, \frac{2}{N-1} \right)^T, \\ E &= \left( -2, \frac{1}{N-1}, \dots, \frac{1}{N-1} \right)^T, \\ S &= \left( -0,5, \frac{1}{N-1}, \dots, \frac{1}{N-1} \right)^T. \end{aligned}$$

Таким образом, в процессе своего движения к экстремуму оптимизируемой функции комплекс на каждой итерации теряет одну наихудшую вершину и приобретает одну новую точку так, что на  $(k+1)$ -й итерации новый комплекс также имеет  $N$  точек-вершин.

В отличие от комплекс-метода, в генетических алгоритмах в результате селекции из популяции одновременно исключаются несколько особей с наихудшими (максимальными) значениями функции приспособленности. В связи с этим представляется целесообразным ввести алгоритм комплекс-метода с отражением, растяжением и сжатием сразу нескольких вершин.

Итак, пусть на  $k$ -й итерации процесса оптимизации имеется комплекс  $x_i(k)$ ,  $i = 1, 2, \dots, N$  с  $P < N$  наихудшими вершинами  $x_{H_p}(k)$ ,  $p = 1, 2, \dots, P$ . Тогда координаты центра тяжести комплекса без вершин  $x_{H_p}(k)$  задаются выражением

$$x_C(k) = \frac{1}{N-P} \left( \sum_{i=1}^N x_i(k) - \sum_{p=1}^P x_{H_p}(k) \right),$$

а процедура отражения описывается системой уравнений

$$\begin{cases} x_{R_1}(k) = x_C(k) + \eta_R(x_C(k) - x_{H_1}(k)), \\ \vdots \\ x_{R_P}(k) = x_C(k) + \eta_R(x_C(k) - x_{H_P}(k)), \end{cases}$$

или

$$\begin{cases} x_{R_1}(k) = (1 + \eta_R)x_C(k) - \eta_R x_{H_1}(k), \\ \vdots \\ x_{R_P}(k) = (1 + \eta_R)x_C(k) - \eta_R x_{H_P}(k). \end{cases}$$

В матричной форме эти системы уравнений могут быть записаны более компактно

$$X_R(k) = X(k)R_P,$$

где  $X(k) = \underbrace{(x_{H_1}(k), \dots, x_{H_P}(k))}_{(n \times P)}, \underbrace{(x_1(k), \dots, x_{N-P}(k))}_{(n \times (N-P))} - (n \times N)$ -матрица,  $X_R(k) = (x_{R_1}(k), \dots, x_{R_P}(k)) - (n \times P)$ -

матрица,  $R_P = \begin{matrix} P \\ N-P \end{matrix} \left\{ \begin{matrix} -\eta_R I_P \\ \hline \frac{1+\eta_R}{N-P} I_{N-P,P} \end{matrix} \right\}$  –  $(N \times P)$ -мат-

рица,  $I_p$  –  $(P \times p)$ -единичная матрица,  $I_{N-P,P}$  –  $((N-P) \times P)$ -матрица, образованная единицами. В случае, если среди отраженных вершин оказывается  $Q \leq P$  наилучших, комплекс растягивается в их направлении согласно уравнениям

$$\begin{cases} x_{E_1}(k) = x_C(k) + \eta_E(x_{R_1}(k) - x_C(k)), \\ \vdots \\ x_{E_Q}(k) = x_C(k) + \eta_E(x_{R_Q}(k) - x_C(k)), \end{cases}$$

или

$$\begin{cases} x_{E_1}(k) = (1 - \eta_E)x_C(k) + \eta_E x_{R_1}(k), \\ \vdots \\ x_{E_Q}(k) = (1 - \eta_E)x_C(k) + \eta_E x_{R_Q}(k), \end{cases}$$

или

$$X_E(k) = X(k)E_Q,$$

где  $X_E(k) = (x_{E_1}(k), \dots, x_{E_Q}(k))$  –  $(n \times Q)$ -матрица,  $E_Q =$

$$= \begin{matrix} Q \\ N-Q \end{matrix} \left\{ \begin{matrix} -\eta_E \eta_R I_Q \\ \hline \left(1 - \frac{\eta_E(1-\eta_R)}{N-P}\right) I_{N-Q,Q} \end{matrix} \right\} \text{ – } (N \times Q)\text{-матрица.}$$

Если, далее, среди отражаемых вершин окажется  $U \leq P$  наихудших, комплекс сжимается в их направлении согласно уравнениям

$$\begin{cases} x_{S_1}(k) = x_C(k) + \eta_S(x_{R_1}(k) - x_C(k)), \\ \vdots \\ x_{S_U}(k) = x_C(k) + \eta_S(x_{R_U}(k) - x_C(k)), \end{cases}$$

или

$$\begin{cases} x_{S_1}(k) = (1 + \eta_S)x_C(k) - \eta_S x_{R_1}(k), \\ \vdots \\ x_{S_U}(k) = (1 + \eta_S)x_C(k) - \eta_S x_{R_U}(k), \end{cases}$$

или

$$X_S(k) = X(k)S_U,$$

где  $X_S(k) = (x_{S_1}(k), \dots, x_{S_U}(k))$  –  $(n \times U)$ -матрица,  $S_U =$

$$= \begin{matrix} U \\ N-U \end{matrix} \left\{ \begin{matrix} -\eta_S \eta_R I_U \\ \hline \left(1 - \frac{\eta_S(1-\eta_R)}{N-U}\right) I_{N-U,U} \end{matrix} \right\} \text{ – } (N \times U)\text{-матрица.}$$

Таким образом, комплекс-метод приобретает черты генетического алгоритма, у которого в результате

селекции на каждой итерации из популяции удаляет- ся несколько наихудших особей.

Объединяя введенную модификацию комплекс-м- ета с холландовской генетической процедурой, при- ходим к алгоритму, реализующему идею искусствен- ного отбора, состоящую в данном случае в том, что из популяции не только удаляются наихудшие особи, но и одновременно создаются их «антиподы», обла- дающие улучшенными свойствами.

Работа такого алгоритма образована последова- тельностью следующих шагов:

- создание начальной популяции, образованной  $P(0)$  особями хромосомами – вершинами комплекса;
- операция кроссовера с увеличением популяции  $P_{CR}(0) > P(0)$ ;
- операция мутации  $P_M(0) > P_{CR}(0)$ ;
- операция инверсии  $P_I(0) > P_M(0)$ ;
- первая селекция (определение наихудших осо- бей) без сокращения популяции  $P_{SEL1}(0) = P_I(0)$ ;
- операция отражения с удалением  $P$  наихудших особей  $P_R(0) < P_{SEL1}(0)$ ;
- операция растяжения без увеличения популяции  $P_E(0) = P_R(0)$ ;
- операция сжатия без увеличения популяции  $P_I(0) = P_E(0)$ ;
- вторая селекция с удалением  $P_W(0)$  наихудших особей  $P_{SEL2}(0) = P_I(0) - P_W(0) = P(1)$  и формиро- вание популяции  $P(1)$  для следующей итерации алго- ритма. Предлагаемый алгоритм может быть пред- ставлен схемой, приведенной на рис. 1.

### РЕЗУЛЬТАТЫ ЭКСПЕРИМЕНТАЛЬНЫХ ИССЛЕДОВАНИЙ

Целью исследования было оценить качество кла- стеризации стандартными генетическими алгоритма- ми и разработанным в ходе исследования генети- ческим алгоритмом с искусственным отбором. Рабо- та предложенного алгоритма оценивалась на выборке новостных статей Reuters-21578, которая является одной из наиболее часто используемых тестовых вы- борок в Text Mining и информационном поиске. Для эксперимента было выбрано 200 документов из топик- ков coffee, crude, sugar и trade выборки Reuters- 21578.

Эксперимент показал, что предложенный генети- ческий алгоритм с искусственным отбором работает быстрее и дает более точные результаты (в среднем 6–8 %) по сравнению со стандартными генетическими алгоритмами и может быть использован для работы с большими массивами текстовых документов.

Таким образом, было установлено, что в задаче кластеризации документов предложенный генетиче- ский алгоритм с искусственным отбором дает более точные результаты, чем стандартные генетические ал- горитмы.

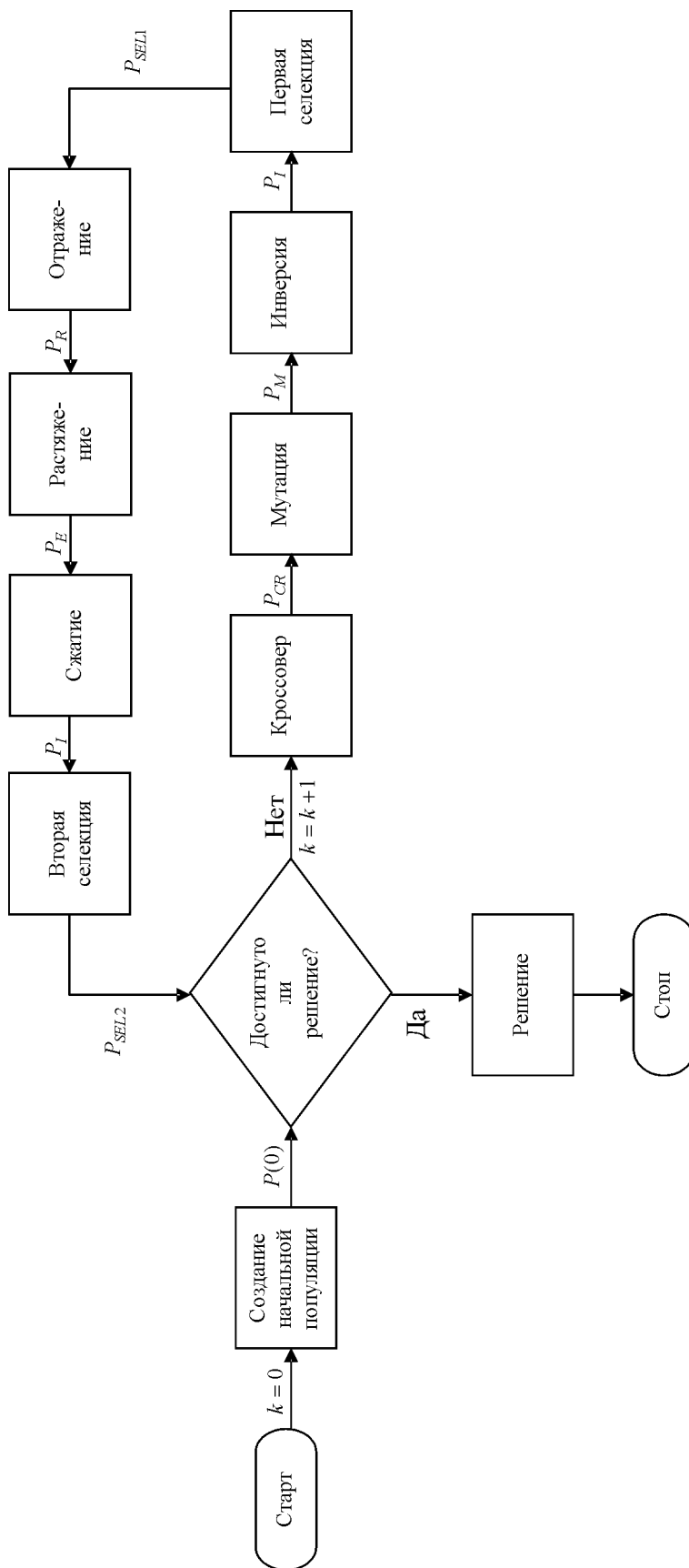


Рисунок 1 – Генетический алгоритм с искусственным отбором на основе последовательного комплекс-метода

**ВИВОДИ**

В роботі пропонується новий генетичний алгоритм з штучним відбором на основі комплексного методу адаптаційної оптимізації, призначений для пошуку екстремуму произвольних функцій великого числа аргументів в умовах суттєвої неопределенності о характері цих функцій. Алгоритм має покращені характеристики порівняно з традиційними генетичними процедурами, прост в реалізації і призначений для використання в Genetic Mining великих масивів текстових документів в режимі послідовної обробки.

**ПЕРЕЧЕНЬ ССЫЛОК**

1. Desjardins G. A genetic algorithm for text mining / G. Desjardins, G. R. Godin, R. Proulx // Sixth International Conference on Data Mining, Text Mining and their Business Applications. – 2005. – P. 133–142.
2. Zhang C. Self-adaptive GA, quantitative semantic similarity measures and ontology-based text clustering / C. Zhang, W. Song, C. Li, W. Yu // 2008. – <http://eprints.rclis.org/14909/> (15.12.2008).
3. Othman R. M. Incorporating semantic similarity measure in genetic algorithm: an approach for searching the gene ontology terms / R. M. Othman, S. Deris, R. M. Illias, H. T. Alashwal, R. Hassan, F. Mohamed // International Journal of Computational Intelligence. – 2006. – № 3. – P. 257–266.
4. Khalessizadeh S. M. Genetic mining: using genetic algorithm for topic based on concept distribution / S. M. Khalessizadeh, R. Zaefarian, S. H. Nasseri, E. Ardil // Proceedings of World Academy of Science, Engineering and Technology – 2006. – 13. – P. 144–147
5. Mani I. Advances in Automatic Text Summarization / I. Mani, M. T. Maybury // Cambridge : MIT Press, 1999. – 442 p.
6. Othman R. M. Automatic clustering of gene ontology by genetic algorithm / R. M. Othman, S. Deris, R. M. Illias, Z. Zakaria, S. M. Mohamad // International Journal of Information Technology. – 2006 – 3.–№ 1. – P.37–46.

7. Rocha F. E. L. A new approach to meaningful learning assessment using concept maps: ontologies and genetic algorithms / F. E. L. Rocha, J. V. da Costa, E. L. Favero // 2004. – <http://cmc.ihmc.us/papers/cmc2004-238.pdf> (15.12.2008)
8. Holland J. H. Genetic algorithms and the optimal allocations of trails // SIAM Journal of Computing. – 1973. – 2. – P. 88–105.
9. Holland J. H. Adaptation in Natural and Artificial Systems. An Introductory Analysis with Application to Biology, Control and Artificial Intelligence. – London : Bradford Book Edition, 1994. – 211 p.
10. Батищев Д. И. Генетические алгоритмы решения экстремальных задач. – Воронеж : Воронежский государственный технический университет, 1995 – 69 с.
11. Курейчик В. М. Генетические алгоритмы. Состояние. Проблемы. Перспективы // Известия РАН. Теория и системы управления. – 1999. – 1. – С. 144–160.
12. Рутковская Д. Нейронные сети, генетические алгоритмы и нечеткие системы / Рутковская Д., Пилиньский М., Рутковский Л. – Москва : Горячая линия – Телеком, 2004. – 452 с.
13. Горский В. Г. Планирование промышленных экспериментов / Горский В. Г., Адлер Ю. П. – Москва : Металлургия, 1974. – 264 с.
14. Химмельблау Д. М. Прикладное нелинейное программирование / Химмельблау Д. М. – Москва : Мир, 1975. – 534 с.
15. Реклейтис Г. Оптимизация в технике : кн. 1 / Реклейтис Г., Рейвиндран А., Рэгсдел К. – Москва : Мир, 1986. – 349 с.

Надійшла 16.03.2009

*Запропоновано новий генетичний алгоритм зі штучним відбором, в основі якого лежить синтез звичайного еволюційного генетичного підходу з ідеями послідовного комплекс-методу пошуку екстремуму довільних функцій багатьох змінних. Алгоритм використовується для кластеризації великих обсягів текстових документів у режимі послідовної обробки.*

*The new genetic algorithm with artificial selection is proposed. The algorithm is based on the synthesis of ordinary evolutionary genetic approach with the ideas of sequential complex-method for extremum searching arbitrary multivariable functions. The algorithm is used for a clusterization of large data collection in a data-processing mode.*

УДК 519.7:004.93

О. О. Олійник, С. О. Субботін

## ОПТИМІЗАЦІЯ НА ОСНОВІ КОЛЕКТИВНОГО ІНТЕЛЕКТУ РОЮ ЧАСТОК З КЕРУВАННЯМ ЗМІНОЮ ЇХНЬОЇ ШВИДКОСТІ

*Досліджено метод оптимізації на основі моделювання поведінки рою часток. Розроблено модифікацію дослідженого методу з керуванням зміною швидкості часток. Проведено експерименти зі знаходження глобального оптимуму багатовимірної функції на основі запропонованої модифікації.*

**ВСТУП**

Градентні методи безумовної оптимізації, що традиційно застосовуються при синтезі моделей складних об'єктів і систем, є високоітеративними та накладають певні вимоги (наприклад, унімодальність, безперервність, монотонність, диференційованість та інші.)

© Олійник О. О., Субботін С. О., 2009



до вигляду цільової функції, а також, як правило, є методами локального пошуку.

Тому актуальним є завдання розроблення нових методів оптимізації, що не накладають подібних вимог на цільові функції. У зв'язку з цим виникає інтерес до методів, заснованих на випадковому пошуку, які характеризуються більшою швидкістю роботи, а також не накладають ніяких вимог на залежності, що досліджуються. До таких методів належать мультиагентні методи інтелектуальної оптимізації, що мають біонічну природу і засновані на моделюванні поведінки колективного інтелекту соціальних живих істот (Swarm Intelligence) [1, 2].

До інтелектуальних методів мультиагентної оптимізації, заснованих на моделюванні суспільного інтелекту, відносяться: метод мурашиних колоній (Ant Colony Optimization, ACO) [3, 4], метод бджолоїної колонії (Bee Colony Optimization, BCO) [5, 6], оптимізація з використанням рою часток (Particle Swarm Optimization, PSO) [7] та інші методи [8]. Дані методи вже ефективно застосовуються для вирішення різних задач оптимізації [3–10].

## 1 ПОСТАНОВКА ЗАВДАННЯ

Метод оптимізації з використанням рою часток (Particle Swarm Optimization, PSO) традиційно використовується для вирішення задачі знаходження глобального оптимуму багатовимірної функції, де показує досить гарні результати [11, 12]. Проте метод характеризується й недоліками, пов'язаними з можливістю занадто сильної зміни швидкості часток, що призводить до збільшення часу пошуку оптимального рішення.

Метою даної роботи є розробка додаткових процедур для методу оптимізації з використанням рою часток для виключення його недоліків, пов'язаних з лавиноподібною зміною швидкості часток, що дозволяє покращити оптимізаційний процес.

Основними завданнями роботи являються:

- дослідження методу оптимізації з використанням рою часток;
- виявлення основних недоліків досліджуваного методу;
- створення методу оптимізації на основі моделювання поведінки рою часток, що враховує головні недоліки базового методу;
- розробка програмного забезпечення, що реалізує запропонований метод, та його тестування.

## 2 МЕТОД ОПТИМІЗАЦІЇ 3 ВИКОРИСТАННЯ РОЮ ЧАСТОК

В PSO-методі, особини, що називаються частками, переміщуються в багатовимірному просторі рі-

шень. Зміни координат часток усередині простору пошуку обумовлюються природною соціально-психологічною тенденцією часток конкурувати між собою. Отже, зміни в стані частки залежать від досвіду й знань її сусідів. У цьому випадку, слово «знання» є синонімом «інформації». Результат такого моделювання полягає в тому, що процес пошуку мотивує частки недетермінованим чином повертатися в оптимальні ділянки простору рішень.

Особини в рої часток мають дуже просту поведінку: вони прагнуть перевершити досягнення сусідніх часток і поліпшити власні. Таким чином, емергентна властивість даної системи полягає в дослідженні оптимальних ділянок багатовимірному простору пошуку.

PSO метод керує роєм часток і кожна частка являє собою потенційне рішення [28, 29]. За аналогією з еволюційними стратегіями, рій можна трактувати як популяцію, а частку як індивіда (або хромосому). Таким чином, частка «літає» у багатовимірному просторі рішень, та її позиція визначається виходячи із власного досвіду і досвіду своїх сусідів.

За допомогою  $x_i(t)$  позначимо позицію частки  $i$  у просторі пошуку у момент часу  $t$  ( $t$  позначає дискретні значення часу). Позиція частки змінюється шляхом додавання швидкості  $v_i(t)$  до поточної позиції:

$$x_i(t + 1) = x_i(t) + v_i(t + 1).$$

Початковий стан визначається в такий спосіб:  $x_i(0) = U(x_{\min}, x_{\max})$ , де  $U(a, b)$  є функцією генерації випадкових чисел з діапазону  $[a, b]$ . Дана формула являє собою вектор швидкості й визначає сам оптимізаційний процес, а також відображує використання, як отриманих знань частки, так й обмін інформацією із сусідніми частками. Власні знання самої частки, що також називаються *когнітивним компонентом* формули швидкості, прямо пропорційні поточній відстані частки від її найкращого положення, що було знайдено з моменту старту її життєвого циклу. А обмін інформацією даної особини з іншими є *соціальним компонентом* формули швидкості.

Відомі два підходи, що є різновидами базового PSO-методу: *gbest* та *lbest*, які відрізняються ступенем зв'язаності часток у просторі пошуку.

В *gbest* PSO-методі швидкість частки розраховується за формулою:

$$v_{ij}(t + 1) = v_{ij}(t) + c_1 r_{1j}(t)[y_{ij}(t) - x_{ij}(t)] + c_2 r_{2j}(t)[y_j^*(t) - x_{ij}(t)],$$

де  $v_{ij}(t)$  – швидкість частки  $i$  у вимірі  $j$  ( $j = 1, \dots, n_x$ ) у момент часу  $t$ ;  $x_{ij}(t)$  – позиція частки  $i$  у вимірі  $j$ ;  $y_j^*(t)$  – найкраща досягнута позиція рою (у просторі рішень);  $c_1$  та  $c_2$  – позитивні константи прискорення, які використовуються для варіювання ваг

когнітивного й соціального компонентів швидкості частки відповідно;  $r_{1j}(t), r_{2j}(t) = U(0, 1)$  є випадковими значеннями з діапазону  $[0, 1]$ . Ці випадкові величини привносять стохастичний елемент у роботу методу.

Величина  $y_i$  відображає найкращу позицію частки  $i$ , що вона відвідувала, починаючи з першої ітерації. Наступна оптимальна позиція частки  $i$  у момент часу  $t + 1$  розраховується за формулою

$$y_i(t+1) = \begin{cases} y_i(t), & \text{якщо } f(x_i(t+1)) \geq f(y_i(t)); \\ x_i(t+1), & \text{якщо } f(x_i(t+1)) < f(y_i(t)), \end{cases}$$

де  $f: \mathfrak{R}^{n_x} \rightarrow \mathfrak{R}$  – фітнес-функція,  $\mathfrak{R}^{n_x}$  – множина значень незалежних змінних,  $\mathfrak{R}$  – множина значень оптимізованої величини. Також, як і в еволюційних підходах, вона є мірою близькості даного рішення до оптимального, також фітнес-функція визначає продуктивність, або якість частки.

В lbest PSO-методі швидкість розраховується за формулою:

$$v_{ij}(t+1) = v_{ij}(t) + c_1 r_{1j}(t)[y_{ij}^*(t) - x_{ij}(t)] + c_2 r_{2j}(t)[y_{ij}^*(t) - x_{ij}(t)],$$

де  $y_{ij}^*(t)$  – найкраща позиція, знайдена сусідами  $i$ -ї частки у вимір  $j$ . Локально найкраща позиція частки  $y_i^*$ , тобто краща позиція, знайдена в сусідстві  $N_i$ , що визначається як

$$y_i^*(t+1) \in \{N_i | f(y_i^*(t+1))\} = \min\{f(x)\}, \forall x \in N_i,$$

де сусідство визначається за формулою

$$N_i = \{y_{i-n_{N_i}}(t), y_{i-n_{N_i}+1}(t), \dots, y_{i-1}(t), y_i(t), y_{i+1}(t), \dots, y_{i+n_{N_i}}(t)\}$$

для сусідства розмірністю  $n_{N_i}$ .

Метод PSO може бути описаний у вигляді послідовності таких кроків.

Крок 1. Якщо умови закінчення виконано, тоді закінчити роботу методу; у протилежному випадку – перейти до кроку 2.

Крок 2. Створити й ініціалізувати  $n_x$ -вимірний рій.

Крок 3. Встановити:  $i = 1$ .

Крок 4. Визначення локально найкращої позиції. Якщо виконано умову  $f(x_i) < f(y_i)$ , тоді встановити:  $y_i = x_i$ .

Крок 5. Визначення глобально найкращої позиції. Встановити:  $y^* = y_i$ .

Крок 6. Встановити:  $i = i + 1$ .

Крок 7. Якщо  $i < n_s$ , тоді виконати перехід до кроку 3, у протилежному випадку – до кроку 8.

Крок 8. Встановити:  $i = 1$ .

Крок 9. Оновити швидкість частки, використовуючи формулу gbest PSO або lbest PSO.

Крок 10. Оновити позицію частки.

Крок 11. Встановити:  $i = i + 1$ .

Крок 12. Якщо  $i < n_s$ , тоді виконати перехід до кроку 9, у протилежному випадку – до кроку 1.

Закінчення виконання PSO-методу відбувається, коли виконується хоча б один із критеріїв зупинення. Існують два аспекти, які повинні бути враховані при виборі подібних мір.

1. Умова зупинення не повинна спричинити передчасну збіжність PSO-методу. У протилежному випадку будуть отримані субоптимальні рішення.

2. При перевірці критеріїв зупинення фітнес-функція має викликатися мінімальну кількість разів, оскільки від цього сильно залежить швидкість роботи PSO-методу.

Як критерії зупинення використовують:

– досягнення максимальної кількості ітерацій;

– знаходження задовільного рішення;

– вихід при відсутності поліпшень протягом багатьох ітерацій;

– вихід за умови, що перша похідна фітнес-функції в околиці частки або часток дорівнює нулю.

### 3 МЕТОД PSO З КЕРУВАННЯМ ЗМІНОЮ ШВИДКОСТІ ЧАСТОК

Ефективний оптимізаційний метод має володіти двома важливими властивостями: повністю досліджувати простір рішень задачі, і фіксувати пошук біля потенційних оптимумів. В PSO-методі це визначається формулою швидкості часток.

В ранніх реалізаціях PSO-методу було виявлено, що швидкості можуть занадто швидко зростати, особливо швидкості тих часток, які знаходяться далеко від свого власного оптимуму або загального оптимуму своїх сусідів. У результаті, такі частки можуть залишати простір пошуку, що дуже негативно впливає на роботу методу в цілому. Для запобігання подібним ситуаціям варто контролювати діапазон зміни швидкостей часток. Якщо швидкість частки перевищує максимально допустиму, то її варто знизити до дозволеного рівня. Нехай  $V_{\max,j}$  – максимально припустима швидкість частки у вимірі  $j$ . Тоді швидкість частки пропонується змінювати у такий спосіб:

$$y_{ij}(t+1) = \begin{cases} v'_{ij}(t+1), & \text{якщо } v_{ij}(t+1) = V_{\max,j}; \\ V_{\max,j}, & \text{якщо } v'_{ij}(t+1) \geq V_{\max,j}, \end{cases}$$

де значення  $v'_{ij}$  розраховується у відповідності до gbest PSO або lbest PSO.

Величина  $V_{\max,j}$  є досить важливою, оскільки вона контролює загальні масштаби пошуку. Більші значення змушують PSO-метод досліджувати простір за-

дачі більш ретельно. Якщо значення  $V_{\max,j}$  є досить невеликим, то PSO-метод може взагалі не покинути локальних ділянок, а також збільшується кількість ітерацій, потрібних для досягнення оптимуму.

Виходячи з зазначеного робимо висновки, що  $V_{\max,j}$  варто підбирати таким чином, щоб балансувати між швидким і повільним пересуваннями часток, і між фіксацією пошуку й загальним дослідженням простору. Пропонується розраховувати  $V_{\max,j}$  у такий спосіб:

$$V_{\max,j} = \delta(x_{\max,j} - x_{\min,j}),$$

де  $x_{\max,j}$  та  $x_{\min,j}$  – величини, які, відповідно, знизу та зверху обмежують діапазон значень виміру  $j$ , а  $\delta \in (0,1]$ . Очевидно, що величина  $\delta$  є проблемно-залежною.

Відзначимо дві важливі властивості запропонованої модифікації PSO-методу.

Величина  $V_{\max,j}$  обмежує не простір пошуку, всередині якого пересуваються частки, а тільки швидкості часток (точніше діапазон їхньої зміни за одну ітерацію).

Максимальна швидкість часток визначається для кожного виміру окремо й обумовлюється тільки його фізичною природою.

Виходячи з запропонованого підходу можна відзначити, що обмеження швидкості частки може викликати зміну напрямку її пересування, що може як позитивно, так і негативно відбиватися на пошуку оптимуму. Цей ефект зображено на рис. 1. У цьому випадку,  $x_i(t+1)$  – це позиція  $i$ -ї частки без обмежень за швидкістю. А величина  $x'_i(t+1)$  вже відображає позицію частки після того, як швидкість стала рівною  $V_{\max,j}$ .

Небажаний ефект може виникнути у випадку, якщо швидкості всіх часток стануть рівними  $V_{\max,j}$ .

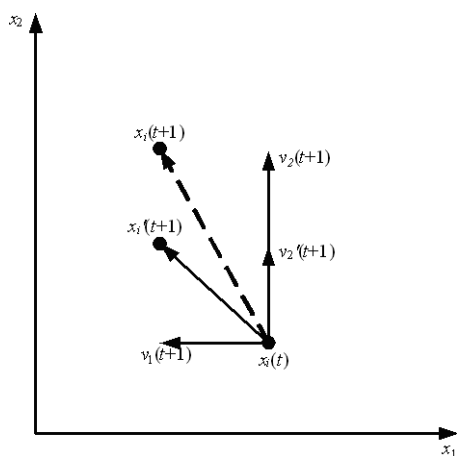


Рисунок 1 – Ефект використання обмеження швидкості частки

Для його запобігання можна запропонувати з ходом ітерацій зменшувати значення  $V_{\max,j}$ . Тоді модифікований PSO-метод буде послідовно звужувати масштаби дослідження простору пошуку, що робить оптимізаційний процес більш ефективним.

Тоді метод PSO з керуванням зміною швидкості часток може бути описаний у вигляді послідовності наступних кроків.

Крок 1. Якщо умови закінчення виконано, тоді припинити роботу методу; у протилежному випадку – перехід до кроку 2.

Крок 2. Створити й ініціалізувати  $n_x$ -вимірний рій. Також розрахувати максимально допустимі швидкості часток:

$$V_{\max,j} = \delta(x_{\max,j} - x_{\min,j}).$$

Крок 3. Встановити:  $i = 1$ .

Крок 4. Визначити локально найкращу позицію. Якщо виконано умову  $f(x_i) < f(y_i)$ , тоді встановити  $y_i = x_i$ .

Крок 5. Визначити глобально найкращу позицію. Встановити:  $y^* = y_i$ .

Крок 6. Встановити:  $i = i + 1$ .

Крок 7. Якщо  $i < n_s$ , то виконати перехід до кроку 3, у протилежному випадку – до кроку 8.

Крок 8. Встановити:  $i = 1$ .

Крок 9. Оновити швидкість частки, використовуючи запропонований підхід:

$$v_{ij}(t+1) = \begin{cases} v'_{ij}(t+1), & \text{якщо } v_{ij}(t+1) = V_{\max,j}; \\ V_{\max,j}, & \text{якщо } v'_{ij}(t+1) \geq V_{\max,j}. \end{cases}$$

Крок 10. Оновити позицію частки.

Крок 11. Встановити:  $i = i + 1$ .

Крок 12. Якщо  $i < n_s$ , то виконати перехід до кроку 9, у протилежному випадку до кроку 13.

Крок 13. Зменшити величину  $\delta$ :  $\delta = \delta^\alpha$ , де  $\alpha \in (0; 1)$ . Зменшення величини  $\delta$  призводить до того, що на наступній ітерації буде зменшена й величина  $V_{\max,j}$ . Перехід до кроку 1.

Розроблений метод PSO з керуванням зміною швидкості часток виключає небажаний ефект занадто великої зміни швидкості часток, що дозволяє покращити оптимізаційний процес за рахунок підвищення швидкості роботи методу.

#### 4 ЕКСПЕРИМЕНТИ Й РЕЗУЛЬТАТИ

Запропонований метод PSO з керуванням зміною швидкості часток було програмно реалізовано у середовищі пакету Matlab 7.0.

Використовуючи створене програмне забезпечення, проводилися експерименти зі знаходження оптимального значення функцій з метою дослідження ефективності запропонованого методу PSO у порів-

Таблиця 1 – Тестові функції та знайдені для них оптимуми

Функція	Фактичний оптимум	Оптимум, отриманий базовим PSO	Оптимум, отриманий запропонованим методом
$f_1 = \sum_{i=1}^n  x_i \cdot \sin(x_i) + 0,1 \cdot x_i , n = 2$ $-5 \leq x_1 \leq 5;$ $-5 \leq x_2 \leq 5$	min: $f_1^*(0; 0) = 0$	$x_1 = 0,004;$ $x_2 = 0,008;$ $f_1^* = 0,0013$	$x_1 = 0,0009;$ $x_2 = 0,005;$ $f_1^* = 0,00062$
$f_2 = 20 + e - 20 \exp\left(-0,2 \sqrt{\frac{1}{n} \sum_{i=1}^n x_i^2}\right) - \exp\left(\frac{1}{n} \sum_{i=1}^n \cos(2\pi x_i)\right), n = 2$ $-5 \leq x_1 \leq 5;$ $-5 \leq x_2 \leq 5$	min: $f_2^*(0; 0) = 0$	$x_1 = 0,000299;$ $x_2 = 0,0002;$ $f_2^* = 0,0010$	$x_1 = 0,00005;$ $x_2 = 0,00028;$ $f_2^* = 0,000807$

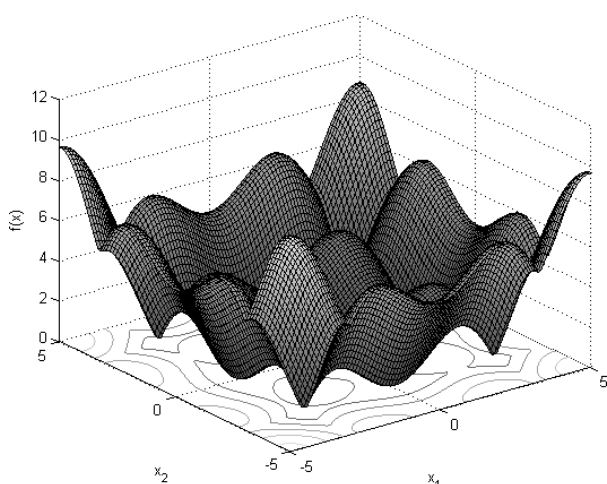


Рисунок 2 – Графік тестової функції  $f_1$

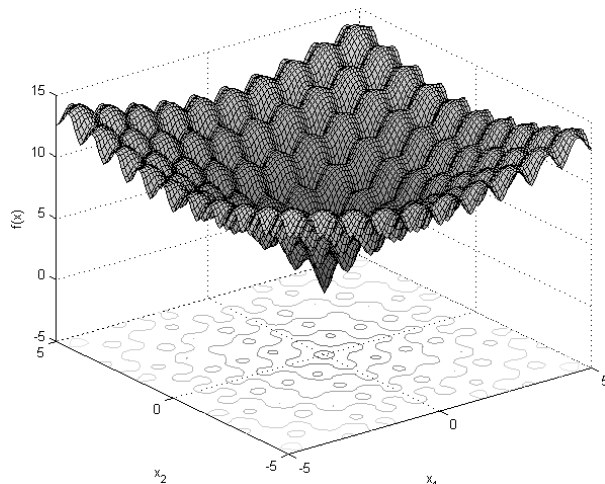


Рисунок 3 – Графік тестової функції  $f_2$

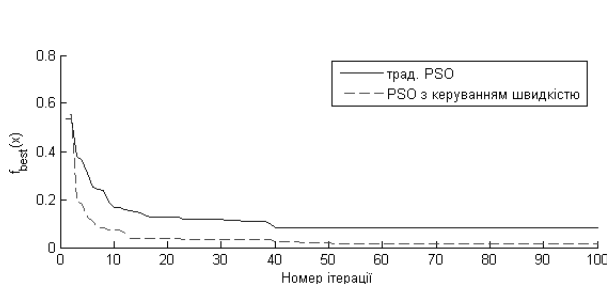


Рисунок 4 – Результати роботи для тестової функції  $f_1$

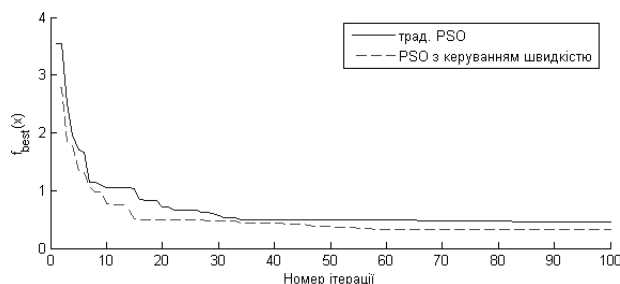


Рисунок 5 – Результати роботи для тестової функції  $f_2$

янні з традиційним методом PSO. В експериментах використовувались тестові функції, які подано в табл. 1, де також наведено отримані результати.

Головною особливістю запропонованих тестових функцій є їхня полімодальність, тобто велика кількість локальних оптимумів, що дозволяє перевірити, наскільки схильний запропонований метод до зацілювання в локальних оптимумах.

Графіки тестових функцій подано на рис. 2 та 3.

Результати роботи розробленого програмного забезпечення для відповідних функцій зображено на рис. 4 та 5.

Виходячи з наведених рисунків, можна зробити висновок, що запропонований метод знаходить більш точно точку оптимуму. При цьому знаходження оптимального значення відбувається швидше у порівнянні з базовим методом оптимізації з використанням рою часток.

## ВИСНОВОК

У роботі вирішено актуальне завдання прискорення роботи методу оптимізації з використанням рою часток.

Наукова новизна роботи полягає в тому, що метод оптимізації з використанням рою часток, на основі якого розроблено метод з керуванням зміною швидкості часток, дістав подальшого розвитку. Запропоновані процедури враховують недоліки базового методу, пов'язані з можливістю занадто великою зміною швидкості часток, що негативно впливає на оптимізаційний процес.

Практична цінність результатів роботи полягає в тому, що розроблено програмне забезпечення, яке реалізує запропонований метод оптимізації на основі моделювання поведінки рою часток з керуванням зміною швидкості часток, і дозволяє провести експерименти по знаходженню оптимуму функцій. Запропонований метод може бути використаний у будь-якій задачі оптимізації, де необхідно знайти значення змінних, за яких залежність, що досліджується, набуває оптимального значення.

Роботу виконано за держбюджетною темою «Науково-методичні основи та математичне забезпечення для автоматизації і моделювання процесів керування та підтримки прийняття рішень на основі процедур розпізнавання й еволюційної оптимізації в нейромрежному і нечіткологічному базисах» (№ держ. реєстрації 0106U008621).

## ПЕРЕЛІК ПОСИЛАНЬ

1. Beni G. Swarm Intelligence / Beni G., Wang J. // Annual Meeting of the Robotics Society : proceedings of Seventh International Conference. – Tokyo : RSJ Press, 1989. – P. 425–428.
2. Bonabeau E. Swarm Intelligence: From Natural to Artificial Systems / Bonabeau E., Dorigo M., Theraulaz G. – New York : Oxford University Press, 1999. – 320 p.
3. Dorigo M. The Ant System: Optimization by a Colony of Cooperating Agents / Dorigo M., Maniezzo V., Color-

- ni A. // IEEE Transactions on Systems, Man, and Cybernetics. – 1996. – Part B, № 26(1). – P. 29–41.
4. Олейник А. А. Сравнительный анализ методов оптимизации на основе метода муравьиных колоний / Олейник А. А. // Комп'ютерне моделювання та інтелектуальні системи : збірник наукових праць / за ред. Д. М. Пізи, С. О. Субботіна. – Запоріжжя : ЗНТУ, 2007. – С. 147–159.
5. Camazine S. A Model of Collective Nectar Source by Honey Bees: Self-organization Through Simple Rules / Camazine S., Sneyd J. // Journal of Theoretical Biology. – 1991. – № 149. – P. 547–571.
6. Sumpter D. J. T. Formalising the Link between Worker and Society in Honey Bee Colonies / Sumpter D. J. T., Broomhead D. S. // Lecture Notes In Computer Science : Proceedings of the First International Workshop on Multi-Agent Systems and Agent-Based Simulation (MABS '98). – Berlin : Springer, 1998. – P. 95–110.
7. Kennedy J. Particle Swarm Optimization / Kennedy J., Eberhart R. C. // International Conference on Neural Networks : Proceedings of the IEEE. – NJ : IEEE Press, 1995. – P. 1942–1948.
8. De Castro L. N. Artificial Immune Systems. Part I. Basic Theory And Applications / De Castro L. N., Von Zuben F. J. // Technical Report No. Rt Dca 01/99. – Brazil : Feec/Unicamp, 1999. – 95 p.
9. Colorni A. Ant System for Job-shop Scheduling / Colorni A., Dorigo M., Maniezzo V., Trubian M. // Belgian Journal of Operations Research, Statistics and Computer Science (JORBEL). – 1994. – № 34. – P. 39–53.
10. Субботин С. А. Отбор информативных признаков на основе модифицированного метода муравьиных колоний / Субботин С. А., Олейник А. А., Яценко В. К. // Радиоелектроніка та інформатика. – 2006. – № 1. – С. 65–69.
11. Engelbrecht A. Computational Intelligence: an Introduction / Engelbrecht A. – Sidney : John Wiley and Sons, 2007. – 597 p.
12. Abraham A. Swarm Intelligence in Data Mining / Abraham A., Grosan G. – Berlin : Springer, 2006. – 267 p.

Надійшла 2.10.2008  
Після доробки 2.12.2008

*Исследован метод оптимизации на основе моделирования поведения роя частиц. Разработана модификация исследованного метода с управлением изменением скорости частиц. Проведены эксперименты по нахождению глобального оптимума многомерной функции на основе предложенной модификации.*

*A particle swarm optimization method has been considered. A modification of PSO-method with velocity control has been developed. The experiments on finding global optimum for multidimensional function on the basis of suggested modification has been carried out.*

УДК 621.365.036

С. Г. Удовенко, А. А. Шамраев, Г. Дибс

## ЦИФРОВОЕ НЕЙРО-НЕЧЕТКОЕ УПРАВЛЕНИЕ БЛОКАМИ ЛИНЕЙНОГО СИЛЬНОТОЧНОГО УСКОРИТЕЛЯ ЭЛЕКТРОНОВ

*Предложена структура и разработаны методы цифрового управления блоками линейного ускорителя с применением нечетких и нейросетевых моделей. Приведены результаты моделирования, подтверждающие эффективность рассмотренных методов.*

© Удовенко С. Г., Шамраев А. А., Дибс Г., 2009

## ВВЕДЕНИЕ

Линейный резонансный ускоритель электронов представляет собой сложную электрофизическую установку, в которой в течение длительного времени

(порядка нескольких миллисекунд) происходит накопление энергии от внешнего маломощного источника (порядка 100 кВт) с последующей быстрой (порядка нескольких микросекунд) трансформацией накопленной энергии в энергию электронного пучка. Поскольку отношение времен накопления и трансформации составляет величину порядка 1000, то мощность ускоренного пучка с учетом к. п. д. = 0,1 будет в 100 раз больше мощности внешнего источника, т. е. составлять 10 МВт. При токе пучка  $I = 1$  А энергия будет составлять 10 МэВ. Преобразование энергии происходит через цепочку: накопитель заряда – коммутатор – импульсный трансформатор – СВЧ-источник – ускоряющая секция. Именно в последнем элементе при ускорении электронного пучка происходит трансформация энергии от внешнего источника в энергию электронного пучка. Ускоряющая секция представляет собой отрезок замедляющей структуры (обычно с фазовой скоростью волны, равной скорости света) с элементами ввода СВЧ мощности. В ряде случаев необходимо проводить регулировку или стабилизацию энергии ускоренных электронов. Отсутствие полной информации об условиях функционирования объектов, а также об их динамических характеристиках и характере действующих помех обуславливают необходимость применения при управлении такими объектами адаптивного и робастного подходов, допускающих возможность использования при синтезе регуляторов упрощенных (в частности, нейросетевых и нечетких) моделей. Наиболее распространенным типом замедляющих структур является диафрагмированный волновод, представляющий собой круглый цилиндрический волновод с равномерно расположенными диафрагмами, отверстия в которых служат как для пролета частиц, так и для передачи СВЧ энергии. Крупнейшим ускорительным центром Украины является Национальный научный центр «Харьковский физико-технический институт». В настоящее время в нем функционирует семь линейных резонансных ускорителей электронов на различные энергии. Одним из них является линейный ускоритель ЛУ-40, который используется для проведения различных облучательных программ. В ряде случаев необходимо проводить регулировку или стабилизацию энергии ускоренных электронов. Отсутствие полной информации об условиях функционирования объектов, а также об их динамических характеристиках и характере действующих помех обуславливают необходимость применения при управлении такими объектами адаптивного и робастного подходов, допускающих возможность использования при синтезе регуляторов упрощенных (в частности, нейросетевых и нечетких) моделей.

## **ПОСТАНОВКА ЗАДАЧИ**

В общем случае построение системы цифрового управления линейным сильноточным ускорителем (СУЛСУ) предполагает необходимость решения ряда взаимосвязанных задач моделирования и разработки регуляторов для отдельных блоков и секций СУЛСУ. В частности, для регулирования энергии ускоренных электронов важной является стабилизация тока анализирующего магнита. При этом целесообразно разработать микропроцессорный блок управления электропитанием магнитных элементов, позволяющий существенно повысить качество работы СУЛСУ, и цифровой регулятор энергии пучка электронов в секциях ускорителя. Ниже приведены результаты решения этих задач с использованием нейросетевых и нечетких методов.

## **ПОДСИСТЕМА НЕЧЕТКОГО УПРАВЛЕНИЯ ЭЛЕКТРОПИТАНИЕМ МАГНИТНЫХ ЭЛЕМЕНТОВ ЛИНЕЙНОГО УСКОРИТЕЛЯ**

Для преодоления недостатков используемой в настоящее время системы цифрового управления магнитными спектрометрами линейных ускорителей была разработана микропроцессорная подсистема управления электропитанием магнитных элементов (СУЭМЭ) для ускорителей типа ЛУ-40. Структурная схема СУЭМЭ показана на рис. 1. СУЭМЭ формирует сигнал опорного напряжения на входе схемы стабилизации тока анализирующего магнита. Напряжение формируется двенадцатиразрядным цифро-аналоговым преобразователем (ЦАП), обеспечивая дискретность шага по току магнита менее 0,1 А или 0,07 % от максимального тока анализирующего магнита. Интерфейсная часть системы выполнена на микроконтроллере MSP430F1611, основными задачами которого являются формирование управляющих сигналов для ЦАП, обслуживание местного пульта управления и обеспечение связи с ЭВМ. Двусторонняя связь контроллера с управляющей ЭВМ осуществляется по оптоволоконному кабелю (ВОЛС) в последовательном формате по протоколу RS232C. Применение оптического канала связи обеспечивает гальваническую развязку контроллера от линии связи, что требуется по условию обеспечения соответствующего уровня безопасности и помехозащищенности при установке опорного напряжения.

Электропитание магнита осуществляется от генератора постоянного тока ПН-145 с независимым возбуждением. Для обеспечения стабилизации и регулировки тока магнита, обмотка возбуждения генератора подключена к управляемому источнику тока, величина тока в обмотке возбуждения генератора задается контроллером магнита. Управляемый источник тока построен на базе линейного источника на-



**НЕЙРОРЕГУЛЯТОР ЕНЕРГІЇ ПУЧКА ЕЛЕКТРОНОВ ЛІНЕЙНОГО УСКОРИТЕЛЯ**

Аналіз показав, що для стабілізації енергії електронів в СУЛСУ для ЛУ-40 цілесообразно використовувати нейросетеву модель (НСМ).

При побудові НСМ односекційного лінійного ускорителя в якості основних параметрів були вибрані чотири: фаза інжекції  $\phi_0$ , ток  $I$  інжекції, початкова енергія (енергія інжекції) частиц  $W_0$  і потужність зовнішнього джерела  $P_0$ . При цьому попередньо була проведена стандартизація змінних, що складається в центруванні і нормуванні вхідних і вихідних змінних:

$$\bar{x}_{jk} = \frac{x_{jk} - m_{x_j}}{\sigma_{x_j}}, \quad \bar{y}_k = \frac{y_k - m_y}{\sigma_y},$$

$$j = \overline{1, 4}, \quad k = \overline{1, 10000},$$

де  $m_{x_j}$  – середнє вибірки  $x_{jk}$ ,  $k = \overline{1, 10000}$ ;  $m_y$  – середнє вибірки  $y_k$ ,  $k = \overline{1, 10000}$ ;  $\sigma_{x_j}$ ,  $\sigma_y$  – дисперсії вибірок  $x_{jk}$  і  $y_k$  відповідно.

При побудові НСМ проводились експерименти з різними структурами мережі, починаючи з найпростішої мережі без зворотних зв'язків, що складається з двох шарів, з 4 нейронами в прихованому шарі і 1 нейроном в вихідному шарі, і закінчуючи рекуррентними мережами, що включають затримки по вхідним сигналам, так і зворотні зв'язки. Для вибору найкращої архітектури мережі використовувалась процедура дискримінації НСМ. В якості функції активації нейронів в прихованому шарі використовувалась гіперболічний тангенс, а в вихідному шарі – лінійна функція. Налаштування мережі здійснювалось на основі 10000 навчальних пар, вимагає точності задавалась на рівні  $1 \cdot 10^{-6}$ . Для навчання мережі використовувалась алгоритм Левенберга – Марквардта. Рис. 2 відображає динаміку навчання НСМ, структура якої показана на рис. 3 і рис. 4. Як видно з рис. 2, а, мережа не досягла вимогової точності, а величина помилки стабілізувалась на рівні  $2,4 \cdot 10^{-5}$ . На рис. 2, б показана крива навчання для найкращої НСМ. Заданої точності вибрана мережа досягла за 132 ітер-

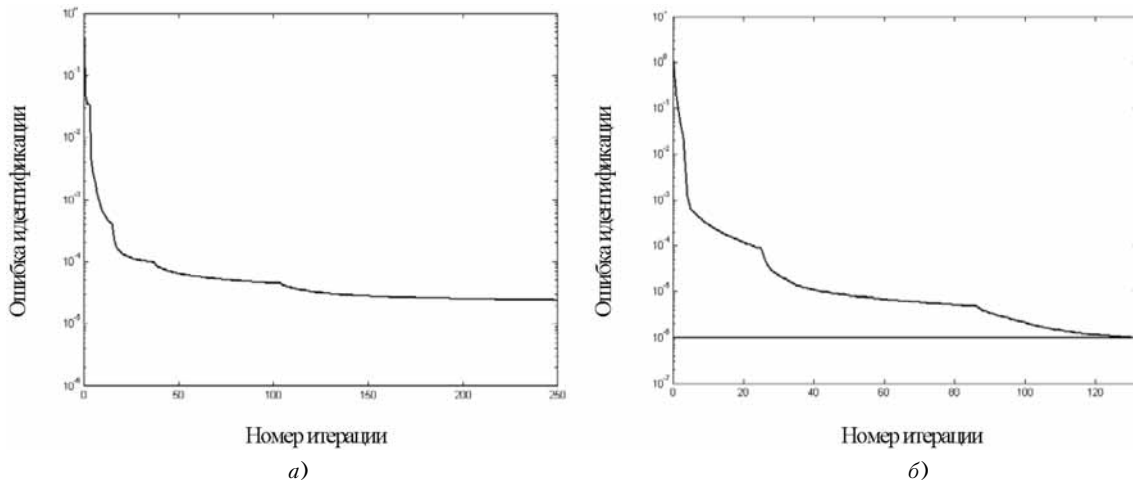


Рисунок 2 – Динаміка навчання нейросетевих моделей:

а – НСМ 1а; б – НСМ 16

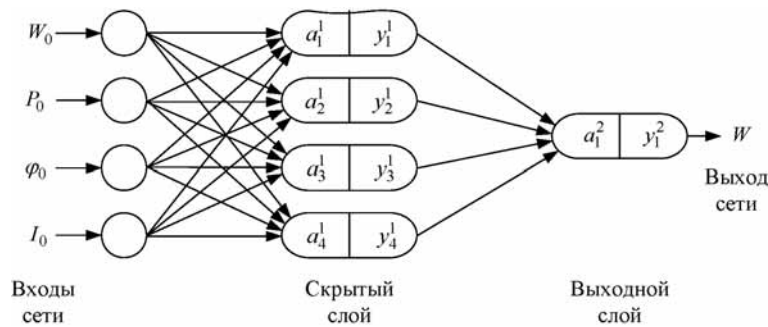


Рисунок 3 – НСМ 1а ускорителя ЛУ-40



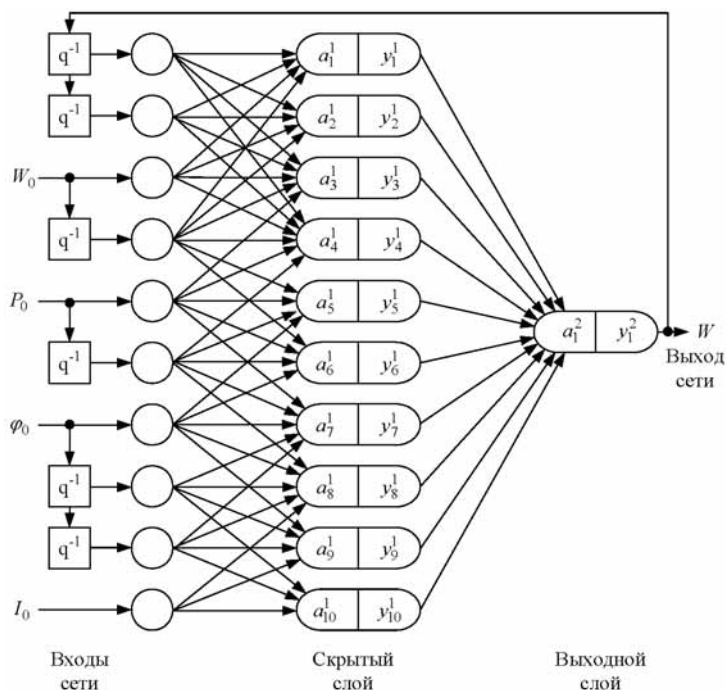


Рисунок 4 – НСМ 16 ускорителя ЛУ-40

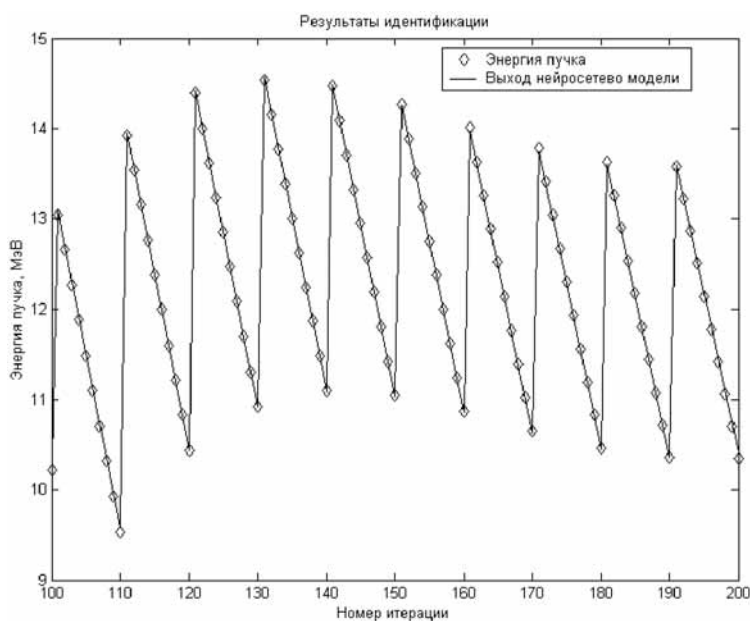


Рисунок 5 – Результаты идентификации

рации. Результаты идентификации объекта представлены на рис. 5. Здесь сплошной линией показаны выходы НСМ, а ромбами – величина энергии пучка электронов. Как видно из рисунка, сеть способна с высокой точностью воспроизводить поведение объекта. На основе полученной НСМ был построен нейро-регулятор, обеспечивающий как стабилизацию энергии пучка электронов, так и ее регулирование. Для

выработки требуемых управляющих воздействий использовался алгоритм, предложенный в [2]. Но, поскольку рассматриваемый объект имеет только один выход, то при реализации алгоритма исключается операция обращения матриц, что обеспечивает стабильность и высокое качество управления. На рис. 6 представлены результаты управления энергией электронов в ускорителе КУТ. Как видно из рисунка,

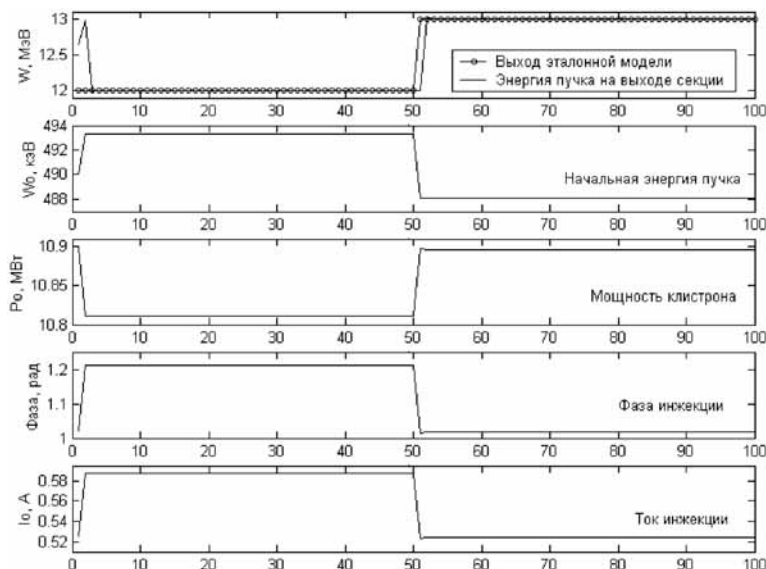


Рисунок 6 – Управление энергией электронов в ускорителе ЛУ-40

нейрорегулятор обеспечивает высокое качество управления при стабилизации энергии пучка электронов. Небольшая задержка обусловлена временем распространения сигнала в сети. Короткие переходные процессы (не более 4 тактов управления), возникающие на выходе сети, вызваны установкой новых целей управления и связаны с изменениями сигналов управления.

## ВЫВОДЫ

Анализ полученных результатов свидетельствует о том, что нечеткие и нейросетевые методы, реализуемые на базе микропроцессорных средств, является эффективным средством построения цифровых систем управления секциями линейных ускорителей электронов. Предпосылкой для использования разработанных методов является внедрение предложенной подсистемы управления электропитанием магнитных элементов.

## ПЕРЕЧЕНЬ ССЫЛОК

1. Довбня А. Н. Определение параметров модели магнитного спектрометра по экспериментальным данным / А. Н. Довбня, С. Г. Удовенко, А. А. Шамраев // Радиоэлектроника, информатика, управління. – 2008. – Вып. 1 (19). – С. 45–49.
2. Efstathiou J. Architectures and techniques of artificial intelligence in process control / J. Efstathiou // Artificial intelligence in real-time control, 1992. – Oxford : Pergamon, 1993. – V. 17. – P. 348–361.

Надійшла 28.04.2009  
Після доробки 11.05.2009

*Запропоновано структуру та розроблено методи цифрового керування блоками лінійного прискорювача з використанням нечітких та нейронних моделей. Наведено результати моделювання, що підтверджують ефективність запропонованих методів.*

*In the work the control structure is offered and the digital control methods for the linear accelerator blocks with use of fuzzy and artificial neural networks models are developed. The simulation data is confirmative efficiency of the considered methods.*

# ПРОГРЕСИВНІ ІНФОРМАЦІЙНІ ТЕХНОЛОГІЇ

## ПРОГРЕССИВНЫЕ ИНФОРМАЦИОННЫЕ ТЕХНОЛОГИИ

### PROGRESSIVE INFORMATION TECHNOLOGIES

---

---

УДК [004.78:33] (075.8)

В. П. Авраменко

#### СИСТЕМА ПОДДЕРЖКИ ПРИНЯТИЯ РЕШЕНИЙ ПО ЛИКВИДАЦИИ НЕШТАТНЫХ СИТУАЦИЙ

---

*Разработаны структура, состав и информационные технологии поддержки принятия решений по ликвидации нештатных и аварийных ситуаций, которые контролируются охранно-пожарной сигнализацией. Приведен пример выбора комплекса аппаратно-программных средств по совокупности локальных критериев, в качестве которых выступают технико-экономические показатели эффективности принятия решений по ликвидации нештатных ситуаций.*

#### **ВВЕДЕНИЕ**

Место решаемой задачи. Успешное функционирование предприятий производственной и непроизводственной сферы, банковских и коммерческих организаций требует обеспечения безопасности объектов деятельности и сотрудников этих организаций на случай возникновения нештатных ситуаций, защиты их от несанкционированного доступа, нападения и принуждения к несанкционированным действиям. Отсюда возникает необходимость разработки системы поддержки принятия решений (СППР), позволяющей повысить достоверность контроля охранно-пожарной сигнализации и ускорить ликвидацию возникшей нештатной ситуации. Современное состояние средств вычислительной техники позволяет оборудовать охра-

няемые объекты современными СППР, позволяющими надежно контролировать состояние охраняемого объекта.

Существующая практика проектирования СППР по совокупности технико-экономических показателей часто наталкивается на противоречие между заданными требованиями и количеством выделенных ресурсов. На начальных этапах проектирования СППР применяются эвристические методы многоальтернативного выбора структуры и состава аппаратно-программных средств создаваемой системы. Численные методы системной оптимизации позволяют выбрать предпочтительные проектные решения по совокупности количественных и качественных показателей, выступающих в качестве локальных критериев. Решение такого класса задач требует разработки моделей и методов, ориентированных на многоальтернативные технологии проектирования СППР.

*Цель решаемой задачи* состоит в создании и апробации методологии и информационных технологий проектирования СППР, направленных на повышение безопасности охраняемых объектов и минимизацию затрат по ликвидации возникших нештатных ситуаций.

## 1 ПОСТАНОВКА ЗАДАЧИ

*Объект исследования* – процессы поддержки принятия решений, направленные на повышение безопасности охраняемых объектов и минимизацию затрат по ликвидации возникших нештатных ситуаций.

*Предмет исследования* – методы, модели и информационные технологии поддержки принятия решений по ликвидации возникших нештатных ситуаций.

*Методы исследования* – системный анализ принятия проектных решений, направленных на повышение безопасности охраняемых объектов; многоальтернативный подход к выбору комплекса технических средств.

**Анализ состояния решаемой задачи.** Научным аспектам создания СППР по ликвидации нештатных ситуаций посвящены работы [1–4]. Принцип многоальтернативности принимаемых решений отражен в работе [1], в которой из множества допустимых решений выбирается более предпочтительный вариант по совокупности технико-экономических показателей безопасного функционирования охраняемых объектов. Для оперативного управления ликвидацией аварийных ситуаций применяются методы оперативной аналитической обработки данных [2]. Методам ликвидации аварийных ситуаций в энергосистемах посвящены работы [3–4]. В настоящее время практически отсутствуют работы по созданию методологий и технологий создания СППР для аварийных и нештатных ситуаций.

## 2 РЕЗУЛЬТАТЫ РЕШЕНИЯ ЗАДАЧИ

**Методология создания СППР безопасности охраняемого объекта.** В связи с ростом числа средних и мелких предприятий и учреждений, занимающихся финансовой и коммерческой деятельностью, возникла настоятельная необходимость создания систем, обеспечивающих их охранную и противопожарную безопасность. Бурный рост выпуска устройств контроля состояния и наблюдения, а также общая тенденция снижения их стоимости, позволяют оборудовать охраняемые объекты современными электронными средствами, позволяющими круглосуточно и с высокой достоверностью контролировать состояние охраняемых объектов.

Процессы управления в производственной и непроизводственной сферах подразделяются на управление в штатных режимах и нештатных ситуациях. Штатным является режим управления, когда параметры и переменные контролируемого объекта не выходят за пределы предписанных ему значений. Управление штатными ситуациями обычно возлагают на систему автоматического управления. Нештатными принято называть ситуации, когда параметры и пере-

менные контролируемого объекта превышают предписываемые ему значения. Ликвидация нештатных ситуациях возлагается, как правило, на диспетчера, в распоряжении которого имеется информационно-советующая система поддержки принятия решений.

На распознавание возникшей нештатной ситуации и принятие управленческого решения по ее устранению отводится весьма ограниченный интервал времени. Для повышения качества принимаемых управленческих решений целесообразно создание СППР, которые позволят обработать значительно большее количество информации и уменьшат влияние человеческого фактора на принятие решения [3–4].

Основными функциональными задачами охранно-пожарных и защитных СППР являются: обслуживание охраняемых объектов с распределенной структурой; высокое быстродействие в реальном масштабе времени; выработка управленческих решений для штатных и нештатных ситуаций; обеспечение требуемой надежности работы оборудования. Основным критерием выбора СППР является малая относительная вероятность ложного срабатывания и высокая вероятность правильного обнаружения аварийной ситуации (пожар, задымление, проникновение в охраняемую зону).

Обеспечение безопасности предприятий связано с проведением ряда организационных и технических мероприятий. К организационным мероприятиям относятся: оценка возможной угрозы безопасности охраняемого объекта; соблюдение факторов, предотвращающих возникновение нештатных ситуаций; допуск в помещения лиц согласно предписанию; контроль охраняемых помещений от взлома, задымления и затопления помещения, возникновения пожара, передвижения в охраняемом помещении людей. К техническим мероприятиям принадлежат: выбор рационального комплекса технических средств, правильная расстановка технических средств охраны, надлежащее сопровождение технических средств.

СППР нештатных ситуаций должны обеспечивать безопасность производственных и непроизводственных объектов, банковских и финансовых учреждений, а также предприятий сферы обслуживания, имеющих сложную структуру управления. СППР должны срабатывать при появлении признаков пожара или попытке проникновения в помещение. Техническое обеспечение безопасности требует принятия компромиссных решений. Например, правоохранительные органы рекомендуют ставить на окна решетки и устанавливать двойные двери, а пожарные инспекции настаивают на том, чтобы в случае необходимости можно было беспрепятственно покинуть помещение. Все замеченные изменения состояний охраняемых объектов должны фиксироваться в журнале событий, происшедших во время дежурства.

**Архитектура СППР ликвидации нештатных ситуаций.** Математическое описание СППР определяется принципами абстрагирования и многоальтернативности, инструментальными средствами построения математических моделей и наличием вычислительных методов для решения задач рассматриваемого класса. Согласно принципу абстрагирования в модель системы должны включаться те компоненты, которые имеют непосредственное отношение к выполнению системой своих функций или своего целевого назначения. Все второстепенные детали желательно опустить, чтобы чрезмерно не усложнять математическое описание системы.

Принцип многоальтернативности моделирования СППР состоит в том, что никакая единственная модель не может в достаточной степени адекватно описать все аспекты функционирования системы. В достаточной степени адекватное математическое описание системы можно получить путем построения множества моделей, каждая из которых составляется под определенным углом зрения и отражает конкретный аспект поведенческой деятельности функционирования системы.

Для метамодели верхнего уровня СППР часто используется математический аппарат теории множеств, теории категорий, теории графов и теории операторов. Операторная модель может иметь вид:

$$Ax = y, x \in R_x^m, y \in R_y^n,$$

где  $x$  и  $y$  – элементы метрических пространств  $R_x^m$  и  $R_y^n$ ;  $A$  – оператор метамодели, переводящий элементы ресурсов  $x \in R_x^m$  в элементы результата производственно-технологической деятельности  $y \in R_y^n$ .

Архитектура СППР может быть представлена тремя составляющими: *объектами* – функциональными и обеспечивающими подсистемами, внешними и внутренними возмущающими воздействиями; *связями*, описывающими точные границы подсистем и их взаимодействия при управлении данным; *поведениями* – изменяющимися состояниями подсистем, зафиксированных в локальных хранилищах данных.

Математическую модель архитектуры СППР (АрхСППР) можно представить кортежем вида:

$$\text{АрхСППР} = \langle \text{МПК}, \text{МВС}, \text{МСП} \rangle,$$

где МПК – множество подсистем и компонент; МВС – множество взаимосвязей между подсистемами и компонентами; МСП – множество составляющих поведения, изменяющихся во времени данных, состояний, управляющих и возмущающих воздействий.

Модель архитектуры СППР при ликвидации нештатных ситуаций можно представить кортежем вида:

$$\text{МОС} = \langle \text{НПД}, \text{ИПР}, \text{УОИ}, \text{ГЭС}, \text{БД}, \text{БЗ}, \text{ПМР}, \text{ВНР}, \text{ЖПР}, \text{ИИ} \rangle,$$

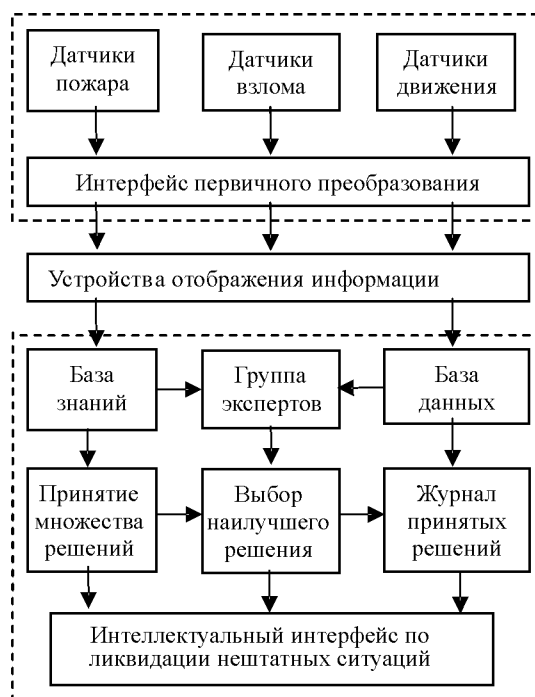


Рисунок 1 – Архитектура СППР по ликвидации нештатной ситуации

где НПД – набор первичных датчиков возникновения пожара, попытки взлома, движения объектов с унифицированными выходами; ИПР – интерфейс первичного преобразования данных; УОИ – устройства отображения информации (звуковые, визуальные, мнемонические); ГЭС – группа экспертов-специалистов по установлению типа и характеристик возникшей нештатной ситуации; БД – база данных; БЗ – база знаний; ПМР – принятие множества допустимых решений; ВНР – выбор наилучшего решения из множества допустимых по ликвидации нештатных ситуаций; ЖПР – журнал принятых решений с указанием лица, ответственного за ликвидацию нештатной ситуации; ИИ – интеллектуальный интерфейс экспертной системы, включающий в себя механизм объяснения, механизм вывода, механизм приобретения знаний (рис. 1).

**Многоальтернативный выбор комплекса технических средств СППР.** Многоальтернативность принятия решений можно представить набором целевых функций  $f_i, i = \overline{1, m}$  и набором альтернатив принимаемых решений  $x = \{x_j\} \subset X, j = \overline{1, n}$  в виде:

$$f_i(x_j) \rightarrow \max_{x \in X}, i = \overline{1, m}, j = \overline{1, n},$$

где  $m$  – количество целевых функций,  $x = \{x_1, x_2, \dots, x_n\} \subset X$  – конечное множество альтернатив принимаемых решений, содержащее  $n$  элементов  $x_j$ . Значения чисел  $m$  и  $n$  должны быть относительно

невелики, поскольку именно они определяют вычислительную сложность решаемой задачи.

После задания частных критериев эффективности используется аддитивный метод линейной свертки

$$J(x_j) = \sum_{i=1}^n \alpha_i f_i(x_j), \quad j = \overline{1, n}$$

для получения исследуемых на оптимальность альтернатив принимаемых решений  $x_1 = A_1$ ,  $x_2 = A_2$ ,  $x_3 = A_3$ .

Для нахождения многоальтернативных решений в условиях неопределенности исходной информации, заданной набором количественных и качественных критериев, применен метод анализа иерархий [1], в основу которого заложен принцип декомпозиции исходной проблемы набором иерархических уровней. Многоальтернативное решение реализуется как процесс поэтапного установления приоритетов. На первом этапе выявляются наиболее важные составляющие задачи; на втором этапе выбирается лучший способ проверки наблюдений, испытания и оценки элементов; на третьем этапе осуществляется выработка и принятие многоальтернативного решения с оценкой его качества.

Глобальный показатель качества выбора комплекса технических средств (КТС) СППР связывается с ее функциональными, техническими и экономическими характеристиками. Такой показатель является достаточно расплывчатым и неопределенным, поэтому желательно воспользоваться набором локальных критериев-заместителей, которые позволяют заменить исходную слабо структурированную задачу другой хорошо структурированной многоальтернативной задачей.

Предположим, что для Региональной системы по чрезвычайным ситуациям требуется выбрать наилучший вариант КТС по совокупности технико-экономических показателей. Альтернативами решений могут выступать следующие аппаратно-программные средства:  $A_1$  – система «Атлант»,  $A_2$  – система «Ай-Ти»,  $A_3$  – система «Мир-АСР»,  $A_4$  – система «Орёл-М»,  $A_5$  – система «Энран». Глобальный критерий эффективности аппроксимируем набором частных критериев, каждый из которых необходимо максимизировать:

–  $f_1(x)$  – гибкость системы, представляющая способность быстро приспосабливаться к изменяющимся условиям и быть использованной как в самостоятельном режиме, так и в составе эксплуатируемой СППР;

–  $f_2(x)$  – производительность, масштабируемость и открытость системы, состоящие в том, что при росте числа пользователей и услуг не должна возникать необходимость доработки программного обеспечения;

–  $f_3(x)$  – аппаратная платформа в виде совокупности согласованных между собой технических средств, обеспечивающих решение заданных функциональных задач с требуемыми качественными показателями;

–  $f_4(x)$  – надежность, характеризующая интегрированное свойство показателей качества инструментальных аппаратно-программных средств системы сохранять работоспособность в процессе их эксплуатации;

–  $f_5(x)$  – экономическая эффективность (экономичность) от внедрения выбранного варианта комплекса аппаратно-программных средств СППР.

Алгоритм метода многоальтернативной оптимизации содержит следующие этапы [1]:

1. Содержательная постановка задачи многоальтернативной оптимизации в условиях неопределенности.

2. Математическая постановка задачи многоальтернативной оптимизации с формированием иерархической структуры обобщенного критерия эффективности и альтернатив принимаемых решений.

3. Ранжирование конечного множества объектов критериев и объектов-альтернатив принимаемых решений

$$p = \{p_1, \dots, p_i, \dots, p_m\}$$

по важности путем задания вектора весовых коэффициентов

$$\alpha = \{\alpha_1, \dots, \alpha_i, \dots, \alpha_m\},$$

значения которых удовлетворяют ограничениям

$$\sum_{i=1} \alpha_i = 1, \quad \alpha_i \geq 0.$$

4. Формирование матрицы парных сравнений  $[Sp_{m \times m}]$  по таким правилам:

– мнение каждого эксперта записывается в виде строки матрицы парных сравнений  $[Sp_{m \times m}]$ ;

– эксперт должен уметь быстро отвечать на вопросы: во сколько раз весовой коэффициент  $\alpha_i$  больше весового коэффициента  $\alpha_j$  или во сколько раз весовой коэффициент  $\alpha_j$  меньше коэффициента  $\alpha_i$ ;

– каждый элемент  $\alpha_{ij}$  матрицы парных сравнений  $[Sp_{m \times m}]$  определяется выражением  $\alpha_{ij} = \alpha_i / \alpha_j$ , где  $\alpha_i$  и  $\alpha_j$  – весовые коэффициенты приоритетности объектов. Размерности сравниваемых коэффициентов  $\alpha_i$  и  $\alpha_j$  должны совпадать, а значения этих коэффициентов не допускают деление на нуль. Если отношение  $\alpha_i / \alpha_j > 1$ , то объект  $p_i$  считается важнее объекта  $p_j$ .

5. Поиск решения задачи многокритериальной оптимизации осуществляется путем поэтапного установления приоритетов. На первом этапе выявляются наиболее важные объекты решаемой проблемы, на втором находится наилучший способ проверки наблюдений, испытания и оценки объектов. На последующих этапах осуществляется выработка рационального решения и оценивание его качества.

Согласованность экспертных мнений является исходной предпосылкой метода анализа иерархий. Для определения согласованности экспертных мнений используется матрица попарных сравнений  $[Sp_{m \times m}]$ .

В качестве меры согласованности используется индекс согласованности. Согласованность матрицы  $[Sp_{m \times m}]$  эквивалентна требованию равенства ее максимального собственного значения  $\lambda_{\max}$  числу сравниваемых объектов  $n$ , то есть  $\lambda_{\max} = n$ . В качестве меры несогласованности мнений экспертов рассматривается нормированное отклонение  $\lambda_{\max}$  от  $n$ , называемое индексом согласованности (ИС):

$$\text{ИС} = \frac{\lambda_{\max} - n}{n - 1}.$$

Получив индекс согласованности и выбрав по таблицам случайный индекс для заданного порядка матрицы, определяется отношение согласованности (ОС)

$$\text{ОС} = \text{ИС} / \text{СИ}.$$

Если величина  $\text{ОС} \leq 0,1$ , то степень согласованности экспертных данных считается приемлемой. В противном случае (если  $\text{ОС} > 0,1$ ) эксперту рекомендуется пересмотреть свои суждения. Для этого необходимо выявить те позиции в матрице суждений, которые вносят максимальный вклад в величину отношения согласованности, и попытаться изменить меру несогласованности в меньшую сторону на основе более глубокого анализа вопроса.

### ВЫВОДЫ

Разработаны математические модели предметной области проектирования СППР охраняемого объекта и инструментальные средства СППР по ликвидации нештатных ситуаций.

Усовершенствованы многоальтернативные процедуры принятия решений с оцениванием человеческого и

технико-экономических факторов в принятии решений по ликвидации нештатных ситуаций.

Разработана математическая модель и вычислительная процедура выбора рационального состава комплекса технических средств СППР.

Разработаны информационное и программное обеспечение процесса поддержки принятия решений оператором по ликвидации нештатных ситуаций.

### ПЕРЕЧЕНЬ ССЫЛОК

1. Саати Т. Принятие решений. Метод анализа иерархий / Т. Саати. – М.: Радио и связь, 1989. – 316 с.
2. Системи підтримки прийняття рішень: навчальний посібник / О. І. Пушкар, В. М. Гіковатий, О. С. Євсєєв, Л. В. Потрашкова / За ред. д-ра екон. наук, проф. Пушкаря О. І. – Х.: «ІНЖЕК», 2006. – 304 с.
3. Ерохин А. Л. Система поддержки принятия решений при авариях в энергосистемах / А. Л. Ерохин // Проблемы бионики. – 1999. – Вып. 50. – С. 157–161.
4. Гриб О. Г. Система поддержки принятия решения при аварийных ситуациях в энергосистемах / О. Г. Гриб, О. Н. Довгалюк, А. Л. Ерохин // Світлотехніка та електроенергетика. – 2008. – № 4. – С. 64–68.

Надійшла 11.05.2009

*Розроблено структура, склад і інформаційні технології підтримки ухвалення рішень по ліквідації нештатних і аварійних ситуацій, які контролюються охоронно-пожежною сигналізацією. Приведено приклад вибору комплексу апаратно-програмних засобів за сукупністю локальних критеріїв, якими виступають техніко-економічні показники ефективності ухвалення рішень по ліквідації нештатних ситуацій.*

*There have been developed a structure, body and information technologies for decision-making support in terms of handling off-nominal situations being controlled by fire and security alarm. There is given an example of choosing a complex of physical and programme means basing on the set of criteria, the latter being technical and economic indicators of decision-making efficiency in terms of handling off-nominal situations.*

УДК 004.056

А. Е. Архипов

## ЭКСПЕРТНО-АНАЛИТИЧЕСКОЕ ОЦЕНИВАНИЕ ИНФОРМАЦИОННЫХ РИСКОВ И УРОВНЯ ЭФФЕКТИВНОСТИ СИСТЕМЫ ЗАЩИТЫ ИНФОРМАЦИИ

---

*Рассматриваются некоторые особенности применения методологии информационных рисков для исследования защищенности информационных систем, в частности, практические подходы к оцениванию количественных показателей, используемых для вычисления информационных рисков, а также показателей защищенности информационных систем и эффективности систем защиты информации.*

© Архипов А. Е., 2009

### ВВЕДЕНИЕ

Действующие международные стандарты [1–3] настойчиво рекомендуют методологию оценивания и управления информационными рисками как действенный инструмент исследования угроз безопасности информационных систем (ИС), оценивания уровня

защищенности этих систем, анализа эффективности функционирования систем защиты информации. К сожалению, высокий потенциал методологии информационных рисков на практике в значительной мере ограничивается необходимостью использования в ней плохо поддающихся определению в условиях практики количественных значений вероятностей реализации тех или иных угроз информации и количественных оценок ущерба (потерь), обусловленных успешной реализацией этих угроз. В максимальной степени это касается угроз несанкционированного доступа (НСД), связанных с наличием плохо прогнозируемого человеческого фактора.

### ПОСТАНОВКА ЗАДАЧИ

Риск  $R(T)$ , определяющий вероятные потери  $q(T)$ , обусловленные реализацией угрозы  $T$ , определяется выражением

$$R(T)=P(T)q(T), \quad (1)$$

где  $P(T)$  – вероятность реализации угрозы  $T$ . Для множества угроз  $\{T_i\}$ ,  $i = \overline{1, N}$  количественные значения  $P(T_i)$ ,  $q(T_i)$  задаются, как правило, путем их прямого экспертного оценивания, что приводит к высокому уровню субъективной ошибки в получаемых оценках параметров. Поэтому в ряде случаев оправданным оказывается применение качественных оценок вероятности ущерба. При этом в обоих случаях уровень погрешностей в итоговой оценке риска резко снижает эффективность и результативность анализа рисков.

Исправить эту ситуацию можно применением экспертно-аналитического подхода к получению оценок  $P(T_i)$ ,  $q(T_i)$ ,  $i = \overline{1, N}$ , суть которого во введении определенной детализации и структуризации процедур нахождения экспертных оценок соответствующих параметров рисков.

### СЦЕНАРНАЯ МЕТОДОЛОГИЯ НАХОЖДЕНИЯ ОЦЕНКИ УЩЕРБА, ОБУСЛОВЛЕННОГО РЕАЛИЗАЦИЕЙ УГРОЗ ИНФОРМАЦИИ

В [4] для оценивания уровня ущерба  $q$  предложено использовать метод сценариев, базирующийся на рассмотрении и интегральном учете последствий ряда возможных сценариев развития событий, обусловленных реализацией угрозы  $T_i$ . Сценарии образуют множество  $\{Sc_l\}$ ,  $l = \overline{1, k}$ , каждому элементу которого ставится в соответствие двойка  $\langle P_l, Q_l \rangle$ , где  $P_l$  – вероятность осуществления  $l$ -го сценария, а  $Q_l$  – характеристика ущерба (потерь), причиняемого ИС либо

информационным ресурсам по завершению развития соответствующего сценария.

Проблемным аспектом оценивания интегрального ущерба в схеме сценариев является построение способа свертки частных потерь  $Q_l$ ,  $l = \overline{1, k}$ , в результирующее значение  $q$ . Обычная свертка вида  $\sum P_l Q_l$  вследствие возможности одновременного параллельного развития ряда сценариев приводит к завышенной оценке интегрального ущерба [4, 5]. Поэтому в общем случае для исходного ансамбля возможных сценариев  $\{Sc_1, Sc_2, \dots, Sc_k\}$  конструируется полная группа возможных исходов  $\{Z_1, Z_2, \dots, Z_m\}$ ,  $m = 2^k$  с указанием их вероятностей  $P(Z_j)$ ,  $j = \overline{1, m}$  и частных ущербов  $Q(Z_j)$ . Так, для ансамбля из трех сценариев  $\{Sc_1, Sc_2, Sc_3\}$  с соответствующим набором элементарных исходов  $\{V_1, V_2, V_3\}$  получаем полную группу событий  $\{Z_1, Z_2, \dots, Z_8\}$ , где  $Z_1 = V_1 \cap V_2 \cap V_3$ ,  $Z_2 = V_1 \cap V_2 \cap \bar{V}_3$ , ...,  $Z_j = \bar{V}_1 \cap V_2 \cap \bar{V}_3$ , ...,  $Z_8 = \bar{V}_1 \cap \bar{V}_2 \cap \bar{V}_3$ .

События  $Z_1, \dots, Z_8$ , составленные из исходов  $V_1, V_2, V_3$  и противоположных им элементов  $\bar{V}_1, \bar{V}_2, \bar{V}_3$ , является попарно несовместимыми, а для их вероятностей справедливы соотношения:  $\sum_{i=1}^m P(Z_i) = 1$ ,  $P(Z_j + Z_g) = P(Z_j) + P(Z_g)$  и  $P(Z_j Z_g) = 0$  при  $j \neq g$ .

В соответствии со структурой событий  $Z_j$ ,  $j = \overline{1, 8}$  рассчитываются их вероятности и частные ущербы [4]:  $P(Z_1) = P_1 P_2 P_3$ ,  $P(Z_2) = P_1 P_2 (1 - P_3)$ , ...,  $P(Z_j) = (1 - P_1) P_2 (1 - P_3)$ , ...,  $P(Z_8) = \prod_{i=1}^3 (1 - P_i)$ ,  $Q(Z_1) = Q_1 + Q_2 + Q_3$ ,  $Q(Z_2) = Q_1 + Q_2$ , ...,  $Q(Z_j) = Q_2$ , ...,  $Q(Z_8) = 0$ .

Для расчета интегрального ущерба  $q$  после перехода к полной группе событий применение свертки вполне корректно:

$$q = \sum_{j=1}^m P(Z_j) Q(Z_j) = \sum_{j=1}^{m-1} P(Z_j) Q(Z_j). \quad (2)$$

В принципе возможна более глубокая проработка сценариев, ведущая к их детализации и представлению в виде графа взаимосвязанных промежуточных сцен с иерархической (древовидной) структурой [4], что позволяет говорить об использовании системного подхода в реализации основных положений изложенной методики нахождения интегральной оценки потерь  $q$ . Конкретизация источников происхождения потерь  $Q(Z_j)$  в реальной ситуации позволит определить процессы образования потерь (ущербов), более четко проанализировать характер и количественные параметры потерь в конкретных предметно-прикладных аспектах, в частности, получить точные и объективные экспертные оценки  $P(Z_j)$  и  $Q(Z_j)$ .



### ОЦЕНИВАНИЕ ВЕРОЯТНОСТНЫХ ПАРАМЕТРОВ МОДЕЛИ УГРОЗ

Для нахождения вероятности  $P(T)$  применение прямой экспертной оценки также не дает приемлемых результатов, при этом уровень субъективных ошибок оценивания оказывается зависимым от подбора экспертов. Эксперты, являющиеся сотрудниками организации, информация которой подлежит защите, т. е. «внутренние эксперты», достаточно хорошо осознают уязвимости системы защиты и защищаемых информационных ресурсов, что и находит свое отражение в их субъективных оценках вероятности  $P(T)$ . Оценки же «внешних» экспертов в большей степени учитывают специфические потребительские свойства защищаемого ресурса, мотивирующего злоумышленника к попыткам овладеть этим информационным ресурсом, уничтожить его либо исказить его содержание. В этом случае перспективным оказывается применение так называемой двухфакторной модели оценки вероятности  $P(T)$ , позволяющей в общей вероятности реализации угрозы выделить два компонента (фактора), один из которых отображает мотивационную составляющую возникновения угрозы, а второй учитывает существующие уязвимости. В итоге имеем [5]:

$$P(T) = p(T)p(a_T), \quad (3)$$

где  $p(T)$  – вероятность возникновения (актуализации) угрозы  $T$ ,  $p(a_T)$  – вероятность наличия совокупности уязвимостей, позволяющих реализовать угрозу  $T$ . По своей сути  $p(a_T)$  – обобщенная вероятность успешного проведения комплекса атак, обусловленных (порождаемых) упомянутой совокупностью уязвимостей защищаемой ИС (включая уязвимости самой системы защиты информации (СЗИ)).

В частности, для оценивания вероятности возникновения угрозы можно использовать предложенное в [6] соотношение, позволяющее в первом приближении полагать

$$p(T) \approx 1 - \frac{D}{d}, \quad (4)$$

где  $D$  – общая стоимость затрат атакующей стороны на реализацию угрозы  $T$ ,  $d$  – полученный при этом «выигрыш», определяемый ценностью защищаемого ресурса  $I$  для злоумышленников.

Ясно, что если ценность ресурса  $I$  для атакующей стороны очень высока, злоумышленники готовы идти на значительные затраты средств для реализации угрозы  $T$ . Поэтому в случае  $d \gg D$  вероятность  $p(T)$  будет практически равна 1. При малых значениях  $d$  мотивированность злоумышленников к реализации угрозы  $T$  низка, в частности при  $d = D$  теоретически  $p(T) = 0$ , а при  $d < D$  формула (4) теряет

смысл. На практике это означает, что вероятность применения для реализации угроз высокочастотных атак низка. Атаки, подготовка, организация и проведение которых сопряжены со значительными затратами, оправданы лишь в случае, если, например, информация  $I$  составляет государственную тайну, т. е. уровень ее значимости может быть чрезвычайно высок, поэтому даже для значительных  $D$  отношение  $D/d < 1$ . Кроме того, важным аспектом в анализе вероятности затратных атак является то, что их организация и проведение связаны со значительными финансовыми рисками, позволить которые себе могут далеко не многие фирмы или организации.

При грубом упрощенном оценивании вероятность  $p(T)$  может задаваться равной 0 или 1. В последнем случае вероятность реализации угрозы  $p(T)$  фактически оказывается тождественной вероятности успешного проведения комплекса атак  $p(a_T)$ .

При оценке вероятности  $p(a_T)$ , учитывая, что ее значение непосредственно связано с наличием и характером возможных уязвимостей ИС и СЗИ, приходим к необходимости детализации обстоятельств, связанных с возможностью реализации угроз, приводящей к некоторой иерархической многоуровневой структуре «источник угроз – уязвимости – атаки – угрозы – потери». При анализе реальной ИС этой структуре будет соответствовать направленный граф, позволяющий зафиксировать связи элементов, составляющих все перечисленные уровни, что, в частности, разрешит построить расчетные соотношения для вычисления вероятности  $P(T)$ . Для этого, выделив совокупность атак  $\{a_{si}\}$ ,  $s = \overline{1, K_i}$ , где  $K_i$  – объем множества атак, направленных на реализацию угрозы  $T_i$ , получаем выражение для расчета вероятности  $p(a_{Ti})$  через вероятности осуществления атак  $p(a_{si})$  [5]:

$$p(a_{Ti}) = 1 - \prod_{s=1}^{K_i} [1 - p(a_{si})]. \quad (5)$$

Вероятности атак  $p(a_{si})$   $s = \overline{1, K_i}$ , так же как и вероятности возникновения угроз  $p(T_i)$ ,  $i = \overline{1, N}$ , задаются путем прямого экспертного оценивания, обеспечивая учет особенностей и свойств реальной ИС. Аналогичным путем оцениваются вероятности осуществления остальных угроз.

Завершающим этапом анализа рисков является вычисление интегрального информационного риска  $R_t$ , обобщающего значения частных рисков, обусловленных реализациями тех или иных угроз в ИС. При расчете  $R_t$  возникают сложности, аналогичные уже рассмотренной выше проблеме свертки частных потерь  $Q_l$ ,  $l = \overline{1, k}$  в оценку интегрального ущерба  $q$ . В частности, группа угроз  $\{T_i\}$ ,  $i = \overline{1, N}$  не образует полной группы событий, достаточной для описания и количественной оценки уровня защищенности ИС [5].

Поэтому традиционно рассматриваемую тройку угроз конфиденциальности, доступности и целостности информации (соответственно  $T_1$ ,  $T_2$ ,  $T_3$ ), к реализации которых обычно стремятся свести имеющееся множество атак, следует преобразовать в полную группу из восьми комплексных угроз, содержащих все возможные сочетания исходных угроз:  $\{T_1T_2T_3\}$ ,  $\{T_1T_2\}$ ,  $\{T_1T_3\}$ ,  $\{T_2T_3\}$ ,  $\{T_1, T_2, T_3\}$ . Для формирования полной группы вводятся события, противоположные угрозам:  $\bar{T}_1$ ,  $\bar{T}_2$ ,  $\bar{T}_3$ , которые формально можно трактовать как события, соответствующие отсутствию факта актуализации (возникновения) соответствующей прямой угрозы конфиденциальности, доступности, целостности. Учитывая возможность существования комплексного события, состоящего в невозможности реализации какой-либо угрозы, получаем полную группу событий, соответствующих восьми комплексным угрозам:  $T_{r_1} = T_1 \cap T_2 \cap T_3$ ,  $T_{r_2} = T_1 \cap T_2 \cap \bar{T}_3$ , ...,  $T_{r_7} = \bar{T}_1 \cap T_2 \cap \bar{T}_3$ ,  $T_{r_8} = \bar{T}_1 \cap \bar{T}_2 \cap \bar{T}_3$ . Для каждого компонента этой полной группы, как это уже выполнялось выше для событий  $\{Z_1, \dots, Z_8\}$ , можно рассчитать соответствующие вероятности и комплексные ущербы, представимые множеством двоек  $\{P(T_{r_l}), Q(T_{r_l})\}$ ,  $l = \bar{1}, 8$ , после чего найти информационный риск ИС в целом. Например,  $P(T_{r_1}) = P(T_1) \times P(T_2)P(T_3)$ ,  $P(T_{r_2}) = P(T_1)P(T_2)[1 - P(T_3)]$ , ...,  $P(T_{r_7}) = [1 - P(T_1)]P(T_2)[1 - P(T_3)]$ , ...,  $P(T_{r_8}) = [1 - P(T_1)][1 - P(T_2)][1 - P(T_3)]$ . Далее, принимая гипотезу аддитивности ущерба, обусловленного реализацией комплексной атаки [5], получаем:  $q(T_{r_1}) = q_1 + q_2 + q_3$ ,  $q(T_{r_2}) = q_1 + q_2$ , ...,  $q(T_{r_7}) = q_2$ , ...,  $q(T_{r_8}) = 0$ , где  $q_1$ ,  $q_2$ ,  $q_3$  – ущербы, обусловленные реализациями соответствующих угроз  $T_1$ ,  $T_2$ ,  $T_3$ .

Окончательная формула для расчета информационного риска, обусловленного угрозами  $T_1$ ,  $T_2$ ,  $T_3$  имеет вид:

$$R_T = \sum_{i=1}^l P(T_{r_i})q(T_{r_i}). \quad (6)$$

### ОЦЕНИВАНИЕ ЗАЩИЩЕННОСТИ ИНФОРМАЦИОННЫХ СИСТЕМ

Осуществление в ИС каких-либо защитных мероприятий ведет к уменьшению вероятностей  $p(a_{si})$  успешного проведения комплекса атак, а, следовательно, к обновлению (уменьшению) значений  $P(T_{r_j})$  в формуле (6) и к новому расчетному значению риска  $R_{TO}$ , называемого в этом случае остаточным. Результативность проведенных защитных мероприятий определяется в первую очередь значением «снятого» риска  $\Delta R = R_T - R_{TO}$ , его приведенными значениями  $\Delta R/R_{TO}$  либо  $\Delta R/R_T = 1 - (R_{TO}/R_T)$ . Последнее является показателем защищенности ИС, характеризующимся удобной измерительной шкалой с диапазоном возможных значений от 1 (случай «абсолют-

ной защищенности,  $T_{T0} = 0$ ) до 0 (нулевая эффективность СЗИ,  $R_T = R_{T0}$ ,  $\Delta R = 0$ ).

Расширенный показатель эффективности системы защиты информации, учитывающий затраты  $C$  на создание и обслуживание системы защиты, имеет вид [7]:

$$w = (\Delta R - C)/C. \quad (7)$$

В соответствие с формулой (7) наиболее эффективной СЗИ будет та, которая обеспечивает максимум предотвращенного ущерба на единицу затрат, обусловленных построением СЗИ и ее обслуживанием в течение определенного промежутка времени.

### ВЫВОДЫ

Результативность применения методологии информационных рисков в исследовании защищенности информационных систем во многом определяется точностью используемых для вычисления информационных рисков количественных параметров, определяющих вероятности угроз и уровни возникающих при их реализации потерь. Оценивание этих параметров на практике сталкивается с серьезными трудностями, вынуждающими применять прямые экспертные методы задания искомых параметров, что нередко приводит к существенному снижению точности результатов исследований. Избежать этих негативных последствий можно, применив изложенные выше экспертно-аналитические методы оценивания количественных параметров, используемых при расчете рисков.

### ПЕРЕЧЕНЬ ССЫЛОК

1. ISO/IEC 2700:2005. Information Technology – Security techniques – Information security management systems – Requirements [Электронный ресурс] : Secretariat ISO/IEC JTC 1/SC 27, DIN Deutsches Institut für Normung Berlin, Germany. – Режим доступа : <http://www.jtc1sc27.din.de/en>. – Назва з екрана.
2. ISO/IEC 27002:2007, Information Technology – Security techniques – Code of practice for information security management [Электронный ресурс] : Secretariat ISO/IEC JTC 1/SC 27, DIN Deutsches Institut für Normung Berlin, Germany. – Режим доступа : <http://www.jtc1sc27.din.de/en>. – Назва з екрана.
3. ISO/IEC 27005, Information Technology – Security techniques – Information security risk management [Электронный ресурс] : Secretariat ISO/IEC JTC 1/SC 27, DIN Deutsches Institut für Normung Berlin, Germany. – Режим доступа : <http://www.jtc1sc27.din.de/en>. – Назва з екрана.
4. Архипов О. Є. Застосування методології передбачення для оцінювання шкоди, заподіяної витоком секретної інформації / О. Є. Архипов, І. П. Касперський // Правове, нормативне та метрологічне забезпечення системи захисту інформації в Україні. – Вип. 2 (15). – К. : 2007. – С. 13–19.
5. Архипов А. Е. Применение среднего риска для оценивания эффективности защиты информационных систем / А. Е. Архипов // Правове, нормативне та метрологічне забезпечення системи захисту інформації в Україні. – Вип. 1 (14). – К. : 2007. – С. 60–67.

6. Архипов А. Е. Применение мотивационно-стоимостных моделей для описания вероятностных соотношений в системе «анализ – защита» / А. Е. Архипов, С. А. Архипова // Правове, нормативне та метрологічне забезпечення системи захисту інформації в Україні. – Вип. 1 (16). – К. : 2008. – С. 57–61.
7. Архипов О. Є. Оцінювання ефективності системи охорони державної таємниці / О. Є. Архипов, І. Т. Бородавко, В. П. Ворожко. – К. : Наук.-вид. від-діл НА СБ України, 2007. – 63 с.

Надійшла 16.02.2009  
Після доробки 17.03.2009

*Розглянуто деякі особливості застосування методології інформаційних ризиків до дослідження захищеності інформаційних систем, зокрема практичні підходи до оцінювання кількісних показників, що використовуються для обчислення інформаційних ризиків, а також показників захищеності інформаційних систем та ефективності систем захисту інформації.*

*In this paper some features of information risk methodology application for information systems security research, in particular practical estimation approaches of the quantity indicators used for calculation information risk, and also parameters of information systems security and efficiency of information protection systems are considered.*

УДК 004.43

В. М. Крищук, О. Ю. Малий, О. Ю. Воропай

## УНІВЕРСАЛЬНА АЛГОРИТМІЧНА МОВА ПРОГРАМУВАННЯ МІКРОКОНТРОЛЕРІВ

*Досліджено існуючі методи створення алгоритмічних мов програмування. Розроблено універсальну мову програмування мікроконтролерів, що дозволяє переводити програми з одного мікроконтролера в інший, а також задавати загальний алгоритм програми при роботі з любим видом мікроконтролерів.*

### ВСТУП

Достатня кількість різних сімейств мікроконтролерів, а також відмінності між мікроконтролерами усередині одного сімейства вимагає розробки методів створення універсального засобу, що могло б дозволити робити налагодження пристроїв на основі мікроконтролерів будь-якого типу. Першим етапом на шляху створення такого засобу є розробка універсальної алгоритмічної мови програмування мікроконтролерів, що дозволить врахувати всі особливості кожного з обраних для роботи мікроконтролерів.

Для створення такої мови необхідно вивчити формальну граматику мов програмування, а також визначити метод універсального опису мікроконтролерів, що дозволив би не тільки описувати існуючі на сьогоднішній день мікроконтролери з усіма виконуваними ними функціями, але і давав можливість розширення достатку виконуваних мікроконтролерами функцій, для опису подальших апаратних розробок структур мікроконтролерів. Для цього пропонується врахувати особливості виконання однієї і тієї ж задачі на різних типах мікроконтролерів, включаючи й апаратну реалізацію задачі, що дозволить провести аналіз і знайти загальні риси роботи. Для чого необхідно проаналізувати особливості перетворення програми, написаної для одного типу мікроконтролера,

в програму для іншого типу чи підтипу усередині одного сімейства [1].

### АНАЛІЗ КОНСТРУКЦІЙ МОВ ПРОГРАМУВАННЯ

Реакція комп'ютера на дані, що вводяться, однозначна – першою справою занести їх до пам'яті, забезпечивши повну схоронність і чекати подальших команд.

Формально мова програмування – це відкрита безліч текстів, написаних за допомогою деякого набору символів – алфавіту мови. За основним своїм призначенням мова програмування – це засіб спілкування між користувачем і комп'ютерною системою.

Синтаксис мов програмування – сукупність вимог, яким повинна задовольняти будь-яка осмислена програма.

Для завдання синтаксичних правил найбільшою популярністю користується апарат форм Бакуса – Наура. Їхнє основне призначення визначити, які саме послідовності символів вважаються програмами в даній мові програмування. Це досягається вказівкою з яких складових частин і яким способом можуть бути побудовані програми [2].

Семантика мови програмування – це правила, що визначають, які операції і у якій послідовності повинна виконати машина, що працює по довільній заданій їй програмі.

Семантика мови програмування в цілому задається вказівкою:

1. Використовуваних у мові типів (тобто множин) простих значень, наприклад цілих і нецілих чисел.

2. Способів побудови складених значень, наприклад комплексних чисел, векторів і чи матриць їхніх аналогів з декількох простих.

3. Операцій над простими, а іноді і складеними значеннями й іншими діями, що відрізняються від операцій, які не виробляють нових числових чи інших значень (наприклад дія, що називається присвоюванням, полягає в тім, що за деяким обчисленням чи обраним значенням закріплюється деяке позначення).

4. Способів завдання складних дій, які складаються з ряду простих, що виконуються у визначеній послідовності.

Прагматика мови програмування – це методологія програмування, тобто опис принципів, методів і прийомів, що дозволяють виходячи з постановки задачі скласти програму її рішення.

У загальному випадку під реалізацією мови програмування на деякій комп'ютерній системі мається на увазі програма, що перекладає текст програм з однієї мови на машинну мову даної комп'ютерної системи [3].

Оператори у відмінності від виразів це конструкції призначені для організації дій більш складних ніж обчислення нових значень. Ці дії зводяться до декількох основних категорій: створити нове ім'я для подальшого використання в програмі (оголошення), обчислити значення деякого виразу і дати йому одне з передбачених у програмі імен, зберігши це значення для використання в майбутньому (присвоєння), установити послідовність, у якій повинні виконуватися інші дії (керування). До операторів керування відносяться: порожній оператор, послідовна композиція, умовна композиція, циклічна композиція, оператори переходу.

### **ЗАПРОПОНОВАНА РЕАЛІЗАЦІЯ ОПИСУ АПАРАТНИХ І ПРОГРАМНИХ ЗАСОБІВ В УНІВЕРСАЛЬНІЙ АЛГОРИТМІЧНІЙ МОВІ ПРОГРАМУВАННЯ МІКРОКОНТРОЛЕРІВ**

Як універсальну мову програмування мікроконтролерів пропонується застосувати мову, що складається з оператора оголошення, оператора присвоєння («=») і двох операторів керування («if», «goto»), це задовольняє умові апарата форм Бакуса – Наура та включає всі необхідні типи операторів.

Для опису апаратних засобів, що входять до складу мікроконтролерів, варто застосувати об'єктний підхід, при якому всі їхні особливості будуть описані як властивості об'єкта з визначеними параметрами унікальними для кожного окремого мікроконтролера, при тому, що об'єкти будуть створені універсальними, тобто придатними до будь-якого мікроконтролера. У виді об'єктів необхідно описувати як програмні так і апаратні частини мікроконтролера [4].

Синтаксис опису властивостей модулів у виді об'єктів застосуємо стандартний, котрий використовується у всіх макромовах програмування, тобто <об'єкт>. <дочірній об'єкт 1-го порядку> ... <властивість> = = значення властивості.

Властивості об'єкта описуються як змінні чи як масиви, при цьому одна і та сама властивість може бути просто змінною й одномірним масивом, одномірним і двомірним масивом і т. д. у тому випадку, коли описується його частина, наприклад один біт цілого байта. (PORT[1]=45h у той час як PORT[1,0]=1, тут одиниця вказує на загальну властивість тобто номер порту, 0 у другому випадку вказує номер біту (номер виводу), при цьому 45h=01000101b, отже 0-й біт=1), це відповідає аналогічному використанню спеціальних регістрів у мікроконтролерах (MOVWF PORTA, однак BSF PORTA, 2).

Розглянемо на прикладі мікроконтролера PIC16F84A опис деяких програмних і апаратних модулів.

Опишемо деякі з них у вигляді об'єктів із властивостями.

1. Об'єкт стек адресів функцій, які викликають процедури(8 level stack) має наступні властивості:

- розрядність 13 біт (Stack.bits=13);
- кількість рівнів (Stack.levels=8);
- робота над стеком виконується тільки командами виклику процедури і повернення з процедури.

Над ним можна робити наступні операції:

- додати дані (командою виклику процедури), при цьому всі попередні дані зміщуються на один рівень униз, а саме нижнє значення віддаляється;
- прочитати дані (командою повернення з процедури), при цьому об'єкт видає верхнє значення і потім стирає його, а інші значення зміщуються на один рівень нагору.

Приклади застосування:

```
.....
015A: CALL 02F1h; виклик процедури яка знаходиться за адресою 02F1h
(Stack.level[1]=PC=015Ah=00000001 01011010b;
For i=8 downto 2 do Stack.level[i]=Stack.level[i-1])
.....
02F1: RETURN; повернення з процедури
(PC=Stack.level[1]; for i=1 to 7 do Stack.level[i] =
=Stack.level[i+1])
.....
```

2. Об'єкт оперативна пам'ять (RAM File registers) має наступні властивості:

- кількість осередків (RAM.bytes=68);
- розрядність (RAM.bits=8);
- діапазон адрес (RAM.beginaddr=0Ch, RAM.endaddr=4Fh);
- має 2 банки пам'яті с різницею адресації 80h (RAM.banks=2, RAM.incbanks=80h).

Над ним можна робити наступні операції:

- записати в обраний осередок дані;

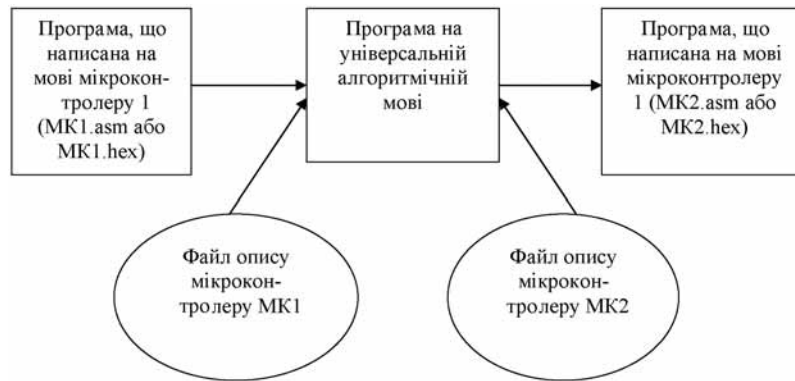


Рисунок 1 – Структурна схема процесу перекладу програми з одного мікроконтролера в інший

– вважати з обраного осередку дані.

Приклади застосування:

MOVLW CFh (w=CFh)

MOVWF 12h (RAM.byte[12]=w)

.....

MOVF 4Fh,w (w=RAM.byte[4F])

При цьому будь-яка виконувана команда змінює які-небудь властивості одного чи декількох об'єктів (змінює вміст масиву пам'яті, значення вектора виводів порту, виводів компаратора і т. д.). Усі команди мікроконтролера зводяться до команд присвоєння («=») і команд перевірки («if») з використанням переходів («goto») універсальної алгоритмічної мови програмування.

### **ВИЗНАЧЕННЯ НЕОБХІДНОЇ КІЛЬКОСТІ ОБ'ЄКТІВ ТА ЇХ ВЛАСТИВОСТЕЙ ШЛЯХОМ ПЕРЕКЛАДУ ПРОГРАМИ З ОДНОГО МІКРОКОНТРОЛЕРА В ІНШИЙ**

Для визначення необхідної кількості об'єктів того чи іншого мікроконтролера та особливостей опису необхідно визначити множини загальних та унікальних для кожного окремого мікроконтролера об'єктів. Визначення множини загальних об'єктів пропонується зробити шляхом перекладу програм одного мікроконтролера в інший, аналіз цих перетворень дозволить визначити загальні властивості мікроконтролерів та їх відмінності.

Для трансляції програм необхідно створити універсальну (загальну для всіх мікроконтролерів) середню частину перекладу тобто запропоновану універсальну мову, що буде враховувати особливості кожного з мікроконтролерів і приводити їх у деякі загальні структури поруч з приватними властивостями, що описують унікальність (рис. 1). При цьому на МК2 особливості МК1 реалізовувати емуляцією за допомогою наявних структур і їхніх властивостей [4].

### **ОСОБЛИВОСТІ ОПИСУ ДЕЯКИХ ЗАГАЛЬНИХ ОБ'ЄКТІВ МІКРОКОНТРОЛЕРІВ УНІВЕРСАЛЬНОЮ АЛГОРИТМІЧНОЮ МОВОЮ**

Проведений аналіз мов програмування мікроконтролерів дав можливість зробити наступні висновки щодо особливостей опису об'єктів:

1. Робота з виводами мікроконтролера повинна зводитися не до кількості портів визначеної розрядності, а описуватися кількістю необхідних висновків з визначеними властивостями.

2. Робота з будь-яким видом пам'яті мікроконтролера повинна вказувати необхідний обсяг використовуваної пам'яті, а не максимальний обсяг пам'яті мікроконтролера.

3. Настроювання мультіпліцированих пристроїв повинні вироблятися автоматично – якщо у визначений момент часу використовуються один пристрій, всі інші (мультіпліцированні з даним) використовуватися не можуть.

4. Використання спеціальних регістрів і регістрів загального призначення не повинне прив'язуватися до визначених адрес у пам'яті, а необхідно тільки вказувати функції регістра.

### **ПРИКЛАД ПЕРЕТВОРЕННЯ ПРОГРАМИ ОДНОГО МІКРОКОНТРОЛЕРА В ПРОГРАМУ ІНШОГО**

Як приклад, за допомогою якого буде надалі зроблений висновок про методи створення універсальної алгоритмічної мови програмування, приводиться перетворення програми мікроконтролера PIC16F84A у програму мікроконтролера PIC12F629. Даний приклад поки не враховує різницю у вбудованих апаратних пристроях, а працює тільки з пристроями, що входять до складу як одного, так і іншого мікроконтролера з метою показати на даному етапі особливості програмного перетворення без програмної

Унікальний серійний номер ключа (32 байти)
Рівень доступу (2 байти)
Бітовий лічильник кількості разів доступу (30 байт)

Рисунок 2 – Карта пам'яті електронного ключа

й апаратної емуляції не вхідних до складу заданих мікроконтролерів пристроїв.

Задана задача наступного змісту:

Розробляється електронний ключ з обмеженою кількістю разів доступу, а також з різними рівнями доступу.

Мікроконтролер у пам'яті EEPROM містить деяку інформацію обсягом 64 байта служить для авторитизації. Серед її 32 байта утримуючих серійний номер ключа (унікальний для кожного), 2 байти рівень доступу і 30 байт лічильник кількості разів доступу (рис. 2). При цьому пристрій, що зчитує, може виконувати тільки 3 команди: читання біта, запис біта (скидання з одиниці в нуль) і встановлення початку області авторитизації. Дані команди виконуються шляхом сполучення двох керуючих сигналів RST і CLK. Читанню біта відповідає високий CLK при низькому RST, запису біта відповідає послідовність високого RST і високого CLK, а встановленню початку області авторитизації – високий CLK при високому RST (рис. 3).

Як видно з особливості задачі лічильник кількості разів доступу побітовий, тобто кількість встановлених біт у лічильнику відповідає дозволеним кількості разів доступу.

Як видно з особливості задачі, лічильник кількості разів доступу побітовий, тобто кількість встановлених біт у лічильнику відповідає дозволеним кількості разів доступу.

Спочатку вирішимо дану задачу на PIC16F84A.

Переходячи до безпосередньої реалізації цієї задачі, необхідно врахувати ту особливість, що порти в

заданих для рішення мікроконтролерах здатні видавати в порт інформацію тільки цілим байтом (на кожен вивід по біту). Ця особливість зважається шляхом бітового зрушення вліво чи вправо байта, при цьому в біті carry спеціального регістра status залишається 1 біт, саме він і буде подаватися в порт. Для сприйняття команд пристрою, що зчитує, використовуємо 2 виводи порту А, один із яких візьмемо тактовий (T1CK0 для підрахунку кількості CLK за допомогою регістра TMR0), другий беремо довільно з тих виводів із входом. Також необхідно взяти один вивід для даних які будуть виходити з ключа в пристрій, що зчитує, (вибираємо будь-який двонаправлений). Після цього пишуться процедури обробки подій, що забезпечують правильну роботу ключа по заданих умовах.

Якщо переходити до перекладу цієї програми в мікроконтролер PIC12F629 варто помітити наступні розходження в характеристиках:

- у 12F629 тільки один шестирозрядний порт вводу/виводу, для даної задачі це не істотно, але може значно вплинути на рішення інших задач, у яких необхідна більша кількість виводів;

- 12F629 має в 2 рази більше пам'яті EEPROM, що в даній задачі вносить необхідність контролю за виходом з першої половини адресації EEPROM пам'яті;

- у 12F629 виводи порту мультипліцированні з внутрішніми пристроями і тому навіть не дивлячись на те, що дані пристрої в задачі використовуватися не будуть, необхідне їх відповідне налаштування;

- 12F629 має відмінну від 16F84A карту пам'яті, у якій спеціальні регістри використовувани в програмі знаходяться по інших адресах, що особливо необхідно врахувати при перекладі програми. Регістри загального призначення також знаходяться в іншому адресному діапазоні, це потрібно враховувати з метою запобігання використання спеціальних регістрів і комірок пам'яті що не можна використовувати.

Далі приведений фрагмент коду програми на мікроконтролерах 16F84A, 12F629 і проміжний код на універсальній алгоритмічній мові.

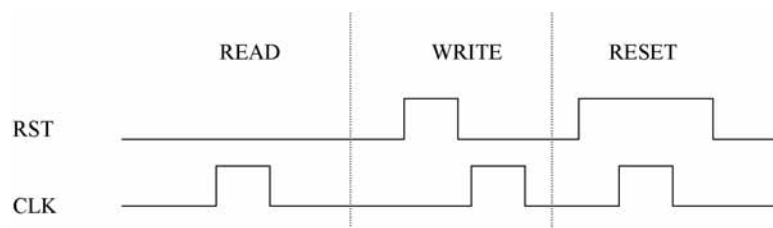


Рисунок 3 – Амплітудно-тимчасова діаграма команд пристрою читання електронного ключа

PIC 16F84A

Універсальна алгоритмічна мова

0089: BTFSC PORTB , 00	if PORT[2,0]=0 then PC=PC+2 else PC=PC;
008A: GOTO 0089	
008B: BSF PORTA , 00	PORT[1,0]=1;PC=PC+1;
008C: MOVWF INTCON	INTCON=w;PC=PC+1;
008D: CLRF TMRO	TMRO=0;PC=PC+1;
008E: NOP	PC=PC+1;
008F: NOP	PC=PC+1;
0090: GOTO 0094	PC=PC+4;
0091: BTFSC PORTA , 04	if PORT[1,4]=0 then PC=PC+2 else PC=PC;
0092: GOTO 0091	
0093: MOVWF PORTA	PORT[1]=w;PC=PC+1;
0094: BTFSS TMRO , 00	if TMRO[0]=1 then PC=PC+2 else PC=PC;
0095: GOTO 0094	
0096: RLF EEDATA , f	EEDATA=EEDATA*2+EEDATA[7]; PC=PC+1;
0097: MOVF STATUS , W	w=STATUS;PC=PC+1;
0098: BTFSC PORTA , 04	if PORT[1,4]=0 then PC=PC+2 else PC=PC;
0099: GOTO 0098	

Універсальна алгоритмічна мова

PIC 12F629

if PORT[2,0]=0 then PC=PC+2 else PC=PC;	007F: BTFSC GPIO , 01
PORT[1,0]=1;PC=PC+1;	0080: GOTO 007F
INTCON=w;PC=PC+1;	0081: BSF GPIO , 00
TMRO=0;PC=PC+1;	0082: MOVWF INTCON
PC=PC+1;	0083: CLRF TMRO
PC=PC+1;	0084: NOP
PC=PC+4;	0085: NOP
if PORT[1,4]=0 then PC=PC+2 else PC=PC;	0086: GOTO 008A
PORT[1]=w;PC=PC+1;	0087: BTFSC GPIO , 02
if TMRO[0]=1 then PC=PC+2 else PC=PC;	0088: GOTO 0087
EEDATA=EEDATA*2+EEDATA[7];	0089: MOVWF GPIO
PC=PC+1;	008A: BTFSS TMRO , 00
w=STATUS;PC=PC+1;	008B: GOTO 008A
if PORT[1,4]=0 then PC=PC+2 else PC=PC;	008C: RLF EEDATA , f
	008D: MOVF STATUS , W
	008E: BTFSC GPIO , 02
	008F: GOTO 008E

На прикладі видно, що використовуються різні види портів, різні спеціальні реєстри опису портів, а також не збігаються адреси фрагментів коду, що не заважає програмі ідентично працювати (робота програми була перевірена практично, шляхом апаратної реалізації).

### ВИСНОВКИ

Методика опису програмних і апаратних модулів з допомогою об'єктів дозволить полегшити розуміння їхньої роботи за рахунок наочності взаємодії блоків системи в цілому.

Використання універсальної алгоритмічної мови написаного по приведених методах дозволить розкласти роботу мікроконтролера на елементарні взаємодії об'єктів, що дасть можливість здійснювати емуляцію роботи деяких апаратних модулів програмно і використовуючи наявні в наявності апаратні модулі.

Універсальна алгоритмічна мова була випробувана в програмному комплексі MC-CAD v0.1, у якому реалізоване дизасемблювання програм чи переклад асемблерної програми будь-якого типу мікроконтролерів в універсальну алгоритмічну мову з метою подальшої компіляції в асемблер чи безпосередньо у файл мікроконтролера, що виконується, іншого довольного типу. Комплекс містить у собі конструктор мікроконтролерів, за допомогою якого можна описати кожний з існуючих мікроконтролерів.

УДК 004.853+004.832

В. В. Литвин

## МУЛЬТИАГЕНТНІ СИСТЕМИ ПІДТРИМКИ ПРИЙНЯТТЯ РІШЕНЬ, ЩО БАЗУЮТЬСЯ НА ПРЕЦЕДЕНТАХ ТА ВИКОРИСТОВУЮТЬ АДАПТИВНІ ОНТОЛОГІЇ

*Досліджено методи побудови та функціонування мультиагентних систем підтримки прийняття рішень, що базуються на прецедентах та використовують адаптивні онтології, які входять у склад інтелектуальних агентів. Розроблено метрику для визначення відстані між прецедентом та поточною ситуацією на основі адаптивних онтологій.*

### АНАЛІЗ ДОСЛІДЖЕНЬ ТА ПОСТАНОВКА ПРОБЛЕМИ В ЗАГАЛЬНОМУ ВИДІ

Згідно до сучасного рівня розвитку інформаційних технологій, а саме систем підтримки прийняття рі-

© Литвин В. В., 2009

### ПЕРЕЛІК ПОСИЛАНЬ

1. *Негода Д. В.* Автоматизация проектирования симуляторов микропроцессоров и микроконтроллеров : дис. ... канд. техн. наук : 05.13.12 / Негода Дмитрий Викторович. – Ульяновск : РГБ, 2005. – 160 с.
2. *Джаныбаев К.* Язык спецификаций для верификации набора операций языков // Теория языков и автоматизация программирования : сб. науч. тр. / АН УССР, науч. совет по пробл. «Кибернетика», Ин-т Кибернетики им. В. М. Глушкова; [Ред-кол.: Е. Л. Ющенко]. – К. : ИК, 1986. – С. 19–24.
3. *Стеблянко В. Г.* Разработка входных языков кроссовых систем подготовки программ реального времени / Стеблянко В. Г., Марченко А. И., Подобед Л. Е. // Теория языков и автоматизация программирования : сб. науч. тр. / АН УССР, науч. совет по пробл. «Кибернетика», Ин-т Кибернетики им. В. М. Глушкова; [Ред-кол.: Е. Л. Ющенко]. – К. : ИК, 1986. – С. 41–48.
4. *Белых А. А.* Унификация архитектур однокристалльных микроконтроллеров и ее применение для разработки программного обеспечения встраиваемых систем : дис. ... канд. техн. наук : 05.13.15 / Белых Андрей Александрович. – М. : РГБ, 2006. – 179 с.

Надійшла 09.02.2009

*Исследованы существующие методы создания алгоритмических языков программирования. Разработан универсальный алгоритмический язык программирования микроконтроллеров, который позволяет переводить программы с одного микроконтроллера в другой, а также задавать общий алгоритм программы при работе с любым видом микроконтроллеров.*

*The existing methods of creation of algorithmic languages of programming are investigated. The universal programming language of microcontrollers that allows to translate the program from one microcontroller to another and also to set common algorithm of the program at work with any kind of microcontrollers is developed.*

шень, розрізняють два напрями розвитку систем логічного виведення базованих на знаннях [1]:

- системи логічного виведення, заснованого на правилах;
- системи логічного виведення, заснованого на прецедентах.

Практично всі ранні експертні системи моделювали хід ухвалення рішення експертом як дедуктивний процес з використанням логічного виведення, заснованого на правилах. Це означало, що в систему закладалася сукупність правил вигляду «якщо... то...», згідно до яких на підставі вхідних даних генерувався



той або інший висновок вирішення проблеми, що цікавив користувача. Вибрана модель була основою для створення експертних систем перших поколінь, які були достатньо зручні як для розробників, так і для користувачів-експертів. Проте з часом було усвідомлено, що дедуктивна модель моделює один з найбільш рідкісних підходів, який використовує експерт під час вирішення проблеми. Ідея виведення за правилами є привабливою, тому що вона ґрунтується на добре формалізованих поняттях, для яких існують наукові методи, що довели свою застосовність і які дозволяють отримати рішення, що не вимагають доведення.

Однак навколишній світ складний. Існує багато слабоформалізованих завдань, для яких, можливо, будуть знайдені рішення. Крім того, існує ряд завдань, для яких ніколи не буде знайдено формального рішення (судочинство, медицина). Актуальність проблеми обумовлена і чисельністю таких завдань, і практичною потребою знайти хоча б одне яке-небудь відповідне рішення там, де через відсутність строгого формалізованого методу не можна знайти всі або найоптимальніші рішення.

Насправді, замість того, щоб вирішувати кожну задачу, виходячи з первинних принципів, експерт часто аналізує ситуацію в цілому і згадує, які рішення ухвалювалися раніше в подібних ситуаціях. Потім він або безпосередньо використовує ці рішення, або, при необхідності, адаптує їх до обставин, що змінилися для конкретної проблеми.

Моделювання такого підходу до вирішення проблем, заснованого на досвіді минулих ситуацій, привело до появи технології логічного виведення, заснованого на прецедентах (англійською – Case-Based Reasoning, або CBR), і надалі – до створення програмних продуктів, що реалізують цю технологію.

У ряді ситуацій метод виведення за прецедентами має серйозні переваги в порівнянні з виведенням, що ґрунтується на правилах, і особливо ефективний, коли:

- основним джерелом знань про завдання є досвід, а не теорія;
- рішення не унікальні для конкретної ситуації і можуть бути використані в інших випадках;
- метою є не гарантоване вірне рішення, а краще з можливих.

Таким чином, виведення, засноване на прецедентах, є методом побудови інтелектуальних систем, які приймають рішення щодо даної проблеми або ситуації за наслідками пошуку аналогій, що зберігаються в базі прецедентів.

Мета роботи – дослідити процес функціонування мультіагентних систем, що базуються на прецедентах та використовують адаптивні онтології для підтримки прийняття рішень.

**Постановка задачі.** Розробити структуру мультіагентної системи, що складається із окремих інтелектуальних агентів, які містять множину прецедентів та адаптивну онтологію. Побудувати алгоритм функціонування такої системи. Ввести метрику для обчислення відстані між прецедентом та поточною ситуацією. На основі отриманої відстані приймати раціональні рішення.

### **ПРОЦЕС ПРИЙНЯТТЯ РІШЕНЬ НА ОСНОВІ ПРЕЦЕДЕНТІВ**

Системи виведення за прецедентами дають дуже добрі результати в найрізноманітніших завданнях, але мають ряд істотних недоліків.

По-перше, вони взагалі не створюють яких-небудь моделей або правил, які узагальнюють попередній досвід, – у виборі рішення вони ґрунтуються на всьому масиві доступних історичних даних, тому неможливо сказати, на основі якої конкретно інформації, системи виведення за прецедентами видають свої рішення.

Можна виділити дві основні проблеми, з якими зустрічаються подібні системи: пошук найбільш відповідних прецедентів і подальша адаптація знайденого рішення.

В основі всіх підходів до відбору прецедентів лежить той чи інший спосіб вимірювання ступеня близькості прецеденту і поточної ситуації. При таких вимірюваннях обчислюється чисельне значення деякої міри, що визначає склад множини прецедентів, які треба опрацювати для досягнення задовільної класифікації або прогнозування. Основним недоліком таких систем є відсутність єдиної методики вибору міри близькості.

Ще один недолік методу пов'язаний з конструкцією прецедентів і призначення ваг їх атрибутам, що зменшує застосовність таких систем в різних предметних областях (ПО).

Однак процес роботи таких систем добре вже відомий (див. рис. 1). Шукаються відстані між поточною ситуацією та наявними прецедентами. Вибирається той прецедент, для якого ця відстань є найменшою. Виконується чи пропонується до виконання те рішення, яке відповідає цьому вибраному прецеденту.

Щоб формально це описати, введемо такі позначення:

- $Pr = \{Pr_1, Pr_2, \dots, Pr_N\}$  – множина прецедентів;
- $R_i$  – рішення, яке приймається згідно з прецедентом  $Pr_i$ ;
- $S$  – поточна ситуація;
- $d_i(Pr_i, S)$  – відстань між  $i$ -м прецедентом та поточною ситуацією;
- $o$  – оцінка прийнятого рішення.

У більшості випадків методи пошуку прецедентів зводяться до індукції дерев рішень або до алгоритм-

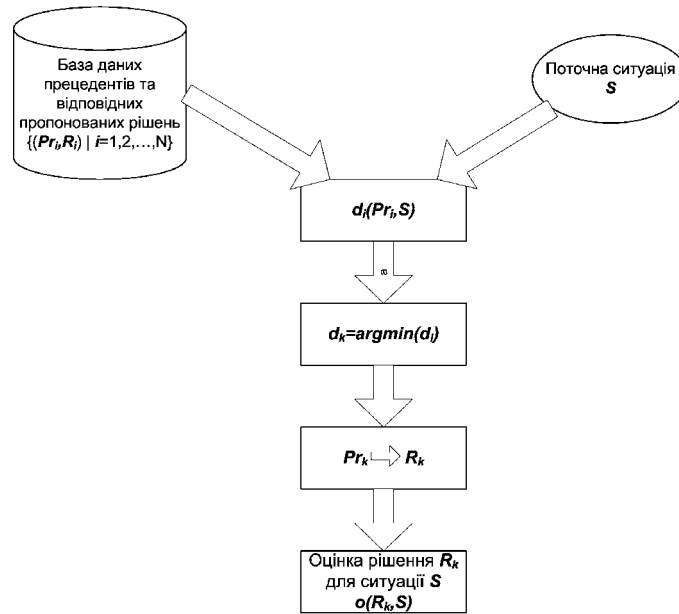


Рисунок 1 – Процес прийняття рішення на основі прецедентів

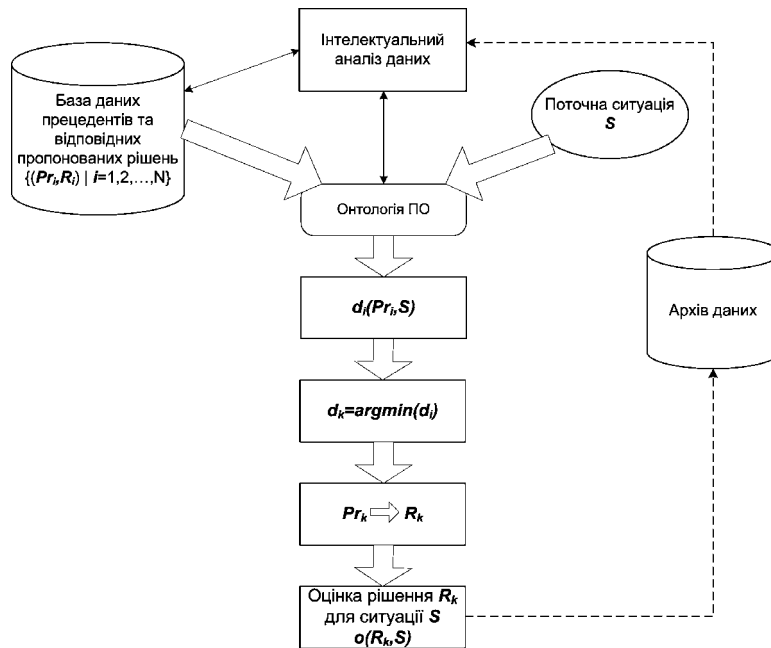


Рисунок 2 – Процес прийняття рішення на основі прецедентів з використанням онтологій

му «найближчого сусіда» [2], доповненого, можливо, використанням знань про предметну область (ПО). Що стосується адаптації і використання знайденого рішення, ця задача до цих пір залишається недостатньо формалізованою і дуже залежною від ПО.

Обидві проблеми – пошук прецедентів і адаптація вибраного рішення – вирішуються (повністю або частково) із залученням фонового знання, іншими сло-

вами, використовуючи знання про ПО (domain knowledge). Існують різні способи отримання інформації про ПО.

– Залучення експертного знання. Воно може виражатися, наприклад, в обмеженнях, що накладаються на діапазони змін ознак об’єктів, або ж у формулюванні множини правил для розбиття бази прецедентів на класи (побудова класифікатора).

– Отримання необхідних знань з набору наявних даних методами видобування знань (англійською – Data Mining). Сюди відносяться всі методи виявлення закономірностей між даними, зокрема, кластеризація, регресія, пошук асоціацій. Використання методів видобування даних може виділити вузьку групу показників, від яких залежить характеристика, що цікавить дослідника, і подати виявлену закономірність в аналітичній формі.

– Формування знань на основі навчальної вибірки, наданої експертом (навчання з учителем). Цей спосіб включає обидва перших.

Сама ж інформація про ПО відображається в її онтології [3]. Тому нами пропонується проектувати прецеденти та поточну ситуацію на онтологію ПО; ввести в межах онтології ПО метрику, за допомогою якої шукати необхідну відстань (див. рис. 2). Крім того, сама онтологія буде адаптуватися до ПО за рахунок методів інтелектуального аналізу даних (ІАД). Щоб ці методи використовувалися, нам необхідно вести історію функціонування інтелектуальної системи, тобто набутий нею досвід розв'язування подібних задач, який буде зберігатися в архіві даних. А саме у цьому архіві даних зберігаються прийняті системою рішення відповідно до ситуацій та оцінка цих рішень. Для процедур ІАД використовуються тільки ті дані, для яких оцінка прийнятих рішень є високою, тобто є вищою за деякий поріг. Проблема оцінювання прийнятих рішень у цій роботі не розглядається, оскільки ця проблема добре досліджена. Для потрібного шкалювання використовуються відомі математичні апарати, такі як теорія нечітких множин та відношень, лінгвістичні змінні тощо [4].

Відзначимо, що використання вищеописаного методу знаходження відстані, дозволяє шукати відстань не лише між прецедентами та поточною ситуацією, а також відстань між різними прецедентами чи різними ситуаціями.

Формальна модель онтології  $O$  визначається як впорядкована трійка такого вигляду:

$$O = \langle X, R, F \rangle,$$

де  $X$  – скінчена множина концептів (понять, термінів) предметної області, яку задає онтологія  $O$ ;  $R$  – скінчена множина відношення між концептами (поняттями, термінами) заданої предметної області;  $F$  – скінчена множина функцій інтерпретації (аксіоматизація), заданих на концептах або відношеннях онтології  $O$ .

### ПОНЯТТЯ АДАПТИВНОЇ ОНТОЛОГІЇ

Ефективність адаптації онтології бази знань до особливостей предметної області визначають закла-

дені в її структуру елементи та механізми її оптимізації (точніше, адаптації) шляхом самонавчання під час експлуатації. Одним з підходів до реалізації таких механізмів є автоматичне зважування понять бази знань (БЗ) та семантичних зв'язків між ними під час самонавчання. Цю роль беруть на себе коефіцієнти важливості понять та зв'язків. Їх розподіл у БЗ має відповідати таким основним вимогам:

- відображати семантичну вагу понять ПО, в якій ця інтелектуальна система реально застосовуватиметься;
- формуватися під час наповнення БЗ та коректуватися відповідно до визначених правил;
- забезпечувати контроль цілісності БЗ;
- відповідати вимогам метрики при їх використанні для порівняння семантичної близькості понять.

Стоїть завдання сформулювати відповідний набір правил присвоєння вагових коефіцієнтів (інформаційної ваги) поняттям та твердженням у моделі БЗ, що забезпечить оцінку актуальної цінності її інформаційного наповнення та досліджуваних поточних ситуацій (наприклад, віднесення текстових документів до класів згідно з УДК, що описано нижче).

Покажемо можливість вирішення сформульованого завдання шляхом введення деяких спрощень і припущень. Подамо базу знань у вигляді іменованого графа, числові семантичні характеристики вершин і ребер якого визначаються за певними правилами. Він є орієнтованим зваженим мультиграфом з такими властивостями:

- 1) на кожний елемент (вершину) може бути довільна кількість посилань;
- 2) кожний елемент може мати зв'язок з будь-якою кількістю інших елементів;
- 3) кожному зв'язку (ребру) у моделі відповідає певний напрям і коефіцієнт важливості зв'язку та достовірності відповідного твердження, кожному поняттю (вершині) – коефіцієнти важливості поняття.

Коефіцієнт важливості поняття (зв'язку) – це чисельна міра, котра характеризує значущість певного поняття (зв'язку) у конкретній предметній області, і динамічно змінюється за певними правилами у процесі експлуатації системи [5].

Наш підхід до подання знань у формі зваженої семантичної мережі (концептуальних графів) полягає у тому, що будь-яке можливе узагальнення, тобто комплексне, складене поняття завжди явним чином артикульоване, назване і як окремий концепт фігурує в базі знань. Тому, якщо деяке узагальнення має спільні властивості чи способи функціонування, вони фізично можуть бути реалізовані через властивості та опрацювання подій відповідного узагальнювального концепту.

Отже, ми розширимо формальне поняття онтології, ввівши в її формальний опис коефіцієнти важ-

ливості понять та відношень. Тому таку онтологію ми будемо визначати як п'ятірку:

$$O = \langle X, R, F, W, L \rangle,$$

де  $W$  – важливість понять  $X$ ,  $L$  – важливість відношень  $R$ .

Визначену таким чином онтологію будемо називати адаптивною, тобто такою, що адаптується до ПО за рахунок модифікації понять та коефіцієнтів важливості цих понять і зв'язків між ними.

Один із способів визначення цих коефіцієнтів важливості та їх зміна будуть описані нижче. Тут тільки зазначимо, що зміна цих коефіцієнтів відбувається відповідно до модифікації знань методами інтелектуального аналізу даних. Метою технології видобування даних є вироблення нового знання, яке користувач може надалі застосувати для поліпшення результатів своєї діяльності. Можна виділити, принаймні, сім методів виявлення і аналізу знань: класифікація; регресія; кластеризація; аналіз асоціацій; прогнозування тимчасових послідовностей (рядів); агрегація (узагальнення); виявлення відхилень.

Методи 1, 2 і 4 використовуються, головним чином, для прогнозування, тоді як останні зручні для опису наявних закономірностей у даних.

**МУЛЬТИАГЕНТНІ СИСТЕМИ ПРИЙНЯТТЯ РІШЕНЬ НА БАЗІ АДАПТИВНИХ ОНТОЛОГІЙ**

Приймати рішення щодо близькості (метрична оцінка) досліджуваної ситуації до прецедента буде інтелектуальний агент (ІА). Згідно з означенням: інтелектуальний агент – це агент, який володіє рядом знань про себе та навколишній світ, і поведінка якого визначається цими знаннями. Очевидно, що для нашого підходу такий ІА повинен складатися з бази даних (БД) прецедентів, архіву даних (тому історії функціонування системи загалом) та онтології ПО, за яку відповідає цей ІА (див. рис. 3).

Інтелектуальний агент здатний у процесі самонавчання адаптуватися до конкретних інформаційних

потреб користувача та виявляти, зберігати і використовувати релевантні до відповідних задач знання [6]. Інтелектуальний агент може бути реалізований в рамках продукційної, об'єктної, нейромережевої парадигми чи їх певного поєднання. Вибір розробника відштовхується від конкретних експлуатаційних вимог. При цьому доцільно максимально відкласти рішення щодо програмно-апаратної реалізації, пов'язані з обмеженнями у виборі операційної системи чи комп'ютерної архітектури. У будь-якому разі основою інтелектуального агента є його база знань, а ядром бази знань є онтологія [5, 7]. Загальні властивості ІА визначаються його онтологією, яка задає спосіб подання знань, механізми міркувань та прийняття рішень.

Оскільки онтологія містить коефіцієнти важливості понять та зв'язків, то очевидно, що ці коефіцієнти мають бути різними залежно від напряму діяльності агента. Для прикладу, така ознака як кашель має різну важливість з точки зору терапевта та окуліста, хоча обидва є лікарями. Тобто окремий напрям (відділ) задається деяким інтелектуальним агентом, а вся лікарня загалом являє собою мультиагентну систему.

Тому наша інтелектуальна система як ціле складається із кількох інтелектуальних агентів, тобто є мультиагентною. Згідно з означенням: мультиагентна система (МАС) – це множина деяких об'єктів (агентів), незалежних, але здатних створювати комунікативне співтовариство і спільно розв'язувати спільні задачі. Крім набору ІА, описаних вище, наша МАС буде містити агента, який вироблятиме нові знання, тим самим буде модифікувати онтології ПО відповідно до наявних концептів, зв'язків та коефіцієнтів їх важливості. Структура пропонованої МАС наведена на рис. 3.

Очевидно, що кожний прецедент однозначно задається онтологією, яка є частиною загальної онтології, і навпаки – онтологією однозначно визначається прецедент. Тобто між ними є однозначна відповідність, яку запишемо у вигляді:

$$Pr_i \Leftrightarrow O_i \subseteq O.$$

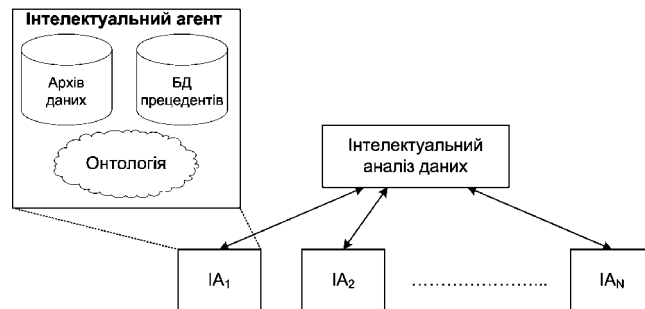


Рисунок 3 – Структура МАС прийняття рішень на основі прецедентів

Інтелектуальний агент на основі методів ІАД модифікує онтологію на основі старої онтології та архіву даних, який зберігає відповідний ІА. Тобто:

$$\text{ІАД: } \langle \mathbf{O}, \text{АД} \rangle \rightarrow \mathbf{O}.$$

Розглянемо методи задання початкових коефіцієнтів онтології ІА та їх модифікацію відповідно до рішень, отриманих методами ІАД, а саме методом побудови дерева рішень [8]. Нехай онтологія містить множину понять  $\mathbf{X} = \{x_1, x_2, \dots, x_n\}$  та множину зв'язків  $\mathbf{R} = \{r_1, r_2, \dots, r_m\}$ .  $\mathbf{W} = \{w_1, w_2, \dots, w_n\}$  – коефіцієнти важливості відповідних понять,  $\mathbf{L} = \{l_1, l_2, \dots, l_m\}$  – коефіцієнти важливості відповідних зв'язків. Спочатку вважаємо, що всі значення  $w_i = 1$ ,  $i = 1, 2, \dots, n$ ;  $l_j = 1$ ,  $j = 1, 2, \dots, m$ .

Нехай у результаті побудови дерева рішень ми отримали  $k$  різних правил  $\mathbf{Rule}_1, \mathbf{Rule}_2, \dots, \mathbf{Rule}_k$ . Кожне правило  $\mathbf{Rule}_s$  складається з підмножини понять, тобто  $\mathbf{Rule}_s = \{x_{s1}, x_{s2}, \dots, x_{st}\}$ . Додаємо до значення  $w_i$  одиницю, якщо поняття  $x_i$  входить у будь-яке правило  $\mathbf{Rule}_s$ ,  $s = 1, \dots, k$ . Аналогічно, якщо між поняттями  $x_i$  та  $x_j$  існує зв'язок  $l_p$ , то значення  $l_p$  збільшуємо на одиницю, якщо обидва поняття  $x_i$  та  $x_j$  входять в одне правило. Тим самим ми збільшуємо важливість понять, які входять у правила, та відповідних зв'язків. Це відповідає уявленню, що поняття (ознаки), які присутні у правилах, є важливішими, ніж інші значення.

### СПОСІБ ВИЗНАЧЕННЯ МЕТРИКИ

Пропонується визначати відстань між прецедентом і ситуацією, як суму відстаней між найважливішими поняттями прецедента та поточною ситуацією. Таких важливих понять може бути одне, два; однак якщо їх є більше-рівне трьох, то нами пропонується вибирати перші три. Ця кількість визначена на основі опитувань експертів різних ПО і вважається ними оптимальною. У такому випадку ми маємо 3 центри ваг  $i$ -го прецедента  $pr_i^1, pr_i^2, pr_i^3$  і 3 центри ваг поточної ситуації  $s^1, s^2, s^3$ . Тоді існує 9 різних відстаней  $d(pr_i^j, s^k)$ ,  $j = 1, 2, 3$ ;  $k = 1, 2, 3$ . Вибираємо 3 найменші з них та їх сумуємо. Отримана таким чином сума й буде відстанню між прецедентом та поточною ситуацією. Очевидно, що визначена таким чином відстань залежатиме від того як ми визначимо відстань між двома суміжними вершинами. Для цього пропонується визначати відстані між вершинами, що з'єднані зв'язком як

$$d_{ij} = \frac{Q}{L_{ij}(W_i + W_j)},$$

де  $W_i$  та  $W_j$  – коефіцієнти важливості вершин  $C_i$  та  $C_j$  відповідно;  $L_{ij}$  – коефіцієнт важливості зв'язку між вершинами;  $Q$  – константа, яка залежить від конкретної онтології. Прийmemo, що  $L_{ii} = \infty$ , тоді  $d_{ii} = 0$ .

Далі знаходимо центри ваг концептуального графа. Це перші три вершини для яких середня відстань  $\bar{d}_i$  є найменшою:

$$\bar{d}_i^* = \min_i \bar{d}_i.$$

Середня відстань  $\bar{d}_i$  для вершини  $C_i$  обчислюється згідно формули

$$\bar{d}_i = \frac{\sum_{j=1, j \neq i} d_{ij}^*}{n-1},$$

де  $n$  – кількість вершин графа,  $d_{ij}^*$  – найкоротший шлях між вершинами  $C_i$  та  $C_j$ , який обчислюється за допомогою відомих алгоритмів, наприклад Форда, Дейкстри, Флойда – Уоршалла [9].

Далі згідно концептуального графа, що задає онтологію прецедента шукаємо відстань від даного прецедента до поточної ситуації. Якщо поняття поточної ситуації не входять в концептуальний граф, то онтологію даного прецедента доповнюємо онтологією всього ІА до якого входить цей прецедент. Якщо ж необхідне поняття далі не входить в онтологію ІА, то його відсутність зумовлює ріст відстані до безмежності, що означає не близькість прецедента із поточною ситуацією.

Зазначимо, що запропонована таким чином відстань задовольняє трьом аксіомам метрики.

Дійсно, згідно визначення відстані, автоматично виконуються дві перші аксіоми:

$$d(C_i, C_i) = 0; \quad d(C_i, C_j) = d(C_j, C_i).$$

Нехай  $R_{ij}^*$  – шлях між вершинами  $C_i$  та  $C_j$ , який відповідає відстані між ними. Тоді  $d_{ij} = d_{ik} + d_{kj}$ , якщо вершина  $C_k$  лежить на шляху  $R_{ij}^*$  і  $d_{ij} < d_{ik} + d_{kj}$ , якщо вершина  $C_k$  не лежить на шляху  $R_{ij}^*$ . А це означає, що виконується третя аксіома метрики.

Відкритим залишається питання, як часто використовувати методи ІАД для модифікації онтології ІА. Тобто агент ІАД виступає ще в ролі координатора, а питання моментів втручання є задачею координації [10]. У цій роботі це питання не розглядатиметься.

Розглянемо запропонований метод на прикладі класифікації текстового документу, тобто його віднесення до відповідного УДК. Набір УДК утворює дерево, ми будемо класифікувати документи за найнижчими рівнями УДК, тобто найнижчий рівень являє собою прецедент. Скільки різних прецедентів буде входить в ІА залежить від того як однаково є важ-

ливими поняття, тобто незмінним є коефіцієнт їх важливості та відповідних зв'язків.

Онтологію ІА в яку входять кілька УДК представимо у вигляді зважених концептуальних графів. Для кожного прецедента (найнижчого УДК) визначимо 3 центри їх інформаційної ваги, які використовуватимемо для обчислення відстаней між прецедентом та поточним текстовим документом (ТД), який ми класифікуємо. Цей ТД ми теж задаємо у вигляді концептуального графа з відповідними коефіцієнтами важливості, які визначаємо як статистична кількість входжень понять та зв'язків у текст. Для цього ТД теж визначаємо 3 центри ваг. Знаходимо відстань між текстом та прецедентами як суму трьох найкоротших відстаней серед 9 можливих. Ідеальний варіант, коли центри ваг прецедента та поточного ТД співпадають. Тоді відстань дорівнюватиме нулю.

Значимо, що таким методом можна не тільки класифікувати текстові документи, а й знаходити відстань між будь-якими двома текстовими документами. Це у свою чергу дозволяє отримати деяку оцінку подібності двох електронних текстових документів, згідно до якої можна впорядкувати ці документи, якщо деякий вважати взірцевим. Тим самим будувати пошукові системи, функціонування яких базується на онтологіях ПО, що дозволяє здійснювати пошук не лише за ключовими словами, але й за контекстом. Приклад такого порівняння описано в [11].

## ВИСНОВКИ

Запропонований метод логічного виведення, заснованого на прецедентах дозволяє розв'язувати задачу класифікації та прийняття відповідного до прецедента рішення. Метод ґрунтується на знаходженні відстані між поточною ситуацією та прецедентами і знаходженні мінімальної відстані. У свою чергу для знаходження відстані використовується онтологія предметної області, яка визначається як п'ятірка елементів, а не як трійка як прийнято згідно стандартного означення. Нами запропоновано враховувати важливість понять та зв'язків між ними, а таку онтологію називати адаптивною. Для правильного функціонування інтелектуальної системи коефіцієнти важливості понять та зв'язків модифікуються на основі досвіду набутого системою, використовуючи методи інтелектуального аналізу даних. Ці коефіцієнти використовуються для обчислення відстані між прецедентом та поточною ситуацією. Показано, що визначена таким чином відстань задовольняє трьома аксіомам метрики.

Даний метод можна використати до класифікації текстових документів згідно УДК або співставляти будь-які два текстові документи на їх подібність. Під час обчислення відстані між текстовими документами

враховується контекст документів і відповідна до контексту семантика вжитих у них термінів та словосполучень. Це дає можливість здійснювати автоматичний пошук документів, котрі найбільше відповідають запиту-прототипу, і відкидати такі, що мають малу вагу і не відповідають предметній області. За допомогою розробленого методу можна здійснювати автоматичний пошук документів, котрі найбільше відповідають запиту-прототипу в мережі Інтернет, а також здійснювати його класифікацію та ранжування за релевантністю до заданого взірця.

## ПЕРЕЛІК ПОСИЛАНЬ

1. Каменнова М. С. Корпоративные информационные системы: технологии и решения / М. С. Каменнова // Системы управления базами данных. – № 3. – 1995. – С. 88–99.
2. Дюк В. А. Data Mining / В. А. Дюк, А. П. Самойленко. – СПб : Питер, 2001. – 368 с.
3. Montes-y-Gomez M. Comparison of Conceptual Graphs / M. Montes-y-Gomez, A. Gelbukh, A. Lypez-Lopez // Mexican International Conference on Artificial Intelligence MICA 2000. – Acapulco, Mexico, April 2000. – Lecture Notes in Artificial Intelligence № 1793. – Springer-Verlag, 2000. – P. 34–39.
4. Аверкин А. Н. Нечеткие множества в моделях управления и искусственного интеллекта / А. Н. Аверкин, И. З. Батыршин, А. Ф. Блишун. – М. : Наука, 1986. – 206 с.
5. Даревич Р. Р. Метод автоматического назначения информационной ваги понять в онтологии бази знань / Р. Р. Даревич, Д. Г. Досин, В. В. Литвин // Відбір та обробка інформації. – 2005. – Вип. 22(98). – С. 105–111.
6. Рассел С. Искусственный интеллект / С. Рассел, П. Норвиг. – М., С.-П., К. : Вильямс, 2006. – 1408 с.
7. Montes-y-Gomez M. Flexible Comparison of Conceptual Graphs / M. Montes-y-Gomez, A. Gelbukh, A. Lypez-Lopez, R. Baeza-Yates // 12th International Conference on Database and Expert Systems Applications DEXA 2001. – Munich, Germany, September, 2001. – Vol. 2113. – Springer-Verlag, 2001. – P. 56–62.
8. Цветков А. М. Разработка алгоритмов индуктивного вывода с использованием деревьев решений / А. М. Цветков // Кибернетика и системный анализ. – № 1. – 1993. – С. 174–178.
9. Свами М. Графы, сети и алгоритмы / М. Свами, К. Тхуласираман. – М. : Наука, 1984. – 408 с.
10. Месарович М. Теория иерархических многоуровневых систем / М. Месарович, Д. Мако, И. Такахаара. – М. : Мир, 1973. – 344 с.
11. Даревич Р. Р. Оцінка подібності текстових документів на основі визначення інформаційної ваги елементів бази знань / Р. Р. Даревич, Д. Г. Досин, В. В. Литвин, З. Т. Назарчук // Искусственный интеллект. – № 3. – 2006. – С. 500–509.

Надійшла 29.04.2009

*Исследуются методы реализации и функционирования мультиагентных систем поддержки принятия решений, базирующиеся на прецедентах и использующие адаптивные онтологии, которые включаются в состав интеллектуальных агентов. Разработана метрика для исчисления расстояния между прецедентом и текущей ситуацией на основе адаптивных онтологий.*

*Methods of developing and functioning multi-agent solution support systems based on precedents and using adaptive ontology which are the part of intelligent agents are analyzed. The method of distance definition between the precedent and current situation based on adaptive ontology is developed.*

І. О. Романенко, В. Ф. Столбов, В. В. Калачова

## ШЛЯХИ ОРГАНІЗАЦІЇ КОНТРОЛЮ ЗНАНЬ В СИСТЕМАХ ДИСТАНЦІЙНОГО НАВЧАННЯ

---

*Робота присвячена дослідженню проблеми підвищення ефективності процесу навчання в системах дистанційного навчання (СДН) завдяки використанню різних способів та технологій контролю знань при підготовці фахівців за дистанційною формою навчання у вищих навчальних закладах (ВНЗ) України в світі положень Болонського процесу. Проаналізовано існуючі шляхи створення продуктивних СДН та запропоновано новий спосіб контролю знань, що дозволяє адекватно оцінити знання тих, хто здійснює дистанційне навчання (ДН) та скоротити час контролю. Запропоновано концептуальну схему підсистеми контролю знань (ПСКЗ) і алгоритм створення тестів, які дозволяють забезпечити високу достовірність процесу контролю знань в СДН.*

### ВСТУП

**Постановка проблеми.** У зв'язку зі швидким зростанням можливостей сучасних технічних засобів, телекомунікаційних та інформаційних технологій відбуваються суттєві зміни форм та сенсу освіти в Україні і в усьому світі. Тенденції зростання ролі професійної та безперервної освіти, зростання долі навчання без відриву від професійної діяльності обумовлює попит на дистанційні освітні послуги, а ДН стає інтегральною формою її отримання.

Впровадження нових форм та новітніх ІТ навчання, до яких належить і дистанційна форма навчання, є актуальним процесом сьогодення.

Дистанційна форма навчання базується на використанні традиційних та інноваційних методів і засобів навчання, які основані на інформаційно-телекомунікаційних технологіях та забезпечують інтерактивну взаємодію учасників навчального процесу, а також отримання, вивчення і контроль засвоєння змісту навчання.

Впровадження дистанційного навчання у процес підготовки та перепідготовки фахівців в освітній системі України обумовлено низкою причин [1, 2]:

– інтеграція України до Європейського Союзу та впровадження європейських норм і стандартів в освіті та науці з урахуванням принципів Болонського процесу;

– інтенсивність розвитку науки потребує постійного удосконалення професійних знань та навичок службовців різних фахів;

– тільки технології ДН спроможні забезпечити своєчасне корегування змісту навчання фахівців за рахунок високої швидкості оновлення знань в інформаційно-освітньому середовищі (ІОС);

– висока економічна ефективність ДН.

Необхідність реформування системи освіти України, її удосконалення і підвищення рівня якості є найважливішою проблемою, яка значною мірою зумовлюється процесами глобалізації та потребами формування позитивних умов для індивідуального розвитку людини і її самореалізації.

Перехід до нової якості освіти є можливим, перш за все, при проведенні оптимізації процесу навчання. З цієї точки зору велику цікавість виявляє впровадження СДН, які передбачають гнучке поєднання самостійної пізнавальної діяльності тих, хто навчається та оперативну, систематичну взаємодію засобами комунікаційних каналів з провідними викладачами навчальних закладів. СДН забезпечують адаптацію процесу навчання до індивідуальних характеристик студентів, спрощують процес представлення учбової інформації і контролю знань, сприяють розробці і впровадженню нових методів контролю знань [3, 4].

Ефективність процесу значно залежить від періодичності, оперативності і якості заходів щодо контролю знань.

На сьогоднішній день існує досить велика кількість шляхів підвищення ефективності процесу навчання в СДН завдяки використанню різних способів та технологій контролю знань [5–8]. Вибір же конкретного варіанту комплексу засобів для здійснення контролю знань на відповідному рівні якості вимагає проведення відповідних досліджень.

Для рішення поставленої проблеми в межах цієї роботи проаналізовано існуючі шляхи створення ефективних СДН та запропоновано новий спосіб контролю знань, що дозволяє адекватно оцінити знання тих, хто навчається та скоротити час контролю за рахунок можливості дистанційного аналізу результатів контролю знань в СДН для здійснення ДН в вищих навчальних закладах України.

**Аналіз літератури.** Суть та зміст положень Болонського процесу та проблеми їх впровадження у навчальних закладах представлено роботами [1, 2].

Питання, зв'язані з теоретичними та практичними аспектами втілення дистанційних технологій навчання (ДН) у сучасну освіту розгорнуто надано в [3, 4]. Сучасні технології та способи здійснення контролю знань в СДН всебічно розглядаються в [5–8].

**Метою цієї роботи** є проведення всебічного аналізу можливих шляхів підвищення ефективності процесу навчання в СДН та внесення пропозиції, щодо способу контролю знань, який дозволить би адекватно оцінити знання студентів дистанційної форми навчання і скоротити час контролю.

### ВИКЛАД ОСНОВНОГО МАТЕРІАЛУ ДОСЛІДЖЕНЬ

Ефективність процесу навчання багато в чому залежить від періодичності, оперативності і якості заходів щодо контролю знань, тому, як і в будь-якій системі навчання, в СДН найважливішою складовою ПСКЗ [3, 4].

ПСКЗ СДН є діалектичним розвитком технічних засобів контролю знань на найбільш високому якісному рівні. Вона характеризується використанням могутнього засобу переробки інформації, активізацією ролі викладача і студента і дозволяє здійснювати оптимізацію контрольних заходів за рахунок використання в ній методів, моделей і сучасних інформаційних технологій навчання [5].

ПСКЗ можна представити у вигляді узагальненої концептуальної схеми (рис. 1), де кожен студент проходить перевірку ступеня його підготовки з відповідної дисципліни засобами тестуючого модуля [6].

Робота ПСКЗ в загальному вигляді може бути представлена чотирикроковим алгоритмом (рис. 2) [7].

1) Формування тесту, що включає етапи:

- складання тестових завдань у відповідності зі встановленими вимогами і особливостями розробленої СДН;

- формування тесту з тестових завдань;

- перевірка валідності, тобто перше попереднє призначення вагових коефіцієнтів значущості завдань в тесті.

2) Створення надійного тесту, що включає етапи:

- формування бази тестових завдань за допомогою редактора тестів;

- формування і занесення тесту в базу неактивних тестів;

- контрольне тестування;

- розрахунок показників ефективності тесту, що складається з перевірки якості тестових завдань і тесту в цілому;

- аналіз показників ефективності тестів;

- друге попереднє призначення вагових коефіцієнтів значущості завдань в тесті.

3) Тестування, що включає етапи:

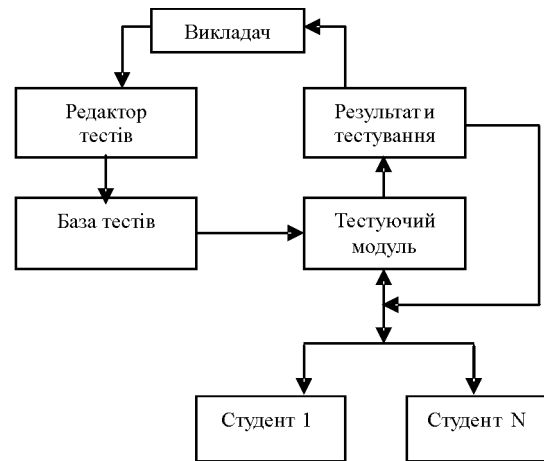


Рисунок 1 – Концептуальна схема ПСКЗ



Рисунок 2 – Алгоритм роботи ПСКЗ

- процес тестування ініціюється викладачем або студентом: вибирається область тестування; вибирається тест з бази активних тестів або динамічно формується і проводиться встановлення параметрів;

- формування бази результатів тестування.

4) Аналіз отриманих результатів. Внесення необхідних корегувань до учбового процесу.

Одним з кроків алгоритму роботи ПСКЗ є створення надійного тесту. Відповідно до теорії тестування, тест можна охарактеризувати як ефективний, якщо він володіє властивістю надійності. Існує декілька чинників, що впливають на надійність: довжина тестового набору; зміст тестового набору; кореляція між результатами виконання завдань; гетерогенність групи; характеристики завдань [8].

Виходячи з даних чинників, формування надійного тесту включає етап формування бази тестових завдань, який може бути охарактеризований схемою, представленою на рис. 3.

Після того, як буде проведено етап контрольного тестування сформованої бази тестових завдань, проводиться розрахунок показників ефективності тесту,



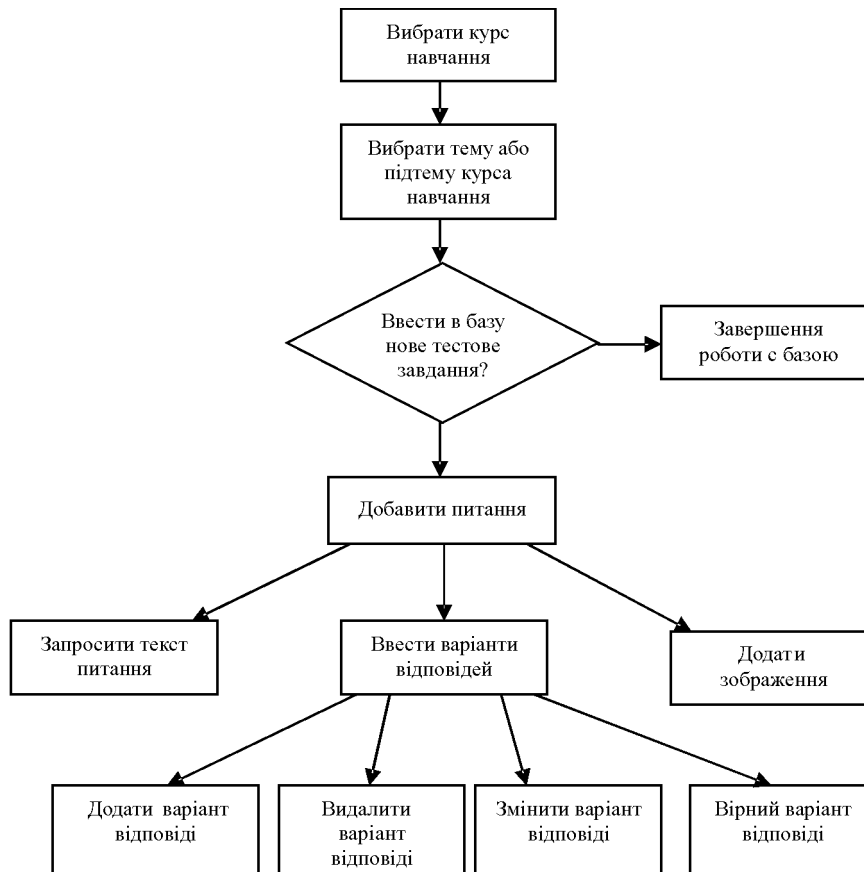


Рисунок 3 – Формування бази тестових завдань

що складається з перевірки якості тестових завдань і тесту в цілому.

Для проведення перевірки якості тестових завдань будується матриця результатів тестування. Проводиться впорядкування по тих, хто був тестованим і завданням. Визначається міра складності завдань. Для цього обчислюються наступні параметри:

- частка правильних відповідей;
- частка неправильних відповідей;
- варіація балів;
- дисперсія балів;
- диференціююча здатність;
- кореляція завдання.

Обчислення параметра «варіація балів» дозволяє виключати з тесту завдання, на які всі, хто проходив тест з контрольних груп відповіли правильно і ті, на які немає жодної правильної відповіді. Дані завдання можуть бути збережені і використані надалі: перші можуть бути використані при першочерговому тестуванні після початку навчання за даним курсом, другі можуть бути допрацьовані провідним викладачем курсу або використані для зміни викладання курсу.

Якщо на якесь завдання правильно відповідають всі тестовані і точно відомо, що склад групи неоднорідний, то таке завдання не диференціює випробову-

ваних. Воно повинне бути виключене з тесту або використовуватися при початковому тестуванні. Якщо ж на завдання не отримана жодна правильна відповідь – завдання сформульоване не коректно.

По закінченню перевірки якості тестових завдань видаються рекомендації по модифікації тестових завдань і проводиться перевірка якості тесту в цілому (проводиться розрахунок показників ефективності активних тестів, формується база даних).

Етап аналізу показників ефективності тестів, має на увазі процес порівняння показника з нормативним його значенням. Якщо показники відповідають нормі, то тест може поповнити базу активних тестів. Якщо показники незадовільні, то видаються рекомендації по коректуванню тестових завдань і тест відправляється на доопрацювання.

Завершальним етапом створення надійного тесту є друге попереднє призначення вагових коефіцієнтів значущості завдань в тесті ( $S_j$ ).

Частка неправильних відповідей ( $r_j$ ) змінюється від 0 до 1, тому цей параметр можна використовувати для ранжування завдань таким чином:

а) Вибирається кількість рівнів складності завдань  $m = 1, 2, \dots, n$  (значення  $n$  визначається розробником тесту і експертом).

б) Відрізок  $(0; 1)$  розділяється на  $m$  частин. Кожна частина відрізка відповідає певному рівню складності. Наприклад, при  $m = 1$  всім завданням призначаються однакові вагові коефіцієнти, наприклад  $S_j = 1$ . Всі завдання мають один рівень складності.

При  $m = 2$ , завдання мають два рівні складності. Завданням, у яких значення  $r_j$  (або  $r_j \pm$  похибка) потрапили в інтервал  $(0; 0,5)$  призначаються  $S_j = 1$ , а для інших (інтервал  $(0,5; 1)$ ) – призначаються  $S_j = 2$  і т. і.

## ВИСНОВКИ

Таким чином, в роботі проаналізовано існуючі шляхи створення ефективних СДН. Запропоновано новий спосіб контролю знань, що дозволяє адекватно оцінити знання тих, хто навчається та скоротити час контролю за рахунок можливості дистанційного аналізу результатів контролю. Запропонована концептуальна схема ПСКЗ і алгоритм створення тестів дозволяють забезпечити високу достовірність контролю знань в СДН.

## ПЕРЕЛІК ПОСИЛАНЬ

1. Основні засади розвитку вищої освіти України в контексті Болонського процесу (документи і матеріали 2003–2004 рр.) / За ред. В. Г. Кременя. – Київ – Тернопіль : ТДПУ, 2004. – 147 с.
2. Концепція розвитку дистанційної освіти в Україні. – Київ, 2001. – 2 с.
3. Образование и виртуальность – 2005. Сборник научных трудов 9-й Международной конференции Украинской ассоциации дистанционного образования / Под общ. Ред. В. А. Гребенюка и В. В. Семенца. – Харьков – Ялта : УАДО, 2005. – 315 с.
4. Романов А. Н. Технология дистанционного обучения / Торопцев В. С., Григорович Д. Б. – М. : ЮНИТИ ДАНА, 2000. – 287 с.

УДК 004.92

С. Ю. Скрупский, К. Н. Касьян

# ПРОГРАММНЫЕ МЕТОДЫ КОДИРОВАНИЯ ВИДЕОПОТОКОВ, ПРИМЕНЯЕМЫЕ В КОМПЬЮТЕРНОЙ ГРАФИКЕ

Выполнено исследование программных методов кодирования видеопотоков с точки зрения результирующего качества, размера, параллелизма и времени кодирования. Представлены результаты экспериментальных исследований популярных видеокодеков, разработан алгоритм выбора метода сжатия видеопотока.

© Скрупский С. Ю., Касьян К. Н., 2009

5. Аванесов В. С. Научные основы тестового контроля знаний. – М. : Иссл. центр, 1994. – 135 с.
6. Журавель В. Ф. Рекомендована практика констрування тестів професійної компетенції випускників вищих навчальних закладів / За загал. ред. Ю. В. Сухарнікова / Ільїн В. В., Кузнецов В. О., Сухарніков Ю. В. – К. : Аграрна освіта, 2000. – 38 с.
7. Аванесов В. С. Теоретические основы разработки педагогических тестов / Дис. на соиск. степени док. пед. наук. – Л., 1995. – 350 с.
8. Crocker L. Introduction to classical and modern test theory / Algina J. – New York : Holt, Rinehart and Winston, 1986. – 356 p.

Надійшла 09.04.2009

Після доробки 20.05.2009

*Работа посвящена исследованию проблемы повышения эффективности процесса обучения в системах дистанционного обучения (СДО) за счет использования различных способов и технологий контроля знаний при подготовке специалистов по дистанционной форме обучения в высших учебных заведениях Украины в свете положений Болонского процесса. Проанализированы существующие пути создания продуктивных СДО и предложен новый способ контроля знаний, который позволяет адекватно оценить знания тех, кто осуществляет дистанционное обучение (ДО) и сократит время контроля. Предложены концептуальная схема подсистемы контроля знаний (ПСКЗ) и алгоритм создания тестов, которые позволяют обеспечить высокую достоверность процесса контроля знаний в СДО.*

*The article researches the problems of improving the teaching efficiency process for the distance studies systems (DSS) at the expense of using different methods and technologies of knowledge control in training the specialists in higher educational institutions of Ukraine as per Bologna Process regulations. The present teaching methods for creating the resultant DSS have been analysed and new method of knowledge control that allows to adequately estimate the knowledge of those who practices the distance learning (DL) and decrease the control time. The conceptual subsystem for knowledge control (SSKC) and algorithm for tests compiling that allow to provide high authenticity of process in knowledge control when DSS are offered.*

## ВВЕДЕНИЕ

В настоящее время значительная часть данных, хранимых в электронном виде, представляет собой видеопоток. Видеопотоком называют набор бит, описывающих количество, положение и цвет пикселей

в кадрах, являющихся составной частью видеоизображения. В отличие от статического изображения, видео представляет собой совокупность статических изображений, меняющихся со скоростью равной частоте кадров. Частота кадров может быть от одного до тридцати, в зависимости от метода видеосъемки и не меняется на протяжении всего видеопотока. Размер кадра в пикселях варьируется от 1×1 до 1920×1080 (видео высокой четкости), однако самыми популярными значениями считаются 640×480 и 720×576. Чем больше размер кадра, тем качественнее видео, однако тем больше бит нужно для описания каждого кадра.

### ПОСТАНОВКА ЗАДАЧИ

Хранить видеопоток *абсолютно несжатым* нерационально ввиду значительного расхода дискового пространства. Для снижения размера видеопотока применяются *видеокодеки*, основной задачей которых является сокращение количества бит, описывающих пиксели кадра и, следовательно, всего видеопотока. Кодеки всегда являются звеном в цепочке программ, которые получают поток данных из источника, производят над ним алгоритмические манипуляции, и помещают в какой-либо другой поток данных. Кодеки могут как кодировать поток/сигнал (часто для передачи, хранения или шифрования), так и декодировать – для просмотра или изменения в формате, более подходящем для этих операций. Кодеки часто используются при цифровой обработке видео и звука. Поток данных на входе может быть как файлом, например несжатым видеофайлом, так и видеопотоком от цифровой камеры или TV. В первом случае входной поток генерируется драйвером файловой системы ОС. Во втором – входной поток генерируется драйвером устройства, которое поставляет видеопоток (например, драйвером TV-тюнера или видеокарты, при записи с видеовхода). В данной работе представлено экспериментальное сравнение современных методов сжатия видеопотоков, разработан алгоритм выбора метода сжатия видеопотока.

### ОБЗОР СОВРЕМЕННЫХ ВИДЕОКОДЕКОВ

*Сжатие* видеопотока было бы невозможно, если бы каждый кадр был уникален и расположение пикселей было полностью случайным, но это не так. Можно сжимать сам кадр или – похожие соседние кадры. В конечном счете, алгоритмы сжатия картинок и видео схожи, если рассматривать видео как трехмерное изображение со временем как третьей координатой. Большинство кодеков используют *сжатие с потерями*, чтобы получать приемлемый размер готового (сжатого) файла. Существуют две основные схемы сжатия с потерями:

– В трансформирующих кодеках используются кадры изображений, разрезаются на небольшие сегменты, трансформируются в новое базисное пространство и производится квантизация (использование не всех пикселей кадра, а только некоторых).

– В предсказывающих кодеках предыдущие и/или последующие данные используются для того, чтобы предсказать текущий кадр изображения. Ошибка между предсказанными данными и реальными вместе с добавочной информацией, необходимой для производства предсказания, квантизуется и кодируется.

В некоторых системах эти две техники комбинируются путем использования трансформирующих кодеков для сжатия ошибочных сигналов, сгенерированных на стадии предсказания. К таким кодекам относятся следующие:

- Flash;
- H.261;
- H.263;
- H.264;
- MNG;
- Motion JPEG;
- MPEG-1 Part 2;
- MPEG-2 Part 2;
- MPEG-4 Part 2;
- Ogg Theora (отличается отсутствием патентных ограничений);
- Sorenson video codec;
- VC-1 – попытка Microsoft выпустить открытую спецификацию для формата WMV.

Существуют также *кодеки, сжимающие без потерь*. Методы сжатия без потерь могут быть классифицированы по типу данных, для которых они были созданы. Три основных типа данных для алгоритма сжатия данных это текст, изображения и звук. В принципе, любой многоцелевой алгоритм сжатия данных без потерь (многоцелевой означает, что он может обрабатывать любой тип бинарных данных) может использоваться для любого типа данных. Большинство программ сжатия без потерь используют два различных типа алгоритмов: один генерирует статистическую модель для входящих данных, другой отображает входящие данные в битовом представлении, используя модель для получения «вероятностных» (то есть часто встречаемых) данных, которые используются чаще, чем «невероятностные». Алгоритмы кодирования через генерирование битовых последовательностей:

- Алгоритм Хаффмана;
- Арифметическое кодирование.

Кодеки, сжимающие без потерь, приведены ниже:

- Animation codec;
- CamStudio Video Codec;
- CorePNG;
- FFV1;
- Huffvuv;
- Lagarith;

- LCL;
- MSU Lossless Video Codec;
- Qbit Lossless Codec;
- SheerVideo;
- TSCC – TechSmith Screen Capture Codec.

Для большинства задач рациональнее применять кодеки с потерями информации, так как малозаметное ухудшение качества оправдывается значительным уменьшением объема данных. Почти единственное исключение – ситуация, когда данные будут подвергаться дальнейшей обработке: в этом случае повторяющиеся потери на кодировании/декодировании окажут серьезное влияние на качество.

Наиболее популярными на данный момент являются кодеки DivX, MPEG-2 DVD, H.264. Все они являются кодеками сжатия с потерями, однако алгоритмы сжатия у них различны.

**DivX** – кодек, поддерживаемый большинством современных DVD-плееров. Применяет общую компенсацию движения – помогает при увеличении («наезде» камеры) и при панорамировании (повороте камеры), если объект сохраняет свою статичность и меняется лишь в размере или местоположении на картинке – разумно при кодировании фильмов о природе и подобных. Деление изображения на макроблоки производится на основе целых долей позволяет более реалистично передавать движение объектов на изображении. Таким образом, DivX – это прогрессивная реализация формата MPEG-4. Он предлагает значительно более высокую степень сжатия видео и несет ряд революционных нововведений. Области применения MPEG-4 включают мультимедийные приложения для мобильного использования и Интернет, цифровое телевидение, домашнее видео и мобильную связь. Кодек поддерживает любую частоту кадров, стандарты PAL и NTSC.

**MPEG-2 DVD** – название группы стандартов цифрового кодирования видеосигналов. Стандарт используется для кодирования видеопотока при вещании, включая спутниковое вещание и кабельное телевидение. MPEG-2 с некоторыми модификациями также активно используется как стандарт для сжатия DVD, используется для общего сжатия движущихся изображений и определяет формат видеопотока, который может быть представлен как три типа кадра – независимосжатые кадры (I-кадры), кадры сжатые с использованием предсказания движения в одном направлении (P-кадры) и кадры, сжатые с использованием предсказания движения в двух направлениях (B-кадры). Использует потоки в 25 или 29,97 кадров в секунду.

**H.264** – стандарт сжатия видео, предназначенный для достижения высокой степени сжатия видеопотока при сохранении высокого качества. Этот относительно молодой кодек обладает рядом преимуществ:

- многокадровое предсказание кадров – использование сжатых ранее кадров в качестве опорных (то есть с заимствованием части материала из них) применяется более эффективно, чем в предыдущих стандартах. Позволяется использование до 32 ссылок на другие кадры. Это поднимает эффективность кодирования, так как позволяет кодеру выбирать для компенсации движения между большим количеством изображений;

- независимость порядка воспроизведения изображений и порядка опорных изображений. В предыдущих стандартах устанавливалась жесткая зависимость между порядком следования изображений для использования при компенсации движения и порядком следования изображений при воспроизведении;

- экстраполяция, выходящая за границы изображения. В MPEG-2 и предшествовавших ему стандартах экстраполяция могла указывать только на пиксели, находящиеся в границах декодированного опорного изображения;

- шеститочечная фильтрация компонента яркости для полупиксельного предсказания с целью уменьшения зубчатости краев и, в конечном счете, обеспечения большей четкости изображения;

- взвешенное предсказание, позволяющее использовать масштабирование и сдвиг после компенсации движения на величины, указанные кодером. Такая методика может чрезвычайно сильно поднять эффективность кодирования для сцен с изменением освещенности, например, при эффектах затемнения, постепенного появления изображения;

- сжатие макроблоков без потерь – позволяет точно описывать определенные области и допускает строгое ограничение на количество закодированных данных для каждого макроблока;

- дополнительное преобразование Адамара, применяемое к дискретно-косинусным коэффициентам основного пространственного преобразования (к коэффициентам яркости, и, в особом случае, цветности) для достижения большей степени сжатия в однородных областях.

Таким образом, DivX – самый популярный и распространенный кодек; MPEG-2 DVD – кодек, поддерживаемый DVD-плеерами, что обеспечивает ему популярность; H.264 – новый прогрессивный кодек, от которого теоретически ожидается низкий битрейт и высокое качество.

### **РЕЗУЛЬТАТЫ ЭКСПЕРИМЕНТАЛЬНОЙ ПРОВЕРКИ ВИДЕОКОДЕКОВ**

Для практической проверки кодеков с видеокamerы был записан ролик продолжительностью 20 минут размером кадра 640×480 25 кадров в секунду без сжатия. Затем несжатый видеопоток кодировался тремя

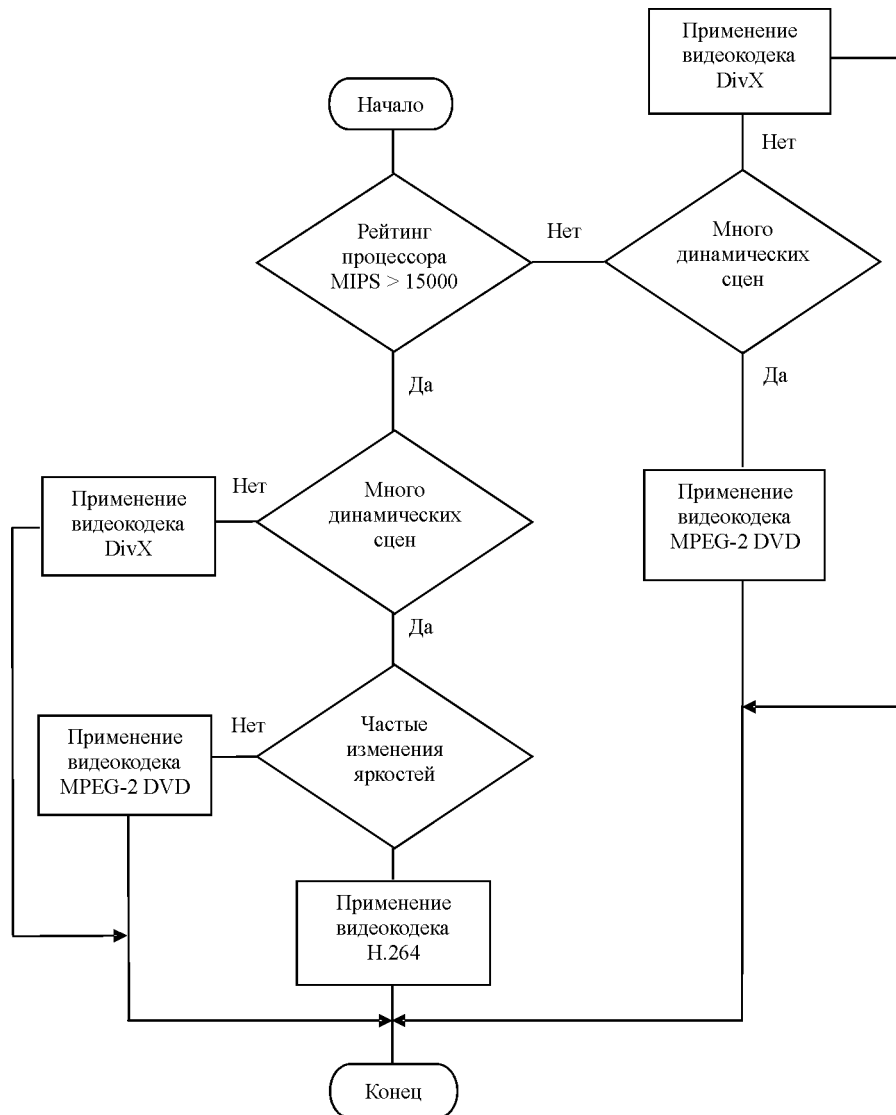


Рисунок 1 – Блок-схема алгоритма рационального выбора метода сжатия видеопотока

вышеперечисленными кодеками с целью фиксации результирующего размера файла, времени кодирования и качества относительно исходного видеопотока. Поскольку основными ресурсами компьютера, которые использует кодек, являются центральный процессор и оперативная память, то с целью проверки оптимизации кодека под многопоточность, измерения выполнялись на двух различных конфигурациях:

- 1 – AMD Sempron (2 ГГц) + DDR 400 (2 Гб).
- 2 – Intel Core 2 Duo (2 ГГц) + DDR2 800 (2 Гб).

Результаты тестирования приведены в табл. 1 и 2. Для оценки качества видеопотока [1] применяется метрика: среднее отношение мощности сигнал/шум (англ. аббр. APSNR) [2, 3]:

$$PSNR = 10 * \log_{10} \frac{\text{MaxErr}^2 * w * h}{\sum_{i=0}^{w-1} \sum_{j=0}^{h-1} (x_{i,j} - y_{i,j})^2}$$

где MaxErr – максимум модуля разности цветовой компоненты,  $w$  – ширина видео,  $h$  – высота видео. MaxErr зависит от битности цветовой компоненты: 255 для 8-битных компонент,  $255 + 3/4$  для 10-битных компонент,  $255 + 63/64$  для 14-битных компонент,  $255 + 255/256$  для 16-битных компонент, 256 для 24-битных компонент.

APSNR – усредненное значение PSNR по всем кадрам видеопотока.

APSNR<sub>исх</sub> несжатого видеопотока составляет 37 дБ. Размер исходного видеопотока – 26,3 Гб. Для всех

кодеків бітрейт встановлен на отметку 4 Мб/с, якість – 100 %.

Таблиця 1 – Результати тестування кодеків на системі 1

Кодек	Розмер результату, Мб	Время кодирования, мин	Качество результата APSNR/ APSNR <sub>исх</sub>
DivX	587	18	0,74
MPEG-2 DVD	586	19	0,67
H.264	586	100	0,89

Таблиця 2 – Результати тестування кодеків на системі 2

Кодек	Розмер результату, Мб	Время кодирования, мин	Качество результата APSNR/ APSNR <sub>исх</sub>	Ускорення
DivX	587	10	0,71	1,8
MPEG-2 DVD	586	11	0,62	1,73
H.264	586	50	0,89	2

## ВИВОДИ

В результаті аналізу і практичного порівняння програмних методів кодування відеопотоків, застосовуваних в комп'ютерній графіці, сформувався алгоритм вибору методу стиснення відеопотока (рис. 1). Для раціонального вибору методу стиснення відеопото-

ка необхідно визначити рейтинг MIPS процесора, який виконуватиме завдання кодування. Якщо CPU MIPS < 15000, тоді застосування H.264 не рекомендується внаслідок занадто великого часу, необхідного для виконання завдання. Далі необхідно проаналізувати вхідний потік на предмет наявності в ньому великої кількості динамічних сцен. Якщо таких сцен небагато, слід застосовувати DivX, інакше – MPEG-2 DVD. Якщо в відеопотоці зустрічаються часті зміни яркостей, тоді слід застосовувати H.264, інакше – MPEG-2 DVD.

## ПЕРЕЧЕНЬ ССЫЛОК

1. Грибунин В. Объективные метрики для оценки качества видеокодеків // В. Грибунин. – М. : Технологии защиты. – 2008. – № 2. – С. 20–24.
2. Winkler S. Digital Video Quality. Vision models and metrics / S. Winkler. – [S. l.] : Wiley, 2005. – 192 p.
3. Avcibas I. Image Quality Statistics and their use in steganalysis and compression / I. Avcibas. – [S. l.] : Wiley, 2001. – 113 p.

Надійшла 21.04.2009

Виконане дослідження програмних методів кодування відеопотоків з погляду результативності, розміру, паралелізму і часу кодування. Наведені результати експериментальних досліджень популярних відеокодеків, розроблено алгоритм вибору методу стиснення відеопотоку.

*There has been performed research on program methods of video streams coding from the view point of the resulting quality, size, parallelism and coding time. Results of experimental research into popular video codecs are presented, the algorithm for choice of a video stream compression method is developed.*

# ТЕОРІЯ І МЕТОДИ АВТОМАТИЧНОГО УПРАВЛІННЯ

## ТЕОРИЯ И МЕТОДЫ АВТОМАТИЧЕСКОГО УПРАВЛЕНИЯ

### THEORY AND METHODS OF AUTOMATIC CONTROL

УДК 681.513

Е. М. Потапенко, А. Е. Казурова

#### КОМПЕНСАЦИЯ ВОЗДЕЙСТВИЯ НЕЛИНЕЙНОГО ТРЕНИЯ НА СИСТЕМЫ УПРАВЛЕНИЯ

*Представлен обзор известных методов компенсации влияния нелинейного трения на системы управления. Предполагается, что измеряются только позиционные сигналы. Предложены редуцированные алгоритмы, оценивающие вектор неопределенности и скорости объекта управления. Осуществляется компенсация влияния вектора неопределенности комбинированным регулятором. Работоспособность алгоритмов управления проиллюстрирована численным моделированием.*

#### **ОБЗОР ИЗВЕСТНЫХ МЕТОДОВ КОМПЕНСАЦИИ ВЛИЯНИЯ НЕЛИНЕЙНОГО ТРЕНИЯ**

Из всех видов возмущений следует выделить нелинейное трение. Нелинейное трение может оказывать сильное отрицательное влияние на процессы управления. Это, в первую очередь, относится к управлению малыми скоростями и позиционному управлению. Существует большое количество моделей трения, отличающихся друг от друга как количеством учитываемых эффектов, так и их математическим описанием [1–7].

1. Кулоново трение описывается уравнением  $f = -f_0 \text{sign} \dot{x}$ , где  $f_0$  – модуль трения,  $\dot{x}$  – относительная скорость трущихся поверхностей.

2. Модель Леонардо да Винчи (предложенная в 1519 г. [1]), наряду с кулоновым трением  $f = -f_0 \text{sign} \dot{x}$ ,  $\forall |\dot{x}| > 0$ , учитывает статическое трение.

3. Модели Гаусса, Лоренца, Тастина [1].

4. Lund Grenoble (LuGre), предложенная в 1995 г. в городах Lund и Grenoble [2].

5. Модели, перечисленные в пп. 3 и 4, помимо описанных выше эффектов, учитывают уменьшение модуля трения в районе нулевой скорости при увеличении ее модуля (Stribeck-эффект). Эти модели отличаются друг от друга математическим описанием.

Зависимость трения от скорости для моделей, перечисленных в п. 3 и 4, имеет вид, представленный на рис. 1. Эту модель, можно описать следующей системой уравнений:

$$\begin{aligned} f &= -m \text{ при } \dot{x} = 0, \quad \forall |m| \leq f_0, \\ f &= -f_0 \text{sign} m \text{ при } \dot{x} = 0, \quad \forall (|m| > f_0), \\ f &= -\varphi(|\dot{x}|) \text{sign} \dot{x} \text{ при } |\dot{x}| > 0, \end{aligned}$$

где  $m$  – сумма всех внешних воздействий на объект управления (ОУ),  $f_0$  – максимальный момент трения

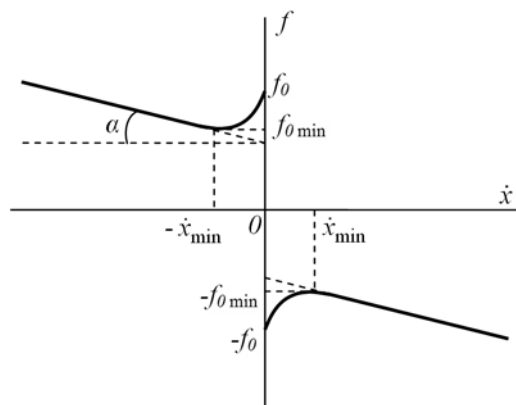


Рисунок 1 – Залежність тріння від швидкості

покою. Моделі из пп. 3, 4 отличаются друг от друга функциями  $|f| = \varphi(\dot{x})$ . Помимо перечисленных эффектов в нелинейном трении присутствуют такие явления как залипание, эффект «щетки» (смещение, предшествующее скольжению), неравномерное движение. Влияние нелинейного трения существенно отличается в зависимости от назначения системы управления, а именно, предназначена система для управления скоростью одного знака или для стабилизации положения, происходящей со знакопеременными скоростями.

При управлении скоростью одного знака трение постоянно находится на одной и той же ветви, пусть, например, на левой ветви рис. 1. На этой ветви при возрастании модуля скорости можно выделить ниспадающую и возрастающую части. При нахождении скорости на восходящей части сила трения стабилизирует движение. При нахождении на нисходящей – дестабилизирует. Вследствие этого в системе могут возникнуть автоколебания скорости между указанными участками. Наиболее сложным в динамическом отношении является режим стабилизации положения с нулевой скоростью. В этом случае проявляются все выше перечисленные эффекты трения. В отличие от управления скоростью, при позиционном управлении трение делает задачу управления существенно нелинейной, так как управление происходит в окрестности нулевой скорости (см. рис. 1). Данный обзор посвящен компенсации влияния трения при позиционном управлении.

Позиционному управлению в присутствии нелинейного трения в отечественной литературе уделено очень мало внимания. Следует отметить работу Ларина В. Б. [8], в которой рассматривается сухое трение (модель Леонардо да Винчи). В этой работе показано, что в случае ПИД-регулятора требуемое положение равновесия асимптотически устойчиво, но процесс затухания медленный и указано на необходимость вычисления и компенсации трения. Для оценки трения, по заключению автора, необходимо измерять

ускорение. Таким образом, для компенсации трения предлагается использовать комбинированное управление. (Как бы в дополнение к работе [8], в работе [5] указано, что при наличии падающих участков характеристики трения системы с ПИД-регуляторами являются неустойчивыми из-за наличия интегральной части регулятора).

В противоположность отечественной литературе, в иностранной литературе уделяется очень большое внимание позиционному управлению с нелинейным трением (см. работы [1–6, 9–17] и библиографию к ним). Рассматриваются следующие методы:

1. Предварительная (off-line) полная или частичная идентификация выбранной модели трения с последующим ее доопределением в реальном времени (on-line) и компенсацией на фоне ПД-регулирования [3–6, 11, 18, 20].

2. Введение уравнений воздействия трения в общую систему уравнений движения и по этим уравнениям синтез законов управления [13, 15, 16, 19].

3. Компенсация трения без учета его математической модели за счет применения таких робастных методов управления как управление с большими коэффициентами усиления (БКУ) или системы с переменной структурой (СПС) [9, 17].

С целью обеспечения работоспособности системы позиционного управления в присутствии нелинейного трения в работах [3–6, 11] предлагается применять комбинированное управление с предварительной идентификацией и последующей компенсацией трения в сочетании с ПД-управлением. Авторы этих работ задаются теми или иными существенно нелинейными моделями трения и предварительно идентифицируют коэффициенты этих моделей, с тем, чтобы по полученным моделям уже в реальном времени вычислять трение и компенсировать его. Этот метод обладает целым рядом существенных недостатков. Во-первых, нет таких моделей, которые учитывают все эффекты. Во-вторых, при наличии других, помимо трения, возмущений идентификация будет не точной. В-третьих, трение не может быть стабильным и компенсация не может быть полной. В-четвертых, ПД-регулятор не в состоянии в достаточной степени уменьшать ошибки от других воздействий.

Метод, описанный в работах [13, 15, 16], базируется на построении функций Ляпунова для модели всей системы, включающей модель трения. Он обладает теми же недостатками, что и предыдущий метод. Кроме того, алгоритмы управления (если их даже удастся синтезировать) получаются очень сложными.

СПС обладают высокой степенью робастности, но им присущ ряд недостатков, а именно: плохая помехозащищенность по отношению к высокочастотным помехам, способность возбуждать колебания упругих элементов, внесение в систему высокочастотных колебаний, ухудшающих точность управления скоростью,



отсутствие робастности на участках достижения линий переключения, повышенные энергозатраты, пониженная надежность, генерация акустических шумов. С целью уменьшения степени (но не полного устранения) каждого из перечисленных недостатков в работе [9] для формирования управляющих воздействий релейные звенья  $\text{sign}$  меняются на линейные звенья с насыщением (saturation), оценивается сумма внешних воздействий, по которой с некоторым гарантированным запасом устанавливается амплитуда управляющих воздействий. На линейных участках звеньев saturation используются ПИД-регуляторы, не рекомендованные в работе [5] для компенсации трения. Кроме того, в работе [5] для оценки внешних воздействий требуется информация о перемещениях, скоростях и ускорениях.

Таким образом, задача точного позиционного управления при наличии нелинейного трения требует лучшего решения. По нашему мнению, решение указанной задачи можно получить за счет робастного комбинированного управления, исследованного в работах [21–23]. Изложим суть этого метода, разработанного для общего случая неопределенностей, применительно к случаю, когда в неопределенность входит нелинейное трение. Объединение всех внешних воздействий, включая трение, и воздействий параметрических неопределенностей в один вектор неопределенности, его идентификация в реальном времени и компенсация на фоне регулирования формируют комбинированную систему управления с наблюдателем неопределенности – КСУНН [21–23]. Однако этот метод ранее не рассматривался применительно к объектам с нелинейным трением. Кроме того, методы оценки вектора скорости в работах [21–23] или сложны, или не учитывают динамику объекта управления. Ниже предлагается новый метод реализации комбинированного управления с наблюдателем вектора неопределенности и вектора состояния объекта управления (ОУ), отличающийся повышенной точностью за счет учета динамики ОУ при сравнительно малом объеме вычислений.

**КОМПЕНСАЦИЯ С ПОМОЩЬЮ КОМБИНИРОВАННОГО УПРАВЛЕНИЯ СОВМЕСТНОГО ВОЗДЕЙСТВИЯ НЕЛИНЕЙНОГО ТРЕНИЯ И ДРУГИХ ВОЗМУЩЕНИЙ**

Рассматривается объект управления

$$\ddot{x} = A_1x + A_2\dot{x} + Bu + Gf, \quad x \in R^n, \quad f \in R^\alpha, \quad (1)$$

где  $x, u$  – векторы состояния и управления,  $A_1, A_2, B, G$  – постоянные известные матрицы соответствующих размеров,  $f = f(x, u, t)$  – ограниченная кусочно дифференцируемая вектор-функция (неопределен-

ность), в которую собраны неизвестные погрешности объекта (1), внешние воздействия на объект, включая трение, и погрешности датчиков. Наблюдатель неопределенности будет строиться таким образом, что его быстрое действие в достаточной степени будет больше быстрого действия управляемого объекта. В этом случае функция  $f$  будет медленно меняющейся в масштабе времени наблюдателя неопределенности, вследствие чего при его синтезе будем полагать

$$\dot{f} = 0. \quad (2)$$

(Работоспособность такого предположения подтверждается численным моделированием.) Зададимся целью построить наблюдатель для системы (1), (2) с измерением только  $x$ , разделяющий скорость  $\dot{x}$  и неопределенность  $f$  (редуцированный наблюдатель). Перепишем уравнение (1) в виде

$$\ddot{x} = A_2\dot{x} + Gf + [A_1x + Bu]. \quad (3)$$

Пусть выполняется условие согласованности управления и возмущения  $G = BQ$ , где  $Q$  – некоторая известная матрица, и матрица  $B$  обратимая. В уравнении (3) вектор в скобках известен. Предположим на время, что скорость  $\dot{x}$  известна. В системе (2), (3) примем в качестве оцениваемых переменных  $\dot{x}$  и  $f$  с измерением

$$y = \dot{x}. \quad (4)$$

Поскольку в соответствии с теоремой 1.31 работы [24] система (2)–(4) является полностью наблюдаемой, построим для нее наблюдатель в виде [25]

$$\ddot{\hat{x}} = A_2\dot{\hat{x}} + G\hat{f} + L_x(\dot{\hat{x}} - \dot{x}) + [A_1x + Bu], \quad (5)$$

$$\dot{\hat{f}} = L_f(\dot{\hat{x}} - \dot{x}), \quad (6)$$

где  $L_x, L_f$  – матрицы коэффициентов передачи наблюдателя. Для устранения необходимости дифференцирования выходного сигнала  $x$  введем новые переменные

$$z_x = \dot{\hat{x}} + L_x x, \quad z_f = \hat{f} + L_f x, \quad (7)$$

откуда следует

$$\dot{\hat{x}} = z_x - L_x x, \quad \hat{f} = z_f - L_f x. \quad (8)$$

С учетом выражений (8) система (5), (6) принимает вид

$$\dot{z}_x = (A_2 + L_x)(z_x - L_x x) + G(z_f - L_f x) + [A_1x + Bu], \quad (9)$$

$$\dot{z}_f = L_f(z_x - L_x x). \quad (10)$$

В системе (9), (10) отсутствует неизмеряемый вектор  $\dot{x}$ . Поэтому необходимость в измерении вектора  $\dot{x}$  отпадает. Закон управления сформируем в виде

$$u = K_1x + K_2\dot{\hat{x}} - Q\hat{f}, \quad (11)$$

где  $\hat{x}$ ,  $\hat{f}$  определены в (8).

Известно, что объем вычислений при интегрировании системы уравнений пропорционален квадрату порядка системы. Пусть каждое из уравнений системы (12), (13) имеет порядок  $n$ . Тогда объем вычислений наблюдателя разделим наблюдатель (9), (10) на два блока с односторонней связью. С этой целью проделаем неособые преобразования системы (9), (10) подобные тому, как было сделано в работе [21], в результате чего получим

$$\dot{z}_x^* = (A_2 + L_x^*)z_x^* + Bu + [A_1 - (A_2 + L_x^*)L_x - GL_f]x, \quad (12)$$

$$\dot{z}_f = L_f(z_x^* + Sz_f - L_x x), \quad (13)$$

где

$$S = -(A_2 + L_x^*)^{-1}G, \quad (14)$$

$$z_x = z_x^* + Sz_f, \quad (15)$$

$$L_x = L_x^* + SL_f. \quad (16)$$

Из уравнений (12), (13) видно, что уравнение (12) независимо от уравнения (13). С учетом квадратичной зависимости объема вычислений от порядка системы объем вычислений  $N = n^2 + n^2 = 2n^2$ , что в два раза меньше, чем в неразложенной системе.

### ПРИМЕР

Рассматривается управляемое движение твердого тела с неизвестной массой, на которое действуют нелинейное трение и нагрузка неизвестных величин. Крутизна характеристики исполнительного органа также не известна. Измеряется только перемещение тела.

ОУ описывается уравнением

$$I_c \ddot{\phi}_c = nc_m i - f_c - f_l, \quad (17)$$

где  $\phi_c$  – угол поворота ОУ,  $I_c$  – его момент инерции,  $i$  – ток якорной цепи двигателя постоянного тока (ДПТ),  $c_m$  – его моментный коэффициент;  $n$  – передаточное число редуктора,  $f_c$  – момент нелинейного трения в ОУ,  $f_l$  – момент нагрузки. В качестве модели трения использовалась модель LuGre (рис. 1).

Момент инерции ОУ и управляющее воздействие представим в виде

$$I_c = I_{c0} + I_{c\delta}, \quad m := nc_m i = nc_m i_0 + nc_m i_\delta. \quad (18)$$

Здесь индексом 0 обозначены известные номинальные части соответствующих величин, а индексом  $\delta$  – погрешности их знания. С учетом обозначений (18) уравнение (17) примет вид

$$I_{c0} \ddot{\phi}_c = nc_m i_0 + f_\Sigma, \quad (19)$$

где суммарная неопределенность  $f_\Sigma$  описывается следующим выражением:

$$f_\Sigma = nc_m i_\delta - f_c - f_l - I_{c\delta} \ddot{\phi}_c. \quad (20)$$

Согласно уравнению (2) запишем

$$\dot{f}_\Sigma = 0. \quad (21)$$

Построим наблюдатель для системы (19), (21), соответствующий наблюдателю (8)–(10)

$$\dot{\hat{\phi}}_c = z_\phi - l_\phi \phi_c, \quad \dot{\hat{f}}_\Sigma = z_f - l_f \phi_c, \quad (22)$$

$$\dot{z}_\phi = I_{c0}^{-1} nc_m i_0 + I_{c0}^{-1} (z_f - l_f \phi_c) + l_\phi (z_\phi - l_\phi \phi_c), \quad (23)$$

$$\dot{z}_f = l_f (z_\phi - l_\phi \phi_c), \quad (24)$$

где  $l_\phi$ ,  $l_f$  – коэффициенты передачи наблюдателя. Закон управления (11) будет иметь вид

$$i = k_1(\phi_c - \phi_{cp}) + k_2(\hat{\phi}_c - \dot{\phi}_{cp}) - (nc_m)^{-1} \hat{f}_\Sigma + (nc_m)^{-1} I_{c0} \dot{\phi}_{cp}, \quad (25)$$

где  $k_1$ ,  $k_2$  – коэффициенты закона управления. Расчет коэффициентов закона управления и наблюдателей осуществлялся по стандартным полиномам с биномиальным распределением корней.

При моделировании принимались следующие параметры:  $I_{c0} = 250$  кг·м<sup>2</sup>, момент инерции считается неизвестным и находится в диапазоне  $I_c = (0, 5 \dots 2)I_{c0}$ ,  $n = 377$ ,  $c_m = 0,062$  Н·м·А<sup>-1</sup>,  $f_l = 0 \dots 500$  Н·м,  $f_c = 200$  Н·м,  $nc_m i = (1 \pm 0,2)nc_m i_0$ , ток ограничивался значениями  $\pm 360$  А.

На рис. 2, 3 показан переходный процесс, возникающий при подаче ступенчатого программного сигнала величиной  $3^\circ$  ( $\phi_{cp} = 0,05236$  рад) при номинальном моменте инерции ОУ. Нагрузка  $f_l = 500$  Н·м в виде ступенчатого сигнала прикладывалась на 0,5 с и снималась на 0,8 с. На рис. 2 показан весь переходный процесс, а на рис. 3 – его фрагмент. Как видно

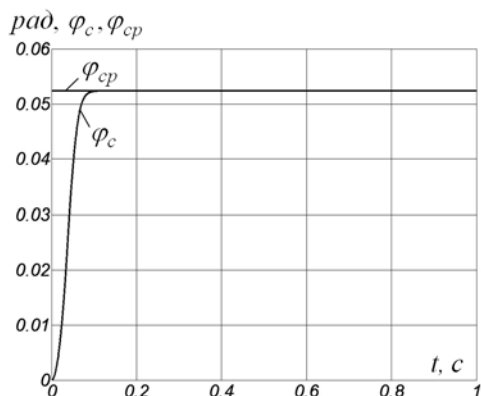


Рисунок 2

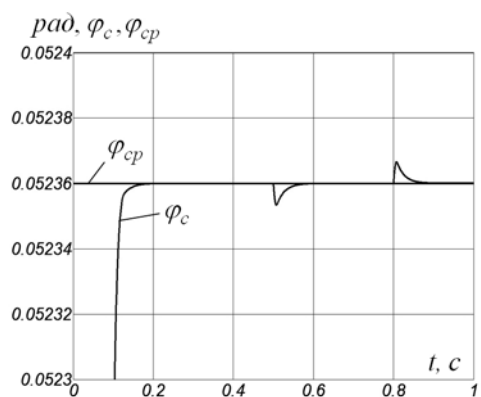


Рисунок 3

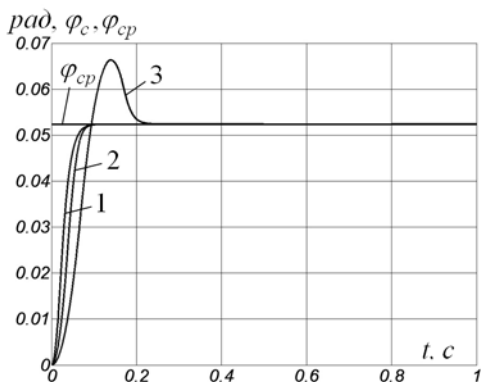


Рисунок 4

из графиков, синтезированная система управления обеспечивает аperiодический переходный процесс, является астатической по отношению к нагрузке, хорошо работает при наличии нелинейного трения, а также в системе отсутствует автоколебательный режим, характерный для СПС. Следует отметить, что ошибка от трения не оказывает существенного влияния на переходной процесс и полностью отсутствует

в установившемся режиме. В то же время, при ПИД-регуляторе система оказывается неустойчивой. На рис. 4 представлены переходные процессы, возникающие при моделировании следующих случаев: 1) с минимальным моментом инерции ОУ при максимальной крутизне моментной характеристики двигателя, 2) с номинальными параметрами, 3) с максимальным моментом инерции при минимальной крутизне моментной характеристики двигателя, которые подтверждают работоспособность системы при оговоренных условиях. Следует обратить внимание на то, что сочетания отклонений момента инерции ОУ и крутизны исполнительного органа выбраны наихудшими. В случае сочетания 3 наблюдается перерегулирование, что объясняется наличием ограничения на напряжение. Как показывают результаты моделирования, осуществляются точная оценка неопределенности и ее компенсация, за счет чего достигается высокая точность управления.

### ВЫВОДЫ

Анализ методов компенсации влияния трения при позиционном управлении позволяет заключить, что наибольшими возможностями обеспечения высокой точности управления при наличии нелинейного трения любого типа (в том числе и типа LuGre) обладает комбинированный метод управления с наблюдателями вектора неопределенности. Высокая точность управления достигается за счет комбинированного принципа действия, который заключается в следующем. В вектор неопределенности объединяются все неизвестные внешние воздействия, включая трение, и воздействия параметрических неопределенностей. С помощью наблюдателя этот вектор идентифицируется и компенсируется одной из составляющих закона управления. Вторая составляющая закона управления обеспечивает управление оставшейся полностью детерминированной линейной стационарной системой. Благодаря такому принципу действия разработанных систем, в отличие от систем с переменной структурой, осуществляется «мягкое» управление с пониженными энергетическими затратами, повышается помехоустойчивость к высокочастотным помехам и в меньшей степени «возбуждается» паразитная динамика. Также следует отметить, что разрабатываемые в данной статье алгоритмы не требуют знания математической модели трения. Модель необходима только для иллюстрации работоспособности рассматриваемого комбинированного управления путем численного моделирования.

Особо следует отметить, что, несмотря на наличие или отсутствие различного вида неопределенностей и возмущений, показатели качества переходных

процессов и точность управления практически не меняются. (Имеет место робастная модальность).

Численное моделирование подтвердило правильность предлагаемого метода синтеза робастной системы управления, а также показало возможность получения с помощью этого метода заданных динамических характеристик.

### ПЕРЕЧЕНЬ ССЫЛОК

1. *Armstrong-Helouvy B.* A survey of models, analysis tools, and compensation methods for the control of machines with friction / B. Armstrong-Helouvy, P. Dupont, and C. Canudas de Wit // *Automatica*. – 1994. – Vol. 30, No. 7. – Pp. 1083–1138.
2. *Canudas-de-Wit C.* A new model for control of systems with friction / C. Canudas-de-Wit, H. Olsson, K.J. Astrom, and P. Lischinsky // *IEEE Transactions on Automatic Control*. – 1995. – Vol. 40. – Pp. 419–425.
3. *Vedagarbha P.* Tracking control of mechanical systems in the presence of nonlinear dynamic friction effects / P. Vedagarbha, D. M. Dawson, and M. Feemster // *IEEE Transactions on Control Systems Technology*. – 1999. – Vol. 7, No. 4. – Pp. 446–456.
4. *Hirschorn R. M.* Control of nonlinear systems with friction / R.M. Hirschorn, G. Miller // *IEEE Transactions on Control Systems Technology*. – 1999. – Vol. 7, No. 5. – Pp. 588–595.
5. *Mallon N.* Friction compensation in a controlled one-link robot using a reduced-order observer / N. Mallon, N. van de Wouw, D. Putra, and H. Nijmeijer // *IEEE Transactions on Control Systems Technology*. – 2006. – Vol. 14, No. 2. – Pp. 374–383.
6. *Marton L.* Modeling, identification, and compensation of stick-slip friction / L. Marton and B. Lantos // *IEEE Transactions on Industrial Electronics*. – 2007. – Vol. 54, No. 1. – Pp. 511–521.
7. *Костиков А. Ю.* Идентификация параметров динамической модели трения в электроприводах технологических машин / А. Ю. Костиков // Труды III Международной конференции «Идентификация систем и задачи управления» SICPRO'04, Москва, 28–30 января 2004 г. – 2004. – С. 571–588.
8. *Ларин В. Б.* О компенсации нелинейного трения в робототехнических системах / В. Б. Ларин // *Прикладная механика*. – 1996. – Т 32, No. 5. – С. 69–74.
9. *Yu-Feng Li.* Sliding mode control of two-mass positioning systems / Yu-Feng Li, B. Eriksson and J. Wikander // *14<sup>th</sup> World Congress of IFAC*, Beijing, P. R. China, 1999. – Paper C-2a-05-2. – Pp. 151–156.
10. *Cho S.-I.* A learning approach to Stick-Slip friction compensation / S.-I. Cho and I.-J. Ha // *14th Triennial World Congress of IFAC*, Beijing, P. R. China. – 1999. – Pp. 169–174.
11. *Shirazi L. R.* Robust friction compensation based on Karnopp model / L.R. Shirazi, A. Besancon-Voda // *Proceedings of the ECC 2001*, 4–7 September, 2001, Seminario de Vilar, Porto, Portugal. – 2001. – Pp. 2558–2563.
12. *Vivas C.* Gain-scheduling control of systems with dynamic friction / C. Vivas, F.R. Rubio, C. Canudas-de-Wit // *Proc. Of the 41<sup>st</sup> IEEE Conf. on Decision and Control*, Las Vegas, Nevada, USA, December, 2002. – Pp. 89–94.
13. *Hung N. V. Q.* Adaptive control for nonlinearly parameterized uncertainties in robot manipulators / N. V. Q. Hung, H.D. Tuan, T. Narikiyo and P. Apkarian // *Proceedings of the 41st IEEE Conference on Decision and Control*, Las Vegas, Nevada USA, December 2002. – 2002. – Pp. 1727–1732.
14. *Grujic L. T.* Control synthesis for robot motion robust exponential stability. Part 1: Modelling and synthesis / L. T. Grujic, A. Kokosy // *In Pros. of IFAC-IFIP-IMACS Conf. Control of Industrial Systems*, Belfort, France, 20–22 May, 1997, Vol. 3. – Pp. 704–709.
15. *Kokosy A.* Control synthesis for robot motion robust exponential stability. Part 2: Application and simulations / A. Kokosy, L.T. Grujic // *In Pros. of IFAC-IFIP-IMACS Conf. Control of Industrial Systems*, Belfort, France, 20–22 May, 1997, Vol. 3. – Pp. 710–715.
16. *Huang Chin-I.* Adaptive backstepping speed/position control with friction compensation for linear induction motor / Chin-I. Huang, and Li-Chen Fu // *In Pros. of the 41<sup>st</sup> IEEE Conf. on Decision and Control*, Las Vegas, Nevada, USA, December, 2002. – 2002. – TuM03-3. – Pp. 474–479.
17. *Alonge F.* Adaptive and robust techniques for friction compensation in motion control of robotic manipulators / F. Alonge, F.D'Ippolito, F.M. Raimondi // *In Pros. of IFAC-IFIP-IMACS Conf. Control of Industrial Systems*, Belfort, France, 20-22 May, 1997, Vol. 3. – Pp. 697–703.
18. *Wen-Fang Xie.* Sliding-mode-observer-based adaptive control for servo actuator with friction / Xie Wen-Fang // *IEEE Transactions on Industrial Electronics*. – 2007. – Vol. 54, No. 3. – Pp. 1517–1527.
19. *Makker C.* Lyapunov-based tracking control in the presence of uncertain nonlinear parameterizable friction / C. Makker, G. Hu, and W. E. Dixon // *IEEE Transactions on Automatic Control*. – 2007. – Vol. 52. – Pp. 1988–1994.
20. *Maolin Jin.* Robust compliant motion control of robot with nonlinear friction using time-delay estimation / Jin Maolin, Hoong Kang Sang, and Hun Chang Pyung // *IEEE Transactions on Industrial Electronics*. – 2008. – Vol. 55, No. 1. – Pp. 258–269.
21. *Потапенко Е. М.* Сравнительная оценка робастных систем управления с различными типами наблюдателей / Е. М. Потапенко // *Изв. РАН. Теория и системы управления*. – 1995. – № 1. – С. 109–117.
22. *Потапенко Е. М.* Робастное управление маневрирующим упругим космическим аппаратом / Е. М. Потапенко, В. Г. Бичай // *Проблемы управления и информатики*. (Киев, НАНУ) – 1998. – № 4. – С. 72–85.
23. *Потапенко Е. М.* Об общности альтернативных робастных систем управления / Е. М. Потапенко, В. Г. Бичай // *Проблемы управления и информатики* (Киев, НАНУ). – 1998. – № 5. – С. 27–30.
24. *Квакернаак Х.* Линейные оптимальные системы управления / Х. Квакернаак, Р. Сиван. – М. : Мир, 1977. – 650 с.
25. *Кузовков Н. Т.* Модальное управление и наблюдающие устройства / Н.Т. Кузовков. – М. : Машиностроение, 1976. – 184 с.

Надійшла 25.05.2009

*Представлено огляд відомих методів компенсації впливу нелінійного тертя на системи керування. Передбачається, що вимірюються лише позиційні сигнали. Запропоновані редуковані алгоритми, що оцінюють вектор невизначеності та швидкості об'єкта керування. Здійснюється компенсація впливу вектора невизначеності комбінованим регулятором. Працездатність алгоритмів керування проілюстровано чисельним моделюванням.*

*The review of known compensation methods for a nonlinear friction effect on control systems is presented. It is assumed that only position signals are measured. The reduced algorithms that estimate an uncertainty vector and speeds for a control object are proposed. The compensation of the effect of the uncertainty vector is performed with a compound controller. The working capacity of the control algorithms is illustrated by a numerical simulation.*

# УПРАВЛІННЯ У ТЕХНІЧНИХ СИСТЕМАХ

# УПРАВЛЕНИЕ В ТЕХНИЧЕСКИХ СИСТЕМАХ

# CONTROL IN TECHNICAL SYSTEMS

УДК 62-55:681.515

В. И. Гостев

## ПРОЕКТИРОВАНИЕ НЕЧЕТКОГО РЕГУЛЯТОРА С ВЫХОДНЫМИ СЖАТЫМИ, ЗАВИСЯЩИМИ ОТ ОДНОГО ПАРАМЕТРА, S-ОБРАЗНОЙ И Z-ОБРАЗНОЙ ФУНКЦИЯМИ ПРИНАДЛЕЖНОСТИ

Получены аналитические выражения для управляющих воздействий на выходе нечеткого регулятора при выходных разведенных сжатых, зависящих от одного параметра, S-образной и Z-образной функциях принадлежности и изложены вопросы проектирования нечеткого регулятора.

### ВВЕДЕНИЕ

В работе [1] изложен новый метод проектирования нечетких регуляторов. В данной работе на основе изложенного метода проектируется нечеткий регулятор, структурная схема которого представлена на рис. 1 [1]. Нечеткий регулятор состоит из трех блоков: блока *формирователя величин*  $A(t)$  и  $B(t)$  (блок 1), блока *сравнения величин*  $A(t)$  и  $B(t)$  и

расчета  $u_c$  (блок 2) и блока *нормировки выходной переменной* (блок 3).

### РЕШЕНИЕ ЗАДАЧИ

Рассмотрим на универсальном множестве  $U = [0, 1]$  два нечетких подмножества с идентичными модифицированными сжатыми совместно используемыми, зависящими от одного параметра S-образной и Z-образной функциями принадлежности для каждой лингвистической величины (см. рис. 2), которые определяются по формулам

$$\mu_1(u) = \frac{(a-u)^2}{a^2}, \quad 0 \leq u \leq a;$$
$$\mu_2(u) = \frac{(u-1+a)^2}{a^2}, \quad 1-a \leq u \leq 1. \quad (1)$$

Допустим, что с выхода *формирователя величин*  $A(t)$  и  $B(t)$  в нечетком регуляторе на блок *сравнения величин*  $A(t)$  и  $B(t)$  и *расчета*  $u_c$  с такими функциями принадлежности в фиксированный момент времени поступают величины  $A$  и  $B$ , как показано на рис. 2.

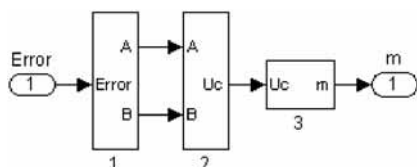


Рисунок 1

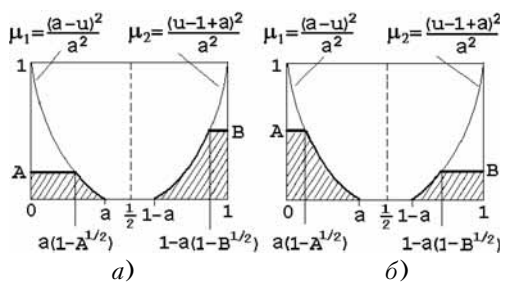


Рисунок 2

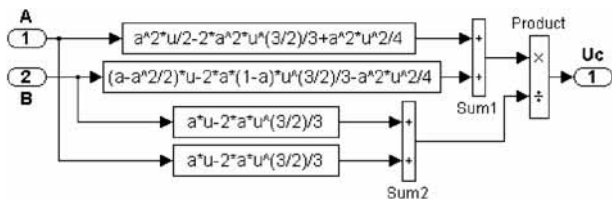


Рисунок 3

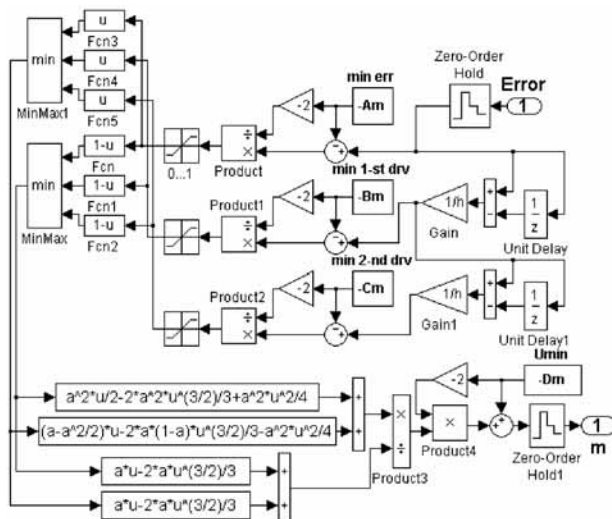


Рисунок 4

Общая формула для определения абсциссы «центра тяжести результирующей фигуры» записывается в виде

$$u_c = \frac{\int_0^1 u \mu(u) du}{\int_0^1 \mu(u) du} \quad (2)$$

Абсцисса «центра тяжести результирующей фигуры» определяется при  $A \geq B$  и при  $A \leq B$  по одной и той же формуле

$$u_c = \frac{A \int_0^{a(1-A^{1/2})} u du + \frac{1}{a^2} \int_{a(1-A^{1/2})}^{1/2} (a-u)^2 u du + \frac{1}{a^2} \int_{1-a}^{1-a(1+B^{1/2})} (u-1+a)^2 u du}{A \int_0^{a(1-A^{1/2})} du + \frac{1}{a^2} \int_{a(1-A^{1/2})}^{1/2} (a-u)^2 du + \frac{1}{a^2} \int_{1-a}^{1-a(1+B^{1/2})} (u-1+a)^2 du + B \int_{1-a(1-B^{1/2})}^1 du}$$

После несложных вычислений находим

$$u_c = \frac{\frac{a^2}{2}A + \left(a - \frac{a^2}{2}\right)B - \frac{2}{3}a[aA^{3/2} + (1-a)B^{3/2}] + \frac{a^2}{4}(A^2 - B^2)}{a(A+B) - \frac{2}{3}a(A^{3/2} + B^{3/2})} \quad (3)$$

при  $A \geq B$  и при  $A \leq B$ .

В частном случае при  $a = 1/2$  имеем

$$u_c = \frac{(A+3B)/8 - (A^{3/2} + B^{3/2})/6 + (A^2 - B^2)/16}{(A+B)/2 - (A^{3/2} + B^{3/2})/3} \quad (4)$$

при  $A \geq B$  и при  $A \leq B$ .

В качестве примера приведем следующие результаты расчетов:

- $a = 0,5: A = 0,2, B = 0,4 \Rightarrow u_c = 0,5942;$
- $A = 0,4, B = 0,2 \Rightarrow u_c = 0,4058.$
- $a = 0,25: A = 0,2, B = 0,4 \Rightarrow u_c = 0,6083;$
- $A = 0,4, B = 0,2 \Rightarrow u_c = 0,3917.$
- $a = 0,1: A = 0,2, B = 0,4 \Rightarrow u_c = 0,6167;$
- $A = 0,4, B = 0,2 \Rightarrow u_c = 0,3833.$

Блок сравнения величин  $A(t)$  и  $B(t)$  и расчета  $u_c$  для нечеткого регулятора с выходными функциями принадлежности, приведенными на рис. 2, показан на рис. 3.

При конструировании нечетких регуляторов на основе структурной схемы, приведенной на рис. 1, блок сравнения величин  $A(t)$  и  $B(t)$  и расчета  $u_c$  (блок 2), приведенный на рис. 1, может использоваться с различными блоками формирователей величин  $A(t)$  и  $B(t)$  (блоками 1 на рис. 1).

Блок сравнения величин  $A(t)$  и  $B(t)$  и расчета  $u_c$  для нечеткого регулятора с выходными функциями принадлежности, приведенными на рис. 2, получается достаточно простым. Если для этого блока (см. рис. 3) использовать наиболее простой формирователь величин  $A(t)$  и  $B(t)$  со входными треугольными функциями принадлежности, то получим простую схему нечеткого регулятора (см. рис. 4). Логика работы нечеткого регулятора (см. рис. 4) для фиксированного момента времени отображена на рис. 5.

В формирователе величин  $A(t)$  и  $B(t)$  на входе нечеткого регулятора переменные ошибка  $\theta^*$ , первая производная ошибки  $\dot{\theta}^*$  и вторая производная ошибки  $\ddot{\theta}^*$ ,

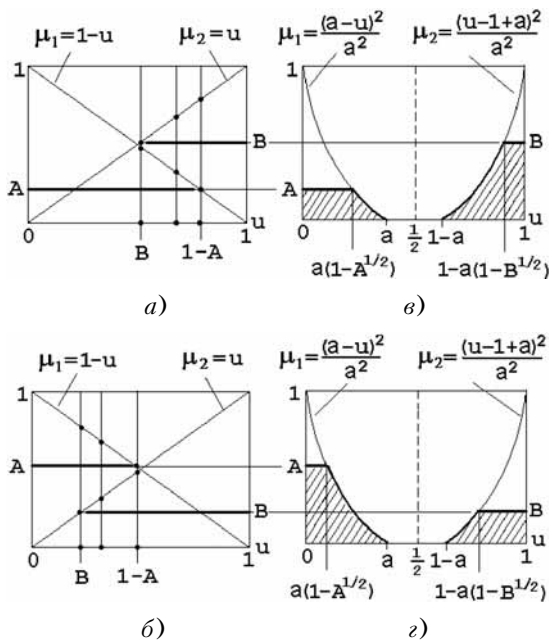


Рисунок 5

поступающие в регулятор с шагом квантования  $h$ , пересчитываются в переменные  $u_1^*$ ,  $u_2^*$ ,  $u_3^*$  по формулам

$$\left. \begin{aligned} u_1^* &= (\theta^* + A_m)/(2A_m); \\ u_2^* &= (\theta^* + B_m)/(2B_m); \\ u_3^* &= (\theta^* + C_m)/(2C_m) \end{aligned} \right\}$$

(это формулы для пересчета значений сигналов в значения элементов универсального множества  $U = [0, 1]$  при симметричных диапазонах изменения переменных) и производится расчет значений входных функций принадлежности (см. рис. 5, а, б)

$$\begin{aligned} \mu_1(u) &= 1 - u; \\ \mu_2(u) &= u; \\ 0 &\leq u \leq 1 \end{aligned}$$

для переменных  $u_1^*$ ,  $u_2^*$ ,  $u_3^*$ . По алгоритму Мамдани определяются величины  $A$  и  $B$  по формулам

$$\begin{aligned} A &= \min[\mu_1(u_1^*), \mu_1(u_2^*), \mu_1(u_3^*)], \\ B &= \min[\mu_2(u_1^*), \mu_2(u_2^*), \mu_2(u_3^*)]. \end{aligned}$$

В схеме *формирователя величин*  $A(t)$  и  $B(t)$  при настройке нечеткого регулятора перестраиваются граничные значения  $A_m = \theta_{\max} = -\theta_{\min}$ ,  $B_m = \theta_{\max} = -\theta_{\min}$ ,  $C_m = \theta_{\max} = -\theta_{\min}$ .

Значения  $A$  и  $B$  откладываются соответствующим образом на выходные функции принадлежности

$$\begin{aligned} \mu_1(u) &= \frac{(a-u)^2}{a^2}, \quad 0 \leq u \leq a; \\ \mu_2(u) &= \frac{(u-1+a)^2}{a^2}, \quad 1-a \leq u \leq 1 \end{aligned}$$

(см. рис. 5, в, г) и производится расчет ненормированного выхода регулятора по формуле (3) в блоке *сравнения величин*  $A(t)$  и  $B(t)$  и расчета  $u_c$ . Далее полученное значение  $u_c$  в блоке *нормировки выходной переменной* пересчитывается в выходное напряжение регулятора по формуле  $m^* = m_{\min}(1 - 2u_c)$ .

В динамике при изменении переменных  $u_1(t)$ ,  $u_2(t)$ ,  $u_3(t)$  текущие величины  $A(t)$  и  $B(t)$  поступают в блок *сравнения* и рассчитанное на каждом шаге  $h$  значение  $u_c(t)$  в блоке *нормировки выходной переменной* преобразуется в выходное напряжение  $m(t)$  по формуле  $m(t) = m_{\min}[1 - 2u_c(t)]$ . Блок *нормировки выходной переменной* (блок 3 на рис. 1) показан на рис. 4. В блоке *нормировки выходной переменной* перестраиваются граничные значения диапазона  $D_m = m_{\max} = -m_{\min}$ . Коэффициент  $a$  обычно задается постоянным.

## ЗАКЛЮЧЕНИЕ

Изложенные теория и практическая схема нечеткого регулятора с выходными сжатыми, зависящими от одного параметра, S-образной и Z-образной функциями принадлежности и входными треугольными функциями принадлежности дают возможность использовать такой регулятор в различных системах автоматического управления и путем настройки параметров регулятора добиваться высокого качества систем управления.

## ПЕРЕЧЕНЬ ССЫЛОК

1. Гостев В. И. Новый метод проектирования одного класса нечетких цифровых регуляторов // Проблемы управления и информатики. – 2007. – № 6. – С. 73–84.

Надійшла 13.10.2008

*Отримано аналітичні вирази для керуючих впливів на виході нечіткого регулятора при вихідних роз'єднаних стислих, залежних від одного параметра S-образної й Z-образної функціях належності й викладені питання проектування нечіткого регулятора.*

*Analytical expressions for controlling impact on a fuzzy controller input are received at outlet separated compressed depending on one parametre figurative and non-figurative membership functions and questions of designing of an fuzzy controller are stated.*

И. А. Орловский, Ю. С. Бут

## УПРАВЛЕНИЕ ВЗАИМОСВЯЗАННЫМ АСИНХРОННЫМ ЭЛЕКТРОПРИВОДОМ МОСТОВОГО КРАНА ДЛЯ ПЕРЕМЕЩЕНИЯ МОСТА БЕЗ ВЗАИМОДЕЙСТВИЯ РЕБОРД КОЛЕС С РЕЛЬСАМИ

*Для многокритериального управления взаимосвязанным многодвигательным электроприводом перемещения мостового крана, позволяющего обеспечить перемещение крана без касания ребордами рельсов, разработана структура регулятора общей обратной связи и получены аналитические выражения для расчета его параметров. Выполнено исследование разработанной системы методом математического моделирования.*

Важнейшей проблемой современного краностроения является проектирование грузоподъемных машин, обеспечивающих максимальную производительность при минимуме энергопотребления. Решению этой проблемы способствует применение многодвигательного взаимосвязанного асинхронного привода перемещения мостового крана (АППМК), позволяющего индивидуально воздействовать на каждое колесо механизма передвижения моста. Основной причиной значительного расхода энергии, а, следовательно, и износа колес являются обусловленные технологическими факторами их взаимные перекосы, вызывающие практически постоянный контакт реборд с рельсами [1]. Задача предотвращения или снижения интенсивности износа тесно связана с задачей обеспечения прямолинейной или близкой к ней траектории движения крана, с высокой точностью, совпадающей с направлением рельсового пути.

Основными факторами, способствующими возникновению сил перекоса и поперечных реакций рельсового пути, являются: установочный перекося ходовых колес в горизонтальной плоскости; смещение грузовой тележки от середины пролета; перекося ходовых колес вследствие упругой деформации моста; различие характеристик приводных двигателей для кранов с раздельным приводом и разница в диаметрах приводных колес кранов [1].

По нашему мнению, снизить влияние нежелательных факторов можно, применив многодвигательный взаимосвязанный АППМК с индивидуальным управлением всех четырех колес. Для этого необходимо разработать структуру и найти параметры системы управления (СУ).

© Орловский И. А., Бут Ю. С., 2009

### АНАЛИЗ ЛИТЕРАТУРНЫХ ИСТОЧНИКОВ

Исследованию построения СУ взаимосвязанных ЭП кранов в научно-технической литературе посвящен ряд работ. В [2] рассмотрен метод синтеза регуляторов электромеханических систем на основе концепции обратной задачи динамики. Однако такой метод не разработан для систем, у которых порядок объекта управления выше трех. В [3] применяется экспоненциальный закон управления моментом, который изменяет соответствующим образом задание скорости, что позволяет снизить динамические нагрузки на асинхронный электропривод механизма поворота крана. В [4] по передаточным функциям получены области устойчивости нелинейной двухмассовой электромеханической системы, состоящей из асинхронного двигателя с управлением по напряжению и с внутренней обратной связью по скорости. В [5] предлагается использовать фаззи-регулятор для оптимального управления электроприводом механизма передвижения крана по быстродействию без раскачивания груза. В [6] предлагается использовать нелинейный ПИД-регулятор, коэффициенты и структура которого меняется в зависимости от векторной ошибки. В этой системе изменение управляющего сигнала формируется с учетом модуля и фазы векторной ошибки на каждом такте регулирования. В статье [7] предложена система подчиненного управления с внешним контуром угла поворота крана и внутренним контуром перемещения центра крана в направлении, поперечном основному движению. Недостатком этой системы является то, что движение моста крана осуществляется с заданным предварительно найденным углом поворота моста, непрерывным переключением моментов двигателей из-за использования релейного регулирования с большими коэффициентами усиления и ограничением выходных сигналов.

В работах [2–7] вопросы настройки параметров СУ взаимосвязанного асинхронного привода перемещения мостового крана с индивидуальным управлением всех четырех колес по критерию минимального взаимодействия реборд колес с рельсами при не точно известных внутренних параметрах крана и возмущающих воздействий не рассматривались.



**ЦЕЛЬ СТАТЬИ**

Разработка структуры и методов расчета параметров СУ взаимосвязанного четырехдвигательного АППМК для обеспечения движения моста без взаимодействия реборд колес с рельсами с нулевым углом поворота при неточно известных внутренних параметрах крана и возмущающих воздействий.

**МАТЕМАТИЧЕСКАЯ МОДЕЛЬ ДВИЖЕНИЯ МОСТА КРАНА**

Рассматривается движение четырехколесного мостового крана с цилиндрическими колесами, с разными углами установочного перекоса колес (рис. 1). В математической модели, подробно изложенной в статье [7], мост имеет три степени свободы: две поступательные и одну вращательную. Движение моста рассматривается в неподвижной системе координат «*xу*». В кинематической схеме моста крана, показанной на рис. 1, приняты следующие обозначения:  $P_i$  – движущая или тормозная сила, развиваемая *i*-м колесом,  $i = 1, 2, 3, 4$  – номер колеса;  $W_i$  – сила сопротивления передвижению *i*-му колесу крана;  $\beta_i$  – угол установочного перекоса *i*-го колеса;  $m$  – масса крана;  $a$  – половина базы крана;  $b$  – расстояние от центра масс крана до колес (в рассматриваемой модели центр масс крана находится в центре моста);  $l$  – половина пролета крана;  $\alpha$  – угол конструкции моста ( $\alpha = \arctg(l/a)$ );  $\delta$  – допустимое отклонение колес по оси «*x*» до касания ребордами рельсов.

Движение крана без касания реборд рельсов осуществляется, если отклонение колес по оси «*x*» про-

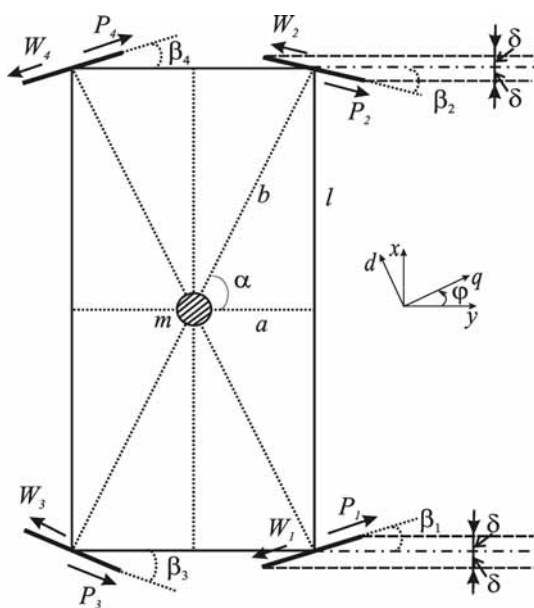


Рисунок 1 – Кинематическая схема моста крана

исходит в допустимом коридоре размером  $2\delta$ , показанном на рис. 1 штриховыми линиями.

В соответствии с классическими уравнениями механики, уравнения движения центра масс крана имеют вид [7]

$$\begin{cases} m\ddot{y} = \sum_{i=1}^4 (P_i - W_i) \cos(\beta_i + \varphi); \\ m\ddot{x} = \sum_{i=1}^4 (P_i - W_i) \sin(\beta_i + \varphi); \\ J\ddot{\varphi} = b(P_1 - W_1) \sin(\alpha + \beta_1) + b(P_2 - W_2) \sin(\beta_2 - \alpha) - \\ - b(P_3 - W_3) \sin(-\alpha + \beta_3) - b(P_4 - W_4) \sin(\alpha + \beta_4), \end{cases} \quad (1)$$

где  $x$  – перемещение центра моста крана по оси «*x*»;  $y$  – перемещение центра моста крана по оси «*y*»;  $\varphi$  – угол поворота моста крана;  $J$  – момент инерции крана относительно центра масс крана.

Силы  $P_i$ , развиваемые колесами, создаются через редукторы асинхронными двигателями (АД) с векторными СУ. С учетом этого, уравнения формирования сил  $P_i$  в операторном виде следующие [8]:

$$P_i = \left( \omega_{zi} - \left( \frac{k_c}{T_{\text{фс}} s + 1} \omega_i \right) \right) W_{\text{pc}} W_T \frac{3}{2} z k \Psi_{r(i)} \frac{k_p}{R_i}, \quad (2)$$

где  $\omega_{zi}$  и  $\omega_i$  – заданное и реальное значения угловой скорости *i*-го АД, соответственно;  $R_i$  – радиус *i*-го колеса;  $k_p$  – передаточное число редуктора;  $z$  – число пар полюсов АД;  $k$  – конструктивная постоянная АД;  $\Psi_{r(i)}$  – модуль вектора потокосцепления ротора *i*-го АД;  $W_T$  – передаточная функция замкнутого контура регулирования активной проекций статорного тока АД;  $W_{\text{pc}}$  – передаточная функция регулятора скорости;  $k_c$  – коэффициент передачи датчика скорости;  $T_{\text{фс}}$  – постоянная времени фильтра по скорости.

**ОБЩАЯ ФУНКЦИОНАЛЬНАЯ СУ ВЗАИМОСВЯЗАННОГО АППМК**

Функциональная схема СУ взаимосвязанного АППМК представлена на рис. 2.

В схеме используются следующие обозначения:  $РСi$  – регуляторы скорости;  $КРTi$  – контуры регулирования токов АД $i$ ;  $V_i$  – проекции линейных скоростей центров колес на плоскости колес;  $ЗИ$  – задатчик интенсивности;  $V_{xi}$  – проекции линейных скоростей колес на ось «*x*»;  $x_i$  – поперечные смещения центров колес (отсчитываемые от средней линии рельса в направлении оси «*x*»).

Общая СУ включает в себя векторные СУ каждого АД и регулятор общей обратной связи (РООС). РООС состоит из трех блоков (рис. 2): блока идентификации параметров, блока коррекции движения по модели объекта, блока формирования требуемых

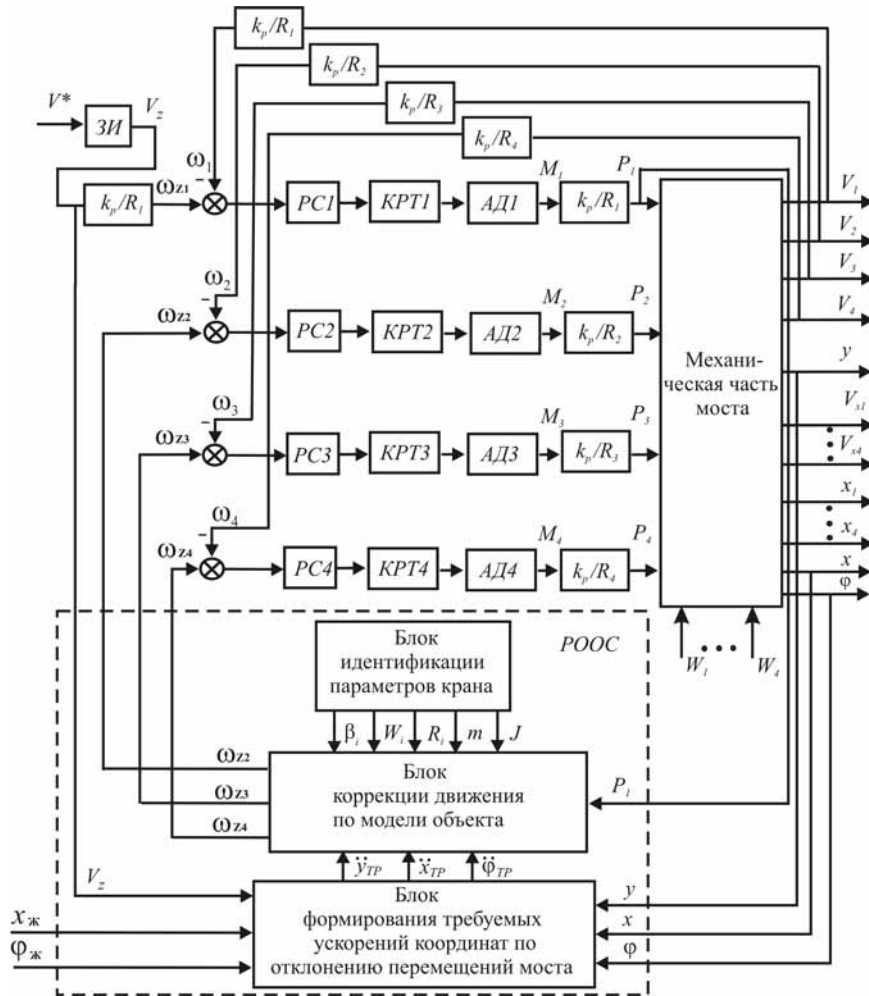


Рисунок 2 – Функциональная схема четырехдвигательного взаимосвязанного АППМК

ускорений координат по отклонению перемещений моста.

Задачей блока идентификации параметров является определение в процессе работы крана значений внутренних параметров АППМК и сил сопротивления колесам моста.

Блок коррекции движения по модели объекта вырабатывает сигналы коррекции заданий угловых скоростей индивидуальных ЭП в зависимости от значения силы  $P_1$  – первого колеса и известных математических моделей АППМК и возмущающих воздействий.

Блок формирования требуемых ускорений координат по отклонению перемещений моста содержит регуляторы координат и обеспечивает выработку требуемых ускорений регулируемых координат.

В работе [9] теоретически исследована возможность (при полностью известной математической модели крана и возмущающих сил) обеспечения про-

дольного движения моста без касания ребрами колес рельсов при отсутствии угла поворота моста относительно заданного направления. На рассмотренных в статье примерах удалось получить очень высокую точность (до  $10^{-13}$  м по « $x$ » и  $2,5 \cdot 10^{-3}$  рад по « $\phi$ ») регулирования перемещений по осям « $x$ » и « $y$ », и углу « $\phi$ » в течение 150 с со скоростью движения по оси « $y$ » 2 м/с.

Достигнутая точность движения обеспечивается блоком коррекции движения по модели объекта, алгоритм работы которого следующий. Задание скорости моста  $V_z$  по оси « $y$ » подается только на ЭП первого двигателя, который создает вращающий момент  $M_1$  и, следовательно, силу  $P_1$  первого колеса. Управление тремя другими двигателями осуществляется так, чтобы на приводимых в движение ими колесах создавались силы  $P_2, P_3, P_4$ , обеспечивающие решение следующей системы уравнений:

$$\begin{cases} m\ddot{y}_{TP} = \sum_{i=1}^4 (P_i - W_i) \cos(\beta_i + \varphi); \\ m\ddot{x}_{TP} = \sum_{i=1}^4 (P_i - W_i) \sin(\beta_i + \varphi); \\ J\ddot{\varphi}_{TP} = b(P_1 - W_1) \sin(\alpha + \beta_1) + b(P_2 - W_2) \sin(\beta_2 - \alpha) - \\ - b(P_3 - W_3) \sin(-\alpha + \beta_3) - b(P_4 - W_4) \sin(\alpha + \beta_4), \end{cases} \quad (3)$$

где  $\ddot{y}_{TP}$ ,  $\ddot{x}_{TP}$ ,  $\ddot{\varphi}_{TP}$  – требуемые значения ускорений по осям « $x$ » и « $y$ » и углу « $\varphi$ », соответственно.

Для упрощения записи решения системы (3) примем следующие обозначения:

$$\begin{cases} A = m\ddot{y}_{TP}, B = m\ddot{x}_{TP}, C = J\ddot{\varphi}_{TP}, P_{11} = P_1 - W_1, \\ P_{22} = P_2 - W_2, P_{33} = P_3 - W_3, P_{44} = P_4 - W_4; \\ c_1 = \cos(\beta_1 + \varphi), c_2 = \cos(\beta_2 + \varphi), \\ c_3 = \cos(\beta_3 + \varphi), c_4 = \cos(\beta_4 + \varphi); \\ s_1 = \sin(\beta_1 + \varphi), s_2 = \sin(\beta_2 + \varphi), \\ s_3 = \sin(\beta_3 + \varphi), s_4 = \sin(\beta_4 + \varphi); \\ b_1 = b \sin(\alpha + \beta_1), b_2 = b \sin(-\alpha + \beta_2), \\ b_3 = (-b) \sin(-\alpha + \beta_3), b_4 = (-b) \sin(\alpha + \beta_4). \end{cases} \quad (4)$$

С учетом этих обозначений система (3) примет вид

$$\begin{cases} A - P_{11} \cdot c_1 = P_{22} \cdot c_2 + P_{33} \cdot c_3 + P_{44} \cdot c_4; \\ B - P_{11} \cdot s_1 = P_{22} \cdot s_2 + P_{33} \cdot s_3 + P_{44} \cdot s_4; \\ C - P_{11} \cdot b_1 = P_{22} \cdot b_2 + P_{33} \cdot b_3 + P_{44} \cdot b_4. \end{cases} \quad (5)$$

Система (5) состоит из линейных алгебраических уравнений и ее решение можно выполнить, используя правило Крамера. Из определителей системы (5)

$$\begin{cases} \Delta = \begin{vmatrix} c_2 & c_3 & c_4 \\ s_2 & s_3 & s_4 \\ b_2 & b_3 & b_4 \end{vmatrix}; & \Delta_2 = \begin{vmatrix} A - P_{11} \cdot c_1 & c_3 & c_4 \\ B - P_{11} \cdot s_1 & s_3 & s_4 \\ C - P_{11} \cdot b_1 & b_3 & b_4 \end{vmatrix}; \\ \Delta_3 = \begin{vmatrix} c_2 & A - P_{11} \cdot c_1 & c_4 \\ s_2 & B - P_{11} \cdot s_1 & s_4 \\ b_2 & C - P_{11} \cdot b_1 & b_4 \end{vmatrix}; & \Delta_4 = \begin{vmatrix} c_2 & c_3 & A - P_{11} \cdot c_1 \\ s_2 & s_3 & B - P_{11} \cdot s_1 \\ b_2 & b_3 & C - P_{11} \cdot b_1 \end{vmatrix} \end{cases} \quad (6)$$

находятся значения сил, развиваемые вторым, третьим и четвертым колесами в виде

$$\begin{cases} P_2 = \frac{\Delta_2}{\Delta} = \frac{(A - P_{11}c_1)s_3b_4 + c_3s_4(C - P_{11}b_1) + c_4(B - P_{11}s_1)b_3 - c_4s_3(C - P_{11}b_1) - c_3(B - P_{11}s_1)b_4 - (A - P_{11}c_1)s_4b}{c_2s_3b_4 + c_3s_4b_2 + c_4s_2b_3 - c_4s_3b_2 - c_3s_2b_4 - c_2s_4b_3}; \\ P_3 = \frac{\Delta_3}{\Delta} = \frac{c_2(B - P_{11}s_1)b_4 + (A - P_{11}c_1)s_4b_2 + c_4s_2(C - P_{11}b_1) - c_4(B - P_{11}s_1)b_2 - (A - P_{11}c_1)s_2b_4 - c_2s_4(C - P_{11}b_1)}{c_2s_3b_4 + c_3s_4b_2 + c_4s_2b_3 - c_4s_3b_2 - c_3s_2b_4 - c_2s_4b_3}; \\ P_4 = \frac{\Delta_4}{\Delta} = \frac{c_2s_3(C - P_{11}b_1) + c_3(B - P_{11}s_1)b_2 + (A - P_{11}c_1)s_2b_3 - (A - P_{11}c_1)s_3b_2 - c_3s_2(C - P_{11}b_1) - c_2(B - P_{11}s_1)b_3}{c_2s_3b_4 + c_3s_4b_2 + c_4s_2b_3 - c_4s_3b_2 - c_3s_2b_4 - c_2s_4b_3}. \end{cases} \quad (7)$$

Для получения вычисленных значений сил  $P_2$ ,  $P_3$ ,  $P_4$ , необходимо сформировать соответствующие им задания скорости двигателей. Эти задания, согласно уравнениям (2) (при отсутствии ограничений в регуляторах), находятся следующим образом:

$$\begin{cases} \omega_{z2} = \left( \left( P_2 \frac{R_2}{k_p} \frac{1}{\frac{3}{2} z k \Psi_{r(i)}} \frac{1}{W_T} \frac{1}{W_{pc}} \right) + \left( \frac{k_c}{T_{\varphi c} s + 1} \omega_2 \right) \right); \\ \omega_{z3} = \left( \left( P_3 \frac{R_3}{k_p} \frac{1}{\frac{3}{2} z k \Psi_{r(i)}} \frac{1}{W_T} \frac{1}{W_{pc}} \right) + \left( \frac{k_c}{T_{\varphi c} s + 1} \omega_3 \right) \right); \\ \omega_{z4} = \left( \left( P_4 \frac{R_4}{k_p} \frac{1}{\frac{3}{2} z k \Psi_{r(i)}} \frac{1}{W_T} \frac{1}{W_{pc}} \right) + \left( \frac{k_c}{T_{\varphi c} s + 1} \omega_4 \right) \right). \end{cases} \quad (8)$$

При наличии неточности параметров в модели крана и РООС, а также при наличии первоначальных смещений по « $x$ » и « $\varphi$ », необходимо обеспечивать требуемые значения ускорений  $\ddot{x}_{TP}$  и  $\ddot{\varphi}_{TP}$  для перемещения моста без взаимодействия реборд колес с рельсами. Поэтому требуемые значения ускорений  $\ddot{y}_{TP}$ ,  $\ddot{x}_{TP}$ ,  $\ddot{\varphi}_{TP}$  формируются с использованием измерения значений регулируемых координат блоком формирования требуемых ускорений координат по следующим формулам:

$$\begin{cases} \ddot{y}_{TP} = K_y(V_z - V); \\ \ddot{x}_{TP} = (x_{ж} - x)K_x + (\dot{x}_{ж} - \dot{x})K_{xv}; \\ \ddot{\varphi}_{TP} = (\varphi_{ж} - \varphi)K_{\varphi} + (\dot{\varphi}_{ж} - \dot{\varphi})K_{\varphi v}, \end{cases} \quad (9)$$

где  $K_y$  – коэффициент усиления ошибки скорости по координате « $y$ »;  $K_x$ ,  $K_{xv}$  – коэффициенты усиления ошибки, соответственно, положения и скорости по координате « $x$ »;  $x_{ж}$ ,  $\dot{x}_{ж}$  – желаемые значения положения и скорости по координате « $x$ »;  $K_{\varphi}$ ,  $K_{\varphi v}$  – коэффициенты усиления ошибки по углу поворота и угловой скорости поворота крана;  $\varphi_{ж}$ ,  $\dot{\varphi}_{ж}$  – желаемые значения угла поворота и угловой скорости поворота крана.

Предполагается, что на кране установлены ультразвуковые датчики для измерения положений колес относительно рельсов по оси « $x$ », а также датчик, измеряющий угол поворота моста.

Во время движения крана происходит непрерывный пересчет требуемых значений ускорений регулируемых координат  $\ddot{y}_{ТР}$ ,  $\ddot{x}_{ТР}$ ,  $\ddot{\phi}_{ТР}$  с использованием которых по уравнениям систем (4) и (7) вычисляются необходимые значения сил второго, третьего и четвертого приводных колес.

### РЕЗУЛЬТАТЫ МОДЕЛИРОВАНИЯ

При моделировании АППМК в системе MATLAB были заданы следующие параметры реального мостового крана (грузоподъемностью 20/5 т):  $m = 47200$  кг;  $J = 2,21 \cdot 10^6$  кгм<sup>2</sup>;  $l = 14,25$  м;  $a = 2,5$  м;  $b = 14,47$  м;  $\alpha = 0,38$  рад; углы установочного перекося колес:  $\beta_1 = 0,009$  рад,  $\beta_2 = -0,005$  рад,  $\beta_3 = -0,003$  рад,  $\beta_4 = 0,003$  рад. Используются двигатели 4A132S6Y3 со следующими параметрами:  $P_{д.н} = 5,5$  кВт,  $U_{снф} = 220$  В,  $I_{снф} = 10,4$  А,  $\omega_n = 100,7$  с<sup>-1</sup>,  $M_n = 49,2$  Нм,  $z = 3$ , амплитуда номинального фазного потокоцепления ротора  $\Psi_{rn} = 0,889$  Вб,  $k_p = 14$ . Задавались различные значения сил сопротивления колесам крана:  $W_1 = 800$  Н,  $W_2 = 640$  Н,  $W_3 = 720$  Н,  $W_4 = 880$  Н.

Для моделирования имеющих различия в радиусах колес изменялся радиус первого колеса на 10 % относительно радиусов других колес ( $R_1 = 0,385$  м,  $R_2 = R_3 = R_4 = 0,35$  м). Максимальное отклонение до взаимодействия реборд колес с рельсами по оси «x» составляет  $\delta = 15$  мм. Моделирование АППМК выполнялось с фиксированным шагом 0,001 с. Задание скорости моста приведено на рис. 3.

Исследовались следующие условия работы модели крана:

1) параметры модели и силы сопротивления колесам точно известны, начальное смещение центра масс крана и угла поворота крана отсутствует, используется только блок коррекции движения по модели объекта;

2) параметры модели и силы сопротивления колесам точно известны, имеется начальное смещение центра масс крана и угла поворота крана, используется блок формирования требуемых ускорений координат;

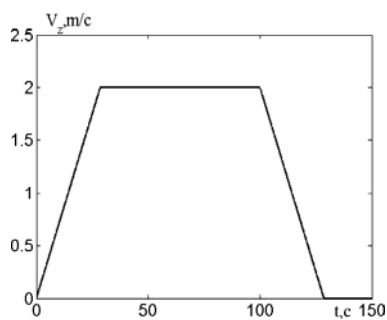


Рисунок 3 – Задание скорости моста

3) параметры модели и силы сопротивления колесам не точно известны, имеется начальное смещение центра масс крана и угла поворота крана, рассматривается источник энергии бесконечной мощности, используется блок коррекции движения по модели объекта;

4) параметры модели и силы сопротивления колесам не точно известны, имеется начальное смещение центра масс крана и угла поворота крана, рассматривается источник энергии бесконечной мощности, используются все блоки РООС;

5) параметры модели и силы сопротивления колесам не точно известны, имеется начальное смещение центра масс крана и угла поворота крана, рассматривается источник энергии конечной мощности, токи двигателей не превышают допустимые значения, используются все блоки РООС.

Значения коэффициентов регуляторов, найденных с помощью исследования на модели, имеют следующие значения:  $K_y = 9440$ ,  $K_x = 700$ ,  $K_{xv} = 700$ ,  $K_\phi = 700$ ,  $K_{\phi v} = 700$ . На рис. 4 и 5 представлены результаты моделирования движения крана. Различные условия работы модели крана представлены на рисунках отдельными столбцами, которые обозначены цифрами «1», «2», «3», «4», «5».

Для первого условия, при отсутствии начального смещения центра масс крана и угла поворота крана и использовании только блока коррекции движения по модели объекта (рис. 4, столбец 1) угол поворота моста в течение 150 с составляет  $\phi = 2,6 \cdot 10^{-14}$  рад. Поперечные смещения центров колес составляют  $0,7 \cdot 10^{-13}$  м. Линейная скорость центра масс крана по оси «y» достигает заданного значения 2 м/с.

Для второго условия, при наличии начального смещения центра масс крана 0,02 м и угла поворота крана 0,001 рад, использовании блока формирования требуемых ускорений координат (рис. 4, столбец 2) угол поворота моста составляет  $\phi = 10^{-3}$  рад. Поперечные смещения центров колес находятся в пределах 0,015 м. Линейная скорость центра масс крана по оси «y» достигает заданного значения 2 м/с. При таком регулировании появляются отрицательные активные составляющие токов, следовательно, и отрицательные движущие (тормозные) силы.

Для третьего условия, при начальном смещении центра масс крана 0,02 м и начальном угле поворота моста 0,001 рад, для заданных параметров крана происходит значительный поворот моста и смещение центров колес, что приводит к взаимодействию реборд колес с рельсами (рис. 5, столбец 3). Линейная скорость центра масс крана по оси «y» достигает заданного значения 2 м/с.

Для четвертого условия (рис. 5, столбец 4), при начальном смещении центра масс крана 0,02 м и начальном угле поворота моста 0,001 рад, использовании всех блоков РООС, угол поворота моста

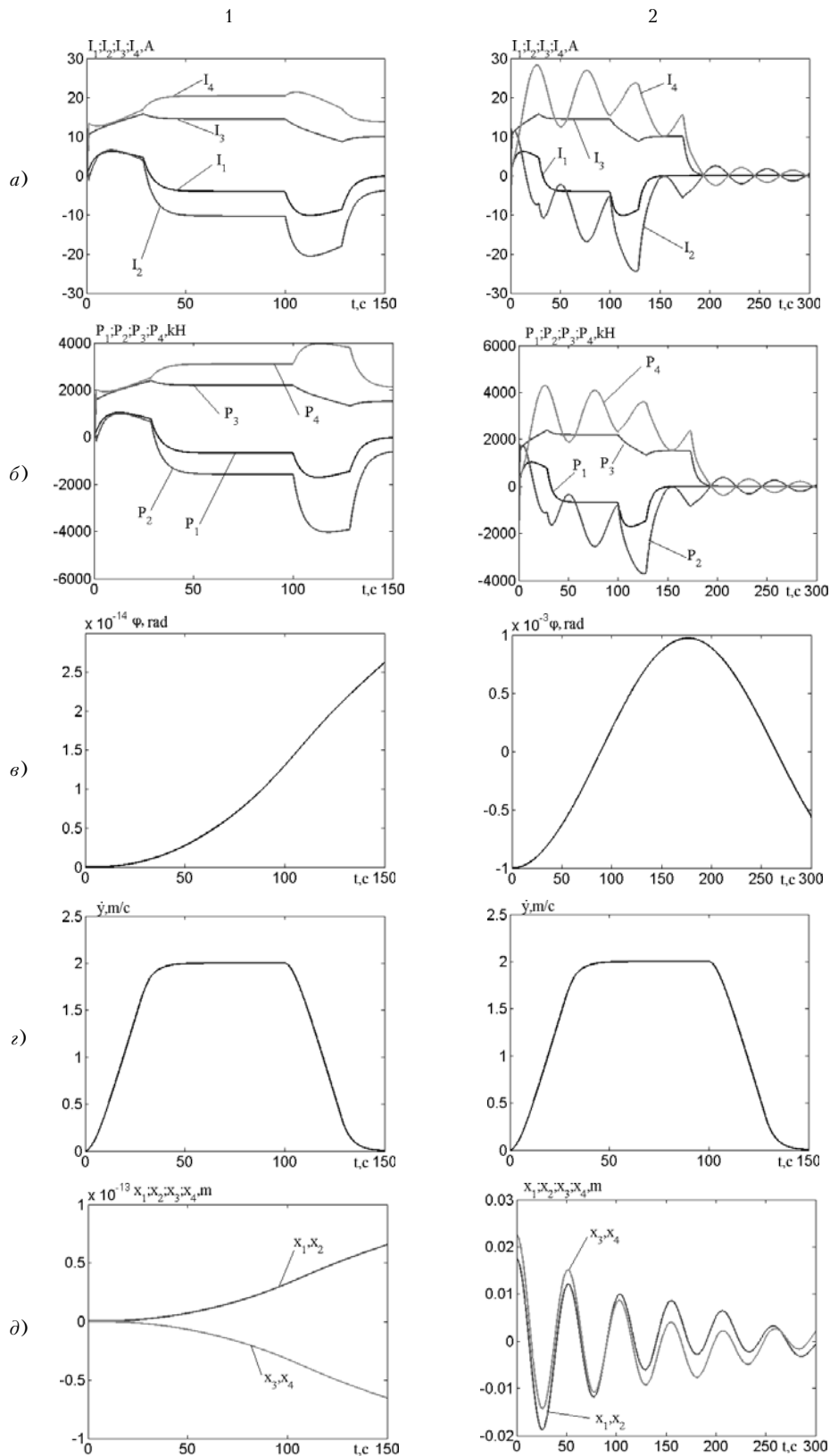


Рисунок 4 – Результаты моделирования движения крана при точно известных параметрах модели и сил сопротивления колесам:

*a* – активные составляющие токов двигателей; *b* – движущие или тормозные силы, развиваемые *i*-м колесом; *c* – угол изгиба моста в горизонтальной плоскости; *d* – линейная скорость центра масс крана по оси «*y*»; *e* – поперечные смещения центров колес от средней линии рельса

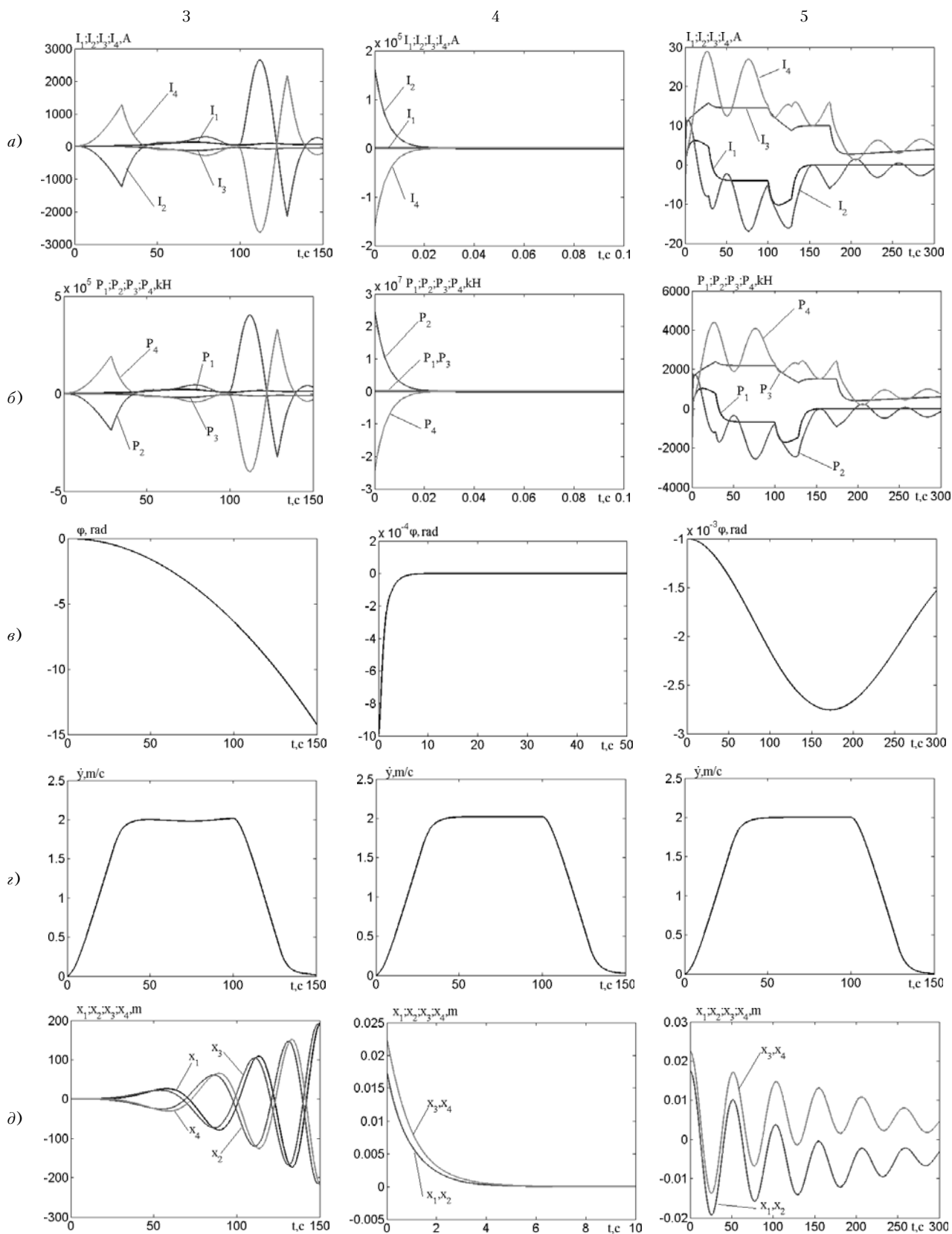


Рисунок 5 – Результаты моделирования движения крана при не точно известных параметрах модели и сил сопротивления колесам:

*a* – активные составляющие токов двигателей; *б* – движущие или тормозные силы, развиваемые *i*-м колесом; *в* – угол изгиба моста в горизонтальной плоскости; *г* – линейная скорость центра масс крана по оси «*y*»; *д* – поперечные смещения центров колес от средней линии рельса

и поперечные смещения центров колес отсутствуют (менее  $10^{-13}$  рад). Линейная скорость центра масс крана по оси «у» достигает заданного значения 2 м/с.

Для пятого условия, при начальном смещении центра масс крана 0,02 м и начальном угле поворота моста 0,001 рад, источнике энергии конечной мощности (токи двигателей не превышают допустимых значений) используются все блоки РООС (рис. 5, столбец 5). При этом угол поворота моста составляет  $\varphi = 2,6 \cdot 10^{-3}$  рад, поперечные смещения центров колес находятся в пределах 0,015 м, линейная скорость центра масс крана по оси «у» достигает заданного значения 2 м/с, появляются отрицательные активные составляющие токов и, следовательно, отрицательные тормозные силы.

## ВЫВОДЫ

1. Предложенная структура РООС и аналитические выражения для расчета его параметров позволили выполнить многокритериальное управление многодвигательным электроприводом перемещения мостового крана.

2. Результаты моделирования разработанной СУ взаимосвязанным асинхронным электроприводом мостового крана показали, что ее применение позволяет регулировать через векторные СУ АД поперечные смещения центров колес от средней линии рельса в пределах 15 мм при не точно известных внутренних параметрах крана, возмущающих воздействий, различных углах установочного перекося колес, моментов сопротивления и изменении радиусов колес на 10 %.

## ПЕРЕЧЕНЬ ССЫЛОК

1. Лобов Н. А. Динамика грузоподъемных кранов / Н. А. Лобов. – М. : Машиностроение, 1987. – 160 с.
2. Островерхов М. Я. Метод синтезу регуляторів електромеханічних систем на основі концепції зворотних задач динаміки в поєднанні з мінімізацією локальних функціоналів миттєвих значень енергії руху / М. Я. Островерхов // Проблеми автоматизованого електропривода. Теорія і практика : вестник Харківського політехнічного університету. 2008. – Вып. 30. – С. 105–110. – (Серія «Електротехніка, електроніка і електропривод»).
3. Герасимьяк Р. П. Управление асинхронным электроприводом механизмов поворота, обеспечивающие снижение динамических нагрузок / Р. П. Герасимьяк, Е. В. Найденко // Проблеми автоматизованого

электропривода. Теория и практика Проблеми автоматизованого електропривода. Теорія і практика : вестник Харківського політехнічного університету. – 2008. – Вып. 30. – С. 111–112. – (Серія «Електротехніка, електроніка і електропривод»).

4. Герасимьяк Р. П. Электромеханическая система подъемных механизмов с замкнутым асинхронным электроприводом / Р. П. Герасимьяк, В. Х. Нгуен, В. А. Лещев // Проблеми автоматизованого електропривода. Теорія і практика : вестник Харківського політехнічного університету. – 2008. – Вып. 30. – С. 326–328. – (Серія «Електротехніка, електроніка і електропривод»).
5. Панкратов А. И. Оптимальное управление электроприводом механизма передвижения крана с учетом изменения длины каната / А. И. Панкратов, А. Ф. Залаятов // Проблеми автоматизованого електропривода : вестник Харківського політехнічного університету. – 2008. – Вып. 30. – С. 221–222. – (Серія «Електротехніка, електроніка і електропривод»).
6. Агамалов О. Н. Альтернативне нелінійне ПІД – управління з використанням векторної помилки [Текст] / О. Н. Агамалов, Н. П. Лукаш // Електротехніка. Електроніка. Електроніка. – 2008. – № 2. – С. 8–13.
7. Орловский И. А. Математическая модель взаимосвязанного электропривода перемещения мостового крана с общей системой управления / И. А. Орловский, Ю. С. Бут // Вісник Кременчуцького державного політехнічного університету ім. М. Остроградського. – Ч. 1. – Кременчук, 2008. – Вып. 4 (51). – С. 145–149.
8. Пивняк Г. Г. Современные частотно-регулируемые асинхронные электроприводы с широтно-импульсной модуляцией / Г. Г. Пивняк, А. В. Волков – Днепропетровск : Национ. горный университет, 2006. – 470 с.
9. Орловский И. А. Многокритериальное управление перемещением мостового крана из его математической модели / И. А. Орловский, Ю. С. Бут // Вісник Кременчуцького державного політехнічного університету ім. М. Остроградського. – Ч. 1. – Кременчук, 2009. – Вып. 4 (57). – С. 21–24.

Надійшла 27.04.2009  
Після доробки 14.05.2009

*Для багатокритеріального керування взаємозалежним багатодвигунним електроприводом переміщення мостового крана, що дозволяє забезпечити переміщення крана без торкання ребрами рейок, розроблена структура регулятора загального зворотного зв'язку й отримані аналітичні вирази для розрахунку його параметрів. Виконано дослідження розробленої системи методом математичного моделювання.*

*For multicriteria control of the interconnected multi-impellent electric drive of moving of the bridge crane, allowing to provide moving the crane without a contact with flanges of rails, the structure of a regulator of the general feedback is developed and analytical expressions for calculation of its parameters are received. Research of the developed system by a method of mathematical modeling is executed.*

---

# АВТОРЫ НОМЕРА

*Авраменко Валерий Павлович*

доктор технических наук, профессор  
Харьковского национального университета  
радиоэлектроники

*Алексин Сергей Геннадиевич*

аспирант Днепропетровского национального университета им. Олеса Гончара

*Архипов Александр Евгеньевич*

доктор технических наук, профессор  
Национального технического университета Украины  
«Киевский политехнический институт»

*Архипова Елена Александровна*

магистрант Национального технического университета Украины «КПИ»

*Бабич Андрей Викторович*

аспирант Запорожского национального технического университета

*Безрук Валерий Михайлович*

доктор технических наук, заведующий кафедрой  
Харьковского национального университета радиоэлектроники

*Белюсов Евгений Леонидович*

ассистент Запорожского национального технического университета

*Бодянский Евгений Владимирович*

доктор технических наук, профессор  
Харьковского национального университета радиоэлектроники

*Бут Юлия Сергеевна*

аспирантка Запорожского национального технического университета

*Буханько Александр Николаевич*

аспирант Харьковского национального университета радиоэлектроники

*Васютин Евгений Васильевич*

кандидат физико-математических наук, ассистент  
Запорожского национального технического университета

*Волкова Валентина Владимировна*

аспирант Харьковского национального университета радиоэлектроники

*Воропай Алексей Юрьевич*

кандидат технических наук, старший преподаватель  
Запорожского национального технического университета

*Гребенник Игорь Валериевич*

доктор технических наук, профессор  
Харьковского национального университета радиоэлектроники

*Гриценко Юрий Алексеевич*

магистрант Харьковского национального университета радиоэлектроники

*Гоменюк Сергей Иванович*

доктор технических наук, декан Запорожского национального университета

*Головач Григорий Петрович*

кандидат физико-математических наук, доцент  
Киевского национального университета им. Тараса Шевченко

*Гостев Владимир Иванович*

доктор технических наук, заведующий кафедрой  
Государственного университета информационных технологий, г. Киев

*Дибче Гассан*

аспирант Харьковского национального университета радиоэлектроники

*Дробахин Олег Олегович*

доктор физико-математических наук, заведующий кафедрой  
Днепропетровского национального университета им. Олеса Гончара

*Дубровин Валерий Иванович*

кандидат технических наук, профессор  
Запорожского национального технического университета



---

*Дуравкин Евгений Владимирович*

кандидат технических наук, доцент  
Харьковского национального университета радиоэлектроники

*Ибрагимов Олег Махмудович*

ведущий инженер Запорожского филиала концерна РРТ, г. Запорожье

*Иванов Денис Валерьевич*

студент Харьковского национального университета радиоэлектроники

*Журавлев Владимир Николаевич*

кандидат технических наук, докторант  
Национального технического университета Украины «КПИ»

*Заболотный Петр Иванович*

младший научный сотрудник Института технической механики НАН и НКА Украины, г. Днепропетровск

*Зависляк Игорь Владимирович*

доктор физико-математических наук, профессор Киевского национального университета им. Тараса Шевченко

*Кабак Владислав Семенович*

кандидат технических наук, декан Запорожского национального технического университета

*Казурова Алина Евгеньевна*

аспирант Запорожского национального технического университета

*Калачова Вероника Валерьевна*

кандидат технических наук, старший научный сотрудник Харьковского университета воздушных сил

*Касьян Константин Николаевич*

кандидат технических наук, доцент Запорожского национального технического университета

*Киприч Татьяна Витальевна*

кандидат технических наук, инженер-программист ГП «Ивченко-Прогресс», г. Запорожье

*Коваль Клавдия Владимировна*

студентка Харьковского национального университета радиоэлектроники

*Коляда Михаил Георгиевич*

кандидат педагогических наук, доцент  
Донецкого национального технического университета

*Крищук Владимир Николаевич*

кандидат технических наук, заведующий кафедрой Запорожского национального технического университета

*Кучеренко Дария Ефимовна*

стажер-исследователь Харьковского национального университета радиоэлектроники

*Лебедев Сергей Геннадиевич*

аспирант Днепропетровского национального университета им. Олеса Гончара

*Лисняк Андрей Александрович*

аспирант Запорожского национального университета

*Литвин Василий Владимирович*

кандидат технических наук, доцент Национального университета «Львовская политехника»

*Малый Александр Юрьевич*

ассистент Запорожского национального технического университета

*Морщавка Сергей Владимирович*

кандидат технических наук, доцент Запорожского национального технического университета

*Олейник Алексей Александрович*

ассистент Запорожского национального технического университета

*Орловский Игорь Анатольевич*

кандидат технических наук, доцент Запорожского национального технического университета

---

*Пица Дмитрий Макарович*

доктор технических наук, проректор  
Запорожского национального технического университета

*Погосов Валентин Вальтерович*

доктор физико-математических наук,  
заведующий кафедрой Запорожского  
национального технического университета

*Попов Максим Александрович*

аспирант Киевского национального университета им. Тараса Шевченко

*Потапенко Евгений Михайлович*

доктор технических наук, профессор Запорожского национального технического университета

*Привалов Евгений Николаевич*

кандидат физико-математических наук,  
заведующий отделом Института технической механики НАН и НКА Украины, г. Днепропетровск

*Притула Анатолий Викторович*

кандидат технических наук, доцент Запорожского национального технического университета

*Романенко Игорь Александрович*

кандидат военных наук, заместитель  
главнокомандующего Воздушных сил  
ВС Украины

*Самер Альмадхоун*

ассистент Харьковского национального университета радиотехники

*Скрупский Степан Юрьевич*

студент Запорожского национального технического университета

*Столбов Владимир Фридрихович*

кандидат юридических наук, начальник  
кафедры Института подготовки  
ответственных кадров для службы безопасности Украины им. Я. Мудрого, г. Харьков

*Сыревич Евгения Ефимовна*

кандидат технических наук, доцент  
Харьковского национального университета радиотехники

*Субботин Сергей Александрович*

кандидат технических наук, доцент  
Запорожского национального технического университета

*Терновой Максим Юрьевич*

кандидат технических наук, доцент  
Национального технического университета Украины «Киевский политехнический институт»

*Удовенко Сергей Григорьевич*

доктор технических наук, профессор  
Харьковского национального университета радиотехники

*Хабаров Александр Юрьевич*

кандидат технических наук, научный  
сотрудник производственно-коммерческой фирмы «РиКо», г. Феодосия

*Харитонов Виктор Николаевич*

инженер-конструктор ГП «Ивченко-Прогресс», г. Запорожье

*Чумаченко Виталий Павлович*

доктор физико-математических наук,  
заведующий кафедрой Запорожского  
национального технического университета

*Чумаченко Ярослав Виталиевич*

кандидат технических наук, доцент  
Запорожского национального технического университета

*Шамраев Анатолий Анатольевич*

кандидат технических наук, доцент  
Харьковского национального университета радиотехники

*Шкиль Александр Сергеевич*

кандидат технических наук, доцент  
Харьковского национального университета радиотехники

# ДЛЯ СВЕДЕНИЯ АВТОРОВ

Журнал «Радиоелектроніка, інформатика, управління» (PIU) призначений для публікації найбільш значимих наукових і практичних результатів досліджень учених вищих навчальних закладів і наукових організацій.

Журнал включений в перелік наукових видань України, в яких можуть публікуватися результати дисертаційних робіт на соискание учених ступенів доктора і кандидата технічних наук і фізико-математических наук по спеціальності «Радиофізика».

Статті, опубліковані в журналі, реферуються в реферативних журналах і базах даних ВІНИТИ (Росія) і «Джерело» (Україна). Інтернет-сторінка журналу:

<http://journal.zntu.edu.ua/ric/index.php?page=index>.

Журнал видається два рази в рік і розповсюджується по підписці в Україні. Підписний індекс 22914.

**Послідовність розміщення матеріала статті:** індекс УДК, назва статті, ініціали і прізвище автора(ів), анотація, текст, перелік посилань. Рукопис статті повинен бути підписаний всіма авторами.

## Текст статті

Приймаються статті на російському, українському і англійському мовах. Розмір статті до 0,5 авторського листа.

## Структура статті

Редакційна колегія приймає до розгляду тільки наукові статті, які містять такі необхідні елементи: постановка проблеми в загальному вигляді і її зв'язок з важливими науковими або практичними завданнями; аналіз останніх досліджень і публікацій, в яких почато рішення даної проблеми, і на які опирається автор; виділення нерешених раніше частин загальної проблеми, котрим присвячується стаття; формулювання цілей статті (постановка задачі); виклад основного матеріалу дослідження з повним обґрунтуванням отриманих наукових результатів, висновки по даному дослідженню і перспективи подальших досліджень в даному напрямку.

## Требования к оформлению

Статті приймаються набрані в редакторі Word for Windows (v.6 і вище).

Параметри сторінки:

- розмір паперу – А4 (210x297);
- орієнтація – книжна;
- шрифт – Times New Roman Cyr, розмір – 12 pt;
- міжрядковий інтервал – полуторний;
- верхнє поле – 20мм, нижнє – 20 мм, лівє – 25 мм, правє – 15 мм.

## Анотація

Анотація є обов'язковою і подається в початку статті на російському, українському (для громадян Ук-

раїни) і англійському мовах. Об'єм анотації не повинен перевищувати 40 слів.

## Рисунки

Рисунки розміщуються в тексті і подаються в окремих файлах (формат.TIF з дозволом не менше 200 dpi). Рисунки нумерують і підписують знизу.

## Таблиці

Заголовки таблиць розміщуються по центру.

## Формули

Формули виконуються з допомогою вбудованого в Word for Windows редактора Microsoft Equation. Формули нумерують в круглих дужках справа:

$$Z(\Theta_{\sim}) = 101 \log \left( \frac{y_2}{y_1} \right). \quad (3)$$

Формули великого розміру записуються в декількох рядках.

## Перелік посилань

Перелік посилань в кінці рукопису подається мовою оригіналу і складається в порядку згадки посилань в тексті і в відповідності до стандарту на бібліографічне описання. Посилання на літературу в тексті позначаються цифрою в квадратних дужках.

**В свіденнях об авторах** необхідно привести прізвище, ім'я, по батькові (повністю), місце роботи, посада, науковий ступінь, робочий і домашній адреса, телефони, e-mail.

## В редакцію журналу необхідно представити:

- 1) надрукований текст статті в 2-х екземплярах;
- 2) експертне обґрунтування про можливість опублікування;
- 3) свідення об авторах;
- 4) супровідне листування-заявка з організації, де була виконана робота;
- 5) назва статті, прізвище, імена і по батькові авторів на англійському мові;
- 6) дискету 3,5" з текстом документа, файлами рисунків, даними про статтю на англійському мові і свіденнями об авторах в форматі Word for Windows (v.6 і вище). Файли можна надіслати електронною поштою в форматі заархівованих (ZIP-архіватором) файлів.

*Статті, які не відповідають вказаним вимогам, не приймаються до розгляду.*

*Рукописи і дискети не повертаються.*

**Адрес редакції:** 69063, г. Запоріжжє, ул. Жуковського, 64, ЗНТУ, редакція журналу "PIU"  
Тел.:

(061) 7-698-2-96 – редакційно-видавничий відділ

(061) 7-644-6-62 – головний редактор

(061) 7-698-2-96 – зам. головного редактора

Факс: (0612) 64-21-41

E-mail: [rvv@zntu.edu.ua](mailto:rvv@zntu.edu.ua), [TQM@zntu.edu.ua](mailto:TQM@zntu.edu.ua)

*Наукове видання*

**Радіоелектроніка,  
інформатика,  
управління**

**№ 2 / 2009**

*Науковий журнал*

Головний редактор – д-р техн. наук Піза Д. М.  
Заст. головного редактора – канд. техн. наук Дубровін В. І.

Комп'ютерне моделювання та верстання

Новікова К. О.

Оригінал-макет підготовлено у редакційно-видавничому відділі ЗНТУ

*Підписано до друку 07.07.2009. Формат 60x84/8.  
Папір офс. Різогр. друк. Ум. друк. арк. 19,1.  
Тираж 300 прим. Зам. № 1215.*

69063 м. Запоріжжя, ЗНТУ, друкарня, вул. Жуковського, 64