

Науковий журнал
Радіоелектроніка
Інформатика
Управління



2001
1 (5)

Видається двічі на рік

Зареєстрований **27 травня 1998 року**
Міністерством інформації України,
свідоцтво - серія **КВ № 3264**.

Засновник і видавник -
Запорізький державний технічний університет

Передплатний індекс - 22914

З М І С Т

| | |
|-----------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|----|
| РАДІОЕЛЕКТРОНІКА | 4 |
| ОПРЕДЕЛЕНИЕ ФУНКЦИЙ ЧУВСТВИТЕЛЬНОСТИ В ДИНАМИКЕ <i>А.И.Вершина, А.Г.Маркин</i> | 4 |
| ВЕСОВЫЕ ФУНКЦИИ ДЛЯ МАЛОТОЧЕЧНЫХ ДПФ В СИСТЕМАХ ПРИЕМА И ОБРАБОТКИ РАДИОЛОКАЦИОННЫХ СИГНАЛОВ <i>Ю.Л.Мейстер, Д.М.Пица</i> | 8 |
| УСТРОЙСТВО ДЛЯ ОПРЕДЕЛЕНИЯ РАССТОЯНИЯ МЕЖДУ ТРАНСПОРТНЫМ СРЕДСТВОМ И АКТИВНОЙ ПУТЕВОЙ СТРУКТУРОЙ В СИСТЕМЕ УПРАВЛЕНИЯ МАГНИТОЛЕВИТИРУЮЩИМИ ТРАНСПОРТНЫМИ СРЕДСТВАМИ <i>В.Н.Привалов, С.В.Плаксин, И.И.Соколовский, А.А.Яшин</i> | 12 |
| ЭКСПЕРИМЕНТАЛЬНЫЕ ИССЛЕДОВАНИЯ АДАПТИВНОЙ КОМПОЗИЦИОННОЙ СИСТЕМЫ ПРОСТРАНСТВЕННО-ВРЕМЕННОЙ ОБРАБОТКИ СИГНАЛОВ И ТРАДИЦИОННОЙ АДАПТИВНОЙ СИСТЕМЫ ПАРАЛЛЕЛЬНОГО ВЕСОВОГО СУММИРОВАНИЯ <i>В.П.Прокофьев, Ю.П.Чиняев</i> | 16 |
| ВЛИЯНИЕ РАЗБРОСА ТЕПЛОФИЗИЧЕСКИХ ПАРАМЕТРОВ ТРАНЗИСТОРОВ НА ХАРАКТЕРИСТИКИ СИММЕТРИЧНЫХ ТРАНЗИСТОРНЫХ СХЕМ <i>В.А.Сергеев</i> | 18 |
| МОДЕЛЮВАННЯ ПРОФІЛІВ РОЗПОДІЛУ ДОМШКОВИХ ІОНІВ, ВАКАНСІЙ ТА МІЖВУЗЛОВИХ АТОМІВ У НАПІВПРОВІДНИКОВОМУ МАТЕРІАЛІ ПРИ ІОННОМУ ЛЕГУВАННІ <i>В.М.Теслюк, О.Р.Корбецький</i> | 22 |

| | | | |
|---------------------------------------------------------------------------------------------------------------------------------------------------------------------------|---------------------------------------------------------------------------------------------------------------------------------------------------------------------|----|-----------|
| ВРАХУВАННЯ ПАРАМЕТРІВ ДЖЕРЕЛА ТА НАВАНТАЖЕННЯ ПРИ РОЗРАХУНКАХ ХАРАКТЕРИСТИК ДВОПОЛЯРНИХ ІМПУЛЬСНИХ ПЕРЕТВОРЮВАЧІВ <i>В.Д.Флора</i> | ОПЕРАТОРИ ІНТЕРВАЛЬНОЇ АРИФМЕТИКИ ДЛЯ РОЗРАХУНКУ ДОПУСКІВ <i>Г.М.Шило</i> | 27 | 32 |
| ІНФОРМАТИКА | | | 35 |
| LINEAR ALGEBRA AND PETRI NETS <i>A.Bourjij, M.Boutayeb</i> | НЕДВІЙКОВЕ КОДУВАННЯ ПОВІДОМЛЕНЬ В КАНАЛАХ З ПОДИНОКИМИ ПОМИЛКАМИ <i>Ю.П.Жураковський</i> | 35 | 82 |
| BUILDING EUROPEAN COMPETENCE CENTRES TO SUPPORT KNOWLEDGE DISTRIBUTION AND LEARNING PROCESSES <i>I.Hamburg, S.Balanica</i> | ВИКОРИСТАННЯ МАТЕМАТИЧНИХ МЕТОДІВ ДЛЯ ОПТИМІЗАЦІЇ ТЕХНОЛОГІЧНИХ ПРОЦЕСІВ ПЕРЕРОБКИ СОЇ <i>В.М.Калініченко</i> | 44 | 86 |
| ОБЗОР СОВРЕМЕННЫХ ТИПОВ НЕЙРОННЫХ СЕТЕЙ <i>М.В.Аникеев, Л.К.Бабенко, О.Б.Макаревич</i> | АЛГОРИТМ МОДЕЛИРОВАНИЯ ФУНКЦИЙ ГРИНА МНОГОСЛОЙНЫХ ДИЭЛЕКТРИЧЕСКИХ СТРУКТУР <i>Л.М.Карпуков</i> | 48 | 87 |
| ЛОГИКО-АЛГЕБРАИЧЕСКИЕ ОСНОВЫ МАТЕМАТИЧЕСКОЙ ТЕОРИИ НАДЕЖНОСТИ <i>Л.И.Волгин</i> | ІНШИЙ ВИВІД ОСНОВНИХ СПІВВІДНОШЕНЬ НЕПЕРЕРВНОГО ВАРІАНТУ ДИНАМІЧНОГО ПРОГРАМУВАННЯ <i>А.Г.Кіку, Т.І.Білоус</i> | 56 | 90 |
| ПОСТРОЕНИЕ АЛГОРИТМА ПОИСКА НЕИСПРАВНОСТЕЙ МНОГОКАНАЛЬНЫХ СИСТЕМ <i>Д.Д.Габриэльян, В.В.Шацкий, Н.В.Шацкий</i> | КОНТРОЛЬ ТЕХНОЛОГИЧЕСКОГО ПРОЦЕССА С ИСПОЛЬЗОВАНИЕМ КАРТЫ ХОТЕЛЛИНГА <i>В.Н.Клячкин</i> | 64 | 92 |
| МОДЕЛИРОВАНИЕ ФУНКЦИИ ЛАНЖЕВЕНА ЦЕПНОЙ ДРОБЬЮ <i>В.И.Грядун, В.П.Пинчук, В.П.Шапвалов</i> | АНАЛИЗ РАБОТЫ СИСТЕМЫ КРОВООБРАЩЕНИЯ, КАК ПРОЦЕССА ПЕРЕДАЧИ ИНФОРМАЦИИ КЛЕТКАМИ КРОВИ, МЕТОДАМИ ВНЕШНЕЙ АЛГЕБРЫ <i>Г.В.Кузнецов, А.А.Яшин</i> | 68 | 94 |
| ПРЕДВАРИТЕЛЬНАЯ НАСТРОЙКА ЯДЕРНОЙ НЕЙРОННОЙ СЕТИ <i>А.Ю.Дорогов</i> | ИДЕНТИФИКАЦИЯ СЛОЖНЫХ ДИНАМИЧЕСКИХ СИСТЕМ ПРИ ПОМОЩИ Д-СЕТИ <i>Е.Н.Литвинов, А.С.Лихозов, Т.А.Лихозова</i> | 71 | 100 |
| ВОЗМОЖНОСТИ ПРИМЕНЕНИЯ НЕЙРОННЫХ СЕТЕЙ В ЗАДАЧАХ ПРОГНОЗИРОВАНИЯ И СОБЛЮДЕНИЕ ПРИНЦИПА СОГЛАСОВАННОСТИ ПОЛУЧЕННЫХ РЕЗУЛЬТАТОВ <i>А.С.Дулесов</i> | ОСОБЕННОСТИ ПРОЦЕДУРНО- ПАРАМЕТРИЧЕСКОЙ ПАРАДИГМЫ ПРОГРАММИРОВАНИЯ <i>А.И.Легалов</i> | 76 | 102 |
| ДИАГНОСТИКА И КОНТРОЛЬ СОСТОЯНИЯ КОЛЕСА ПЕРВОЙ СТУПЕНИ РКНД АНСАМБЛЕМ НЕЙРОСЕТЕЙ <i>С.В.Жернаков</i> | | 79 | |

| | | | |
|---------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|-----------------------------------------------------------------------------------------------------------------------------------------------------------------|-----|-----|
| АНАЛИЗ ФОРМАЛЬНО-ЛОГИЧЕСКОЙ МОДЕЛИ НЕЙРОНА С ПОМОЩЬЮ ДИСКРЕТНОГО ПРОСТРАНСТВА СОСТОЯНИЙ <i>В.Н.Лотин</i> | ИМИТАЦИОННЫЙ ПРОЕКТ АВТОМОБИЛЬНОГО ДОРОЖНОГО ДВИЖЕНИЯ <i>В.Н.Томашевский, Д.С.Печенежский</i> | 106 | 117 |
| НОВЫЕ ИНФОРМАЦИОННО-ВЫЧИСЛИТЕЛЬНЫЕ ТЕХНОЛОГИИ ДЛЯ РАСЧЕТА И АНАЛИЗА РЕЖИМОВ ТЕПЛОСНАБЖАЮЩИХ СИСТЕМ <i>Н.Н.Новицкий, В.В.Токарев, Э.И.Шалагинова</i> | КУБИЧЕСКОЕ МОДЕЛИРОВАНИЕ НЕИСПРАВНОСТЕЙ ЦИФРОВЫХ ПРОЕКТОВ НА ОСНОВЕ FPGA, CPLD <i>В.И.Хаханов, И.Ю.Сысенко, Хак Х.М.Джахирул, Масуд М.Д.Мехеда</i> | 108 | 123 |
| ИМИТАЦИОННОЕ МОДЕЛИРОВАНИЕ НЕЙРОСЕТЕЙ ДЛЯ РЕШЕНИЯ УРАВНЕНИЙ МАТЕМАТИЧЕСКОЙ ФИЗИКИ ЛОКАЛЬНО-АСИНХРОННЫМИ МЕТОДАМИ <i>М.А.Новотарский</i> | ТОРСИОННЫЕ ПОЛЯ И НЕЙРОКОМПЬЮТЕР <i>В.Д.Цыганков</i> | 113 | 130 |
| УПРАВЛЕНИЯ | | | 135 |
| О ВЫБОРЕ ВЕСОВЫХ КОЭФФИЦИЕНТОВ ОПТИМИЗИРУЕМОГО ФУНКЦИОНАЛА В ТЕОРИИ АНАЛИТИЧЕСКОГО КОНСТРУИРОВАНИЯ ОПТИМАЛЬНЫХ РЕГУЛЯТОРОВ <i>И. Е. Александрова, Т. Е. Александрова</i> | ОПТИМИЗАЦИЯ ПАРАМЕТРОВ ЦИФРОВЫХ НЕЧЕТКИХ РЕГУЛЯТОРОВ <i>В. И. Гостев, И. П. Лесовой, А. Е. Чуприн</i> | 135 | 148 |
| СОСТОЯНИЕ, ТЕНДЕНЦИИ И ПРОБЛЕМЫ В ОБЛАСТИ МЕТОДОВ УПРАВЛЕНИЯ АСИНХРОННЫМИ ДВИГАТЕЛЯМИ <i>В. Г. Бичай, Д. М. Пица, Е. Е. Потепенко, Е. М. Потепенко</i> | ИНСТРУМЕНТАЛЬНЫЕ СРЕДСТВА ВСЕОБЩЕГО УПРАВЛЕНИЯ КАЧЕСТВОМ: КОНТРОЛЬНЫЕ КАРТЫ <i>Ю. Н. Дорошенко, В. И. Дубровин</i> | 138 | 151 |
| ФОРМАЛИЗОВАННАЯ ИНТЕГРАЦИЯ ИНФОРМАЦИОННОЙ И ИСПОЛНИТЕЛЬНОЙ ПОДСИСТЕМ МНОГОСВЯЗНЫХ ОБЪЕКТОВ <i>В. Н. Буков, В. В. Косьяничук, В. Н. Рябченко</i> | ОНЛАЙНОВЫЕ МЕТОДЫ УПРАВЛЕНИЯ КАЧЕСТВОМ: ГИБРИДНАЯ ДИАГНОСТИКА НА ОСНОВЕ НЕЙРОННЫХ СЕТЕЙ <i>В. И. Дубровин, С. А. Субботин</i> | 144 | 158 |
| АВТОРЫ НОМЕРА | | | 164 |

І. РАДІОЕЛЕКТРОНІКА

УДК 621.38

ОПРЕДЕЛЕНИЕ ФУНКЦИЙ ЧУВСТВИТЕЛЬНОСТИ В ДИНАМИКЕ

А.И.Вершина, А.Г.Маркин

Рассматриваются пути повышения эффективности определения функций чувствительности при моделировании радиоэлектронных схем в динамике при значительном количестве варьируемых параметров. Получены формулы, позволяющие решать эту задачу.

Розглядаються шляхи підвищення ефективності визначення функцій чутливості при моделюванні радіоелектронних схем у динаміці при значній кількості варіативних параметрів. Отримано формули, що дозволяють вирішувати цю задачу.

Presented are the ways for increasing the efficiency of determining functions of sensitivity while semi-lating radio-electronic networks in dynamics for many variables parameters. Obtained formulas the solution of this problem.

ВВЕДЕНИЕ

Определение функций чувствительности в динамике, под которыми понимаются значения изменяющихся во времени частных производных выходных переменных по параметрам элементов, представляет собой довольно трудоемкую задачу, связанную с многократным интегрированием систем нелинейных дифференциальных уравнений высокого порядка. Необходимость нахождения функций чувствительности связана с решением задач оптимизации, надежности, допусков и т.д.. Это определяет актуальность нахождения методов повышения эффективности определения функций чувствительности для динамических процессов.

ПОСТАНОВКА ЗАДАЧИ

Радиоэлектронная схема описывается системой нелинейных алгебро-дифференциальных уравнений

$$F(\dot{x}, x, p) = 0, \quad (1)$$

где x и \dot{x} - вектор-столбец неизвестных переменных и их производных по времени; p - параметры элементов радиоэлектронной схемы.

Производная по параметрам запишется следующим образом:

$$\frac{\partial F}{\partial \dot{x}} \frac{\partial}{\partial p} \dot{x} + \frac{\partial F}{\partial x} \frac{\partial x}{\partial p} + \frac{\partial F}{\partial p} = 0. \quad (2)$$

Представим производную $\frac{\partial}{\partial p} \dot{x}$ в виде

$$\frac{\partial}{\partial p} \dot{x} = \frac{\partial}{\partial p} \left(\frac{\partial x}{\partial t} \right) = \frac{\partial}{\partial t} \left(\frac{\partial x}{\partial p} \right), \quad (3)$$

отсюда, переходя к конечным приращениям, можно записать

$$\frac{\partial}{\partial t} \left(\frac{\partial x}{\partial p} \right) \approx \frac{\frac{\partial x_{i+1}}{\partial p} - \frac{\partial x_i}{\partial p}}{\Delta t} = \frac{1}{h} \left(\frac{\partial x_{i+1}}{\partial p} - \frac{\partial x_i}{\partial p} \right), \quad (4)$$

где $h = \Delta t$.

После подстановки в выражение (2) получим

$$\frac{1}{h_i} \frac{\partial F_i}{\partial \dot{x}} \left(\frac{\partial x_{i+1}}{\partial p} - \frac{\partial x_i}{\partial p} \right) + \frac{\partial F_i}{\partial x} \frac{\partial x_{i+1}}{\partial p} + \frac{\partial F_i}{\partial p} = 0 \quad (5)$$

или

$$\frac{1}{h_i} \frac{\partial F_i}{\partial \dot{x}} \frac{\partial x_{i+1}}{\partial p} - \frac{1}{h_i} \frac{\partial F_i}{\partial \dot{x}} \frac{\partial x_i}{\partial p} + \frac{\partial F_i}{\partial x} \frac{\partial x_{i+1}}{\partial p} + \frac{\partial F_i}{\partial p} = 0. \quad (6)$$

Отсюда

$$\frac{1}{h_i} \frac{\partial F_i}{\partial \dot{x}} \frac{\partial x_{i+1}}{\partial p} + \frac{\partial F_i}{\partial x} \frac{\partial x_{i+1}}{\partial p} = \frac{1}{h_i} \frac{\partial F_i}{\partial \dot{x}} \frac{\partial x_i}{\partial p} - \frac{\partial F_i}{\partial p}; \quad (7)$$

$$\frac{\partial x_{i+1}}{\partial p} = \left(\frac{1}{h_i} \frac{\partial F_i}{\partial \dot{x}} + \frac{\partial F_i}{\partial x} \right)^{-1} \left(\frac{1}{h_i} \frac{\partial F_i}{\partial \dot{x}} \frac{\partial x_i}{\partial p} - \frac{\partial F_i}{\partial p} \right). \quad (8)$$

Данное выражение представляет собой неявный одношаговый метод интегрирования для определения функций чувствительности.

Ограничим набор выходных переменных элементами вектора x и рассмотрим эффективность алгоритмов вычислений в зависимости от количества выходных переменных и варьируемых параметров.

Для одного варьируемого параметра организация процесса вычисления функций чувствительности по всем

выходным переменным не представляет особых трудностей и осуществляется одновременно с процессом интегрирования в динамике. При этом используется разложение

матрицы $\frac{1}{h_i} \frac{\partial F}{\partial \dot{x}} + \frac{\partial F}{\partial x}$ на треугольные, получаемое на

последней итерации Ньютона, и трудоемкость расчета по одному варьируемому параметру определяется прямым и обратным ходом решения систем линейных алгебраических уравнений. Такому подходу соответствует метод моделей чувствительности [1-3].

При росте числа варьируемых параметров объем вычислений растет пропорционально их количеству.

Представим выражение (8) в развернутом виде

$$\begin{aligned} \frac{\partial x_{i+1}}{\partial p} &= \left(\frac{1}{h_i} \frac{\partial F_i}{\partial \dot{x}} + \frac{\partial F_i}{\partial x} \right)^{-1} \frac{1}{h_i} \frac{\partial F_i}{\partial \dot{x}} \frac{\partial x_i}{\partial p} - \left(\frac{1}{h_i} \frac{\partial F_i}{\partial \dot{x}} + \frac{\partial F_i}{\partial x} \right)^{-1} \frac{\partial F_i}{\partial p} = \\ &= - \left(\frac{1}{h_i} \frac{\partial F_i}{\partial \dot{x}} + \frac{\partial F_i}{\partial x} \right)^{-1} \frac{\partial F_i}{\partial p} + \\ &+ \left(\frac{1}{h_i} \frac{\partial F_i}{\partial \dot{x}} + \frac{\partial F_i}{\partial x} \right)^{-1} \frac{1}{h_i} \frac{\partial F_i}{\partial \dot{x}} \left(\frac{1}{h_{i-1}} \frac{\partial F_{i-1}}{\partial \dot{x}} + \frac{\partial F_{i-1}}{\partial x} \right)^{-1} \times \\ &\times \left(\frac{1}{h_{i-1}} \frac{\partial F_{i-1}}{\partial \dot{x}} \frac{\partial x_{i-1}}{\partial p} - \frac{\partial F_{i-1}}{\partial p} \right) = \\ &= - \left(\frac{1}{h_i} \frac{\partial F_i}{\partial \dot{x}} + \frac{\partial F_i}{\partial x} \right)^{-1} \frac{\partial F_i}{\partial p} - \\ &- \left(\frac{1}{h_i} \frac{\partial F_i}{\partial \dot{x}} + \frac{\partial F_i}{\partial x} \right)^{-1} \frac{1}{h_i} \frac{\partial F_i}{\partial \dot{x}} \left(\frac{1}{h_{i-1}} \frac{\partial F_{i-1}}{\partial \dot{x}} + \frac{\partial F_{i-1}}{\partial x} \right)^{-1} \frac{\partial F_{i-1}}{\partial p} - \\ &- \dots - \left(\frac{1}{h_i} \frac{\partial F_i}{\partial \dot{x}} + \frac{\partial F_i}{\partial x} \right)^{-1} \frac{1}{h_i} \frac{\partial F_i}{\partial \dot{x}} \times \dots \times \\ &\times \frac{1}{h_1} \frac{\partial F_1}{\partial \dot{x}} \left(\frac{\partial F_0}{\partial x} \right)^{-1} \frac{\partial F_0}{\partial p}. \end{aligned} \quad (9)$$

Это выражение можно записать следующим образом:

$$\begin{aligned} \frac{\partial x_{i+1}}{\partial p} &= \\ &= - \left(\frac{1}{h_i} \frac{\partial F_i}{\partial \dot{x}} + \frac{\partial F_i}{\partial x} \right)^{-1} \left(\frac{\partial F_i}{\partial p} - \right. \\ &- \frac{1}{h_i} \frac{\partial F_i}{\partial \dot{x}} \left(\frac{1}{h_{i-1}} \frac{\partial F_{i-1}}{\partial \dot{x}} + \frac{\partial F_{i-1}}{\partial x} \right)^{-1} \left(\frac{\partial F_{i-1}}{\partial p} - \right. \\ &- \frac{1}{h_{i-1}} \frac{\partial F_{i-1}}{\partial \dot{x}} \left(\frac{\partial F_{i-1}}{\partial p} - \dots - \frac{1}{h_1} \frac{\partial F_1}{\partial \dot{x}} \left(\left(\frac{\partial F_0}{\partial x} \right)^{-1} \frac{\partial F_0}{\partial p} \right) \dots \right) \left. \right) \left. \right). \end{aligned} \quad (10)$$

Отсюда видно, что для определение функции чувствительности одной выходной переменной по всем варьируемым параметрам процесс вычислений может быть организован следующим образом:

- находится строка обратной матрицы $\left(\frac{1}{h_i} \frac{\partial F_i}{\partial \dot{x}} + \frac{\partial F_i}{\partial x} \right)^{-1}$,

которая соответствует выходной переменной;

- производные $\frac{\partial F_i}{\partial p}$ умножаются на соответствующие элементы этой строки и запоминаются в векторе формирования функций чувствительности;

- строка обратной матрицы умножается на матрицу

$$\frac{1}{h_i} \frac{\partial F_i}{\partial \dot{x}};$$

- полученная после умножения строка умножается на

$$\left(\frac{1}{h_{i-1}} \frac{\partial F_{i-1}}{\partial \dot{x}} + \frac{\partial F_{i-1}}{\partial x} \right)^{-1} \text{ и т.д.}$$

Такой подход соответствует определению функций чувствительности методом присоединенных схем, которому, как это было отмечено в [3], присущи значительные вычислительные затраты, так как для определения функций чувствительности в момент $t = t_i$, необходимо производить расчеты в обратном масштабе времени для каждой интересующей нас временной точки. Несмотря на это, для практики более предпочтительно эффективно определять функцию чувствительности по всем варьируемым параметрам сразу.

На основании изложенного, поиск путей повышения эффективности методов определения функций чувствительности по многим варьируемым параметрам является актуальной задачей при исследовании динамических процессов в радиоэлектронных схемах.

МЕТОДЫ И АЛГОРИТМЫ РЕШЕНИЯ

1. Рассмотрим случай, когда исследуемая схема содержит только реактивные элементы. Ему соответствует

нулевое значение $\frac{\partial F}{\partial x}$ и выражение (2) может быть записано следующим образом:

$$\frac{\partial F}{\partial \dot{x}} \frac{\partial}{\partial p} \dot{x} + \frac{\partial F}{\partial p} = 0, \quad (11)$$

отсюда с учетом (4) и (8) получим

$$\frac{\partial x_{i+1}}{\partial p} = \frac{\partial x_i}{\partial p} - \left(\frac{1}{h_i} \frac{\partial F_i}{\partial \dot{x}} \right)^{-1} \frac{\partial F_i}{\partial p}, \quad (12)$$

так как

$$\left(\frac{1}{h_i} \frac{\partial F_i}{\partial \dot{x}} \right)^{-1} \frac{1}{h_i} \frac{\partial F_i}{\partial \dot{x}} = E, \quad (13)$$

где E - единичная матрица.

Выражение (12) показывает, что для схемы, представленной только реактивными элементами, определение функции чувствительности по всем варьируемым параметрам не представляет сложности, так как функции чувствительности можно определять в процессе получения значений выходных переменных в очередной временной точке. При этом, в соответствии с интересующими нас выходными переменными, определяются нужные

строки обратной матрицы $\left(\frac{1}{h_i} \frac{\partial F_i}{\partial \dot{x}}\right)^{-1}$.

2. Для радиоэлектронной схемы, в которой нет реактивных элементов, формула для определения функций чувствительности совпадает с выражением, используемым для статических режимов:

$$\frac{\partial x}{\partial p} = \left(\frac{\partial F}{\partial x}\right)^{-1} \frac{\partial F}{\partial p}. \quad (14)$$

В динамике, однако, присутствует время как независимый параметр.

Целесообразность применения того или иного метода определения функции чувствительности для этого случая подробно рассмотрены в работе [4].

3. Трудоемкость вычислений функций чувствительности по всем варьируемым параметрам с использованием выражения (10) представляет значительные трудности.

Рассмотрим возможность упрощения вычислений.

Из выражения (8) можно записать

$$\frac{\partial x_{i+1}}{\partial p} = \left(\frac{1}{h_i} \frac{\partial F_i}{\partial \dot{x}} + \frac{\partial F_i}{\partial x}\right)^{-1} \frac{1}{h_i} \frac{\partial F_i}{\partial \dot{x}} \frac{\partial x_i}{\partial p} - \left(\frac{1}{h_i} \frac{\partial F_i}{\partial \dot{x}} + \frac{\partial F_i}{\partial x}\right)^{-1} \frac{\partial F_i}{\partial p}. \quad (15)$$

Первое слагаемое может рассматриваться как составляющая функции чувствительности, зависящая от функций чувствительности предыдущих шагов, а второе - составляющая функции чувствительности, определяемая текущим шагом.

В отличие от схемы, которая содержит только реактивные элементы и первая составляющая у которой представлена функцией чувствительности предыдущего шага одной независимой переменной, для общего случая первая составляющая зависит от функций чувствительности всех независимых переменных. Это и создает трудности вычислений по всем варьируемым параметрам, так как отсутствует возможность ограничиваться определением только части выходных переменных.

Преобразуем (15) следующим образом:

$$\begin{aligned} \frac{\partial x_{i+1}}{\partial p} &= \left(\frac{1}{h_i} \frac{\partial F_i}{\partial \dot{x}} + \frac{\partial F_i}{\partial x}\right)^{-1} \frac{1}{h_i} \frac{\partial F_i}{\partial \dot{x}} \frac{\partial x_i}{\partial p} - \left(\frac{1}{h_i} \frac{\partial F_i}{\partial \dot{x}} + \frac{\partial F_i}{\partial x}\right)^{-1} \frac{\partial F_i}{\partial p} = \\ &= \left(\frac{1}{h_i} \frac{\partial F_i}{\partial \dot{x}} + \frac{\partial F_i}{\partial x}\right)^{-1} \left[\left(\frac{1}{h_i} \frac{\partial F_i}{\partial \dot{x}} + \frac{\partial F_i}{\partial x}\right) \frac{\partial x_i}{\partial p} - \frac{\partial F_i}{\partial p} \right] \end{aligned}$$

$$\begin{aligned} - \left(\frac{1}{h_i} \frac{\partial F_i}{\partial \dot{x}} + \frac{\partial F_i}{\partial x}\right)^{-1} \frac{\partial F_i}{\partial p} &= \\ = \frac{\partial x_i}{\partial p} - \left(\frac{1}{h_i} \frac{\partial F_i}{\partial \dot{x}} + \frac{\partial F_i}{\partial x}\right)^{-1} \frac{\partial F_i}{\partial x} \frac{\partial x_i}{\partial p} - \left(\frac{1}{h_i} \frac{\partial F_i}{\partial \dot{x}} + \frac{\partial F_i}{\partial x}\right)^{-1} \frac{\partial F_i}{\partial p}. \quad (16) \end{aligned}$$

Если выполняется условие

$$\left(\frac{1}{h_i} \frac{\partial F_i}{\partial \dot{x}} + \frac{\partial F_i}{\partial x}\right)^{-1} \frac{\partial F_i}{\partial x} \frac{\partial x_i}{\partial p} \ll \frac{\partial x_i}{\partial p}, \quad (17)$$

то при расчетах можно использовать выражение

$$\frac{\partial x_{i+1}}{\partial p} \approx \frac{\partial x_i}{\partial p} - \left(\frac{1}{h_i} \frac{\partial F_i}{\partial \dot{x}} + \frac{\partial F_i}{\partial x}\right)^{-1} \frac{\partial F_i}{\partial p}. \quad (18)$$

4. Во многих случаях вполне приемлемой является оценка значений функций чувствительности с точностью до 20-50%, а иногда достаточно знать только порядок этой величины. Это делает обоснованным поиск приближенных выражений для вычисления функций чувствительности по многим варьируемым параметрам.

Естественно предположить, что наибольший вклад в значение первой составляющей функции чувствительности в момент времени $t = t_{i+1}$ вносит составляющая функции чувствительности предыдущего шага в момент времени $t = t_i$.

Пренебрежем влиянием $\frac{\partial x_{i+1}}{\partial p}$, тогда выражение (9)

для функции чувствительности будет иметь вид

$$\begin{aligned} \frac{\partial x_{i+1}}{\partial p} &\approx - \left(\frac{1}{h_i} \frac{\partial F_i}{\partial \dot{x}} + \frac{\partial F_i}{\partial x}\right)^{-1} \frac{\partial F_i}{\partial p} - \\ &- \left(\frac{1}{h_i} \frac{\partial F_i}{\partial \dot{x}} + \frac{\partial F_i}{\partial x}\right)^{-1} \frac{1}{h_i} \frac{\partial F_i}{\partial \dot{x}} \left(\frac{1}{h_{i-1}} \frac{\partial F_{i-1}}{\partial \dot{x}} + \frac{\partial F_{i-1}}{\partial x}\right)^{-1} \frac{\partial F_{i-1}}{\partial p}. \quad (19) \end{aligned}$$

Хотя оценка погрешности использования данного выражения и условия, при которых результат не будет превышать заданной погрешности, в данном случае не рассматривается, такой подход может дать определенный эффект. Однако для использования этой формулы необходимо сохранить результаты предыдущего шага, вклю-

чающие в себя производные $\frac{\partial F_{i-1}}{\partial p}$ и разложение мат-

рицы схемы предыдущего шага $\frac{1}{h_{i-1}} \frac{\partial F_{i-1}}{\partial \dot{x}} + \frac{\partial F_{i-1}}{\partial x}$ на

треугольные, что не совсем удобно.

5. По-видимому, для некоторых радиоэлектронных схем влияние предыдущих шагов может быть незначительным и тогда выражение (18) можно ограничить

только вторым слагаемым:

$$\frac{\partial x_{i+1}}{\partial p} \approx - \left(\frac{1}{h_i} \frac{\partial F_i}{\partial \dot{x}} + \frac{\partial F_i}{\partial x} \right)^{-1} \frac{\partial F_i}{\partial p}. \quad (20)$$

6. Рассмотрим произведения $\left(\frac{1}{h_i} \frac{\partial F_i}{\partial \dot{x}} + \frac{\partial F_i}{\partial x} \right)^{-1} \frac{1}{h_i} \frac{\partial F_i}{\partial \dot{x}} \frac{\partial x_i}{\partial p}$ и

$$\left(\frac{1}{h_i} \frac{\partial F_i}{\partial \dot{x}} + \frac{\partial F_i}{\partial x} \right)^{-1} \frac{\partial F_i}{\partial p},$$

представляющие собой составляющие функции чувствительности, зависящие от предыдущих шагов и определяемые только текущим шагом. При определении значений функции чувствительности одной выходной переменной по всем параметрам сначала найдем соответствующую строку обратной матрицы

$$\left(\frac{1}{h_i} \frac{\partial F_i}{\partial \dot{x}} + \frac{\partial F_i}{\partial x} \right)^{-1}.$$

В первом случае нужно умножить ее на матрицу $\frac{1}{h_i} \frac{\partial F_i}{\partial \dot{x}}$, а затем на матрицу $\frac{\partial x_i}{\partial p}$, а во втором - на

$$\frac{\partial F_i}{\partial p}.$$

Строка обратной матрицы в общем случае содержит почти все ненулевые элементы и в то же время

$$\text{матрица } \frac{1}{h_i} \frac{\partial F_i}{\partial \dot{x}} \text{ в зависимости от количества и схемы}$$

включения реактивных элементов может иметь большое число нулевых столбцов. В результате произведение будет представлять собой матрицу строку, у которой множество нулевых элементов. Произведение этой строки на

$$\text{матрицу } \frac{\partial x_i}{\partial p} \text{ позволяет сделать вывод о зависимости}$$

функции чувствительности интересующей нас выходной переменной от функций чувствительности только некоторых независимых переменных, которым соответствуют ненулевые значения. Если количество ненулевых элементов незначительно, то можно в дополнение к интересующей нас выходной переменной находить функции чувствительности только соответствующих независимых переменных и этим существенно повысить эффективность расчетов без упрощения алгоритма.

В случае, если искомая выходная переменная совпадает с ненулевой переменной, расчет также не требует дополнительных затрат и трудоемкость сравнима с применением метода моделей.

7. Упрощение алгоритма можно производить тогда, когда мы игнорируем переменными в связи с малыми их значениями. Оценка этих значение требует дополнительных исследований, однако как показывают расчеты на этот процесс можно влиять выбором шага интегрирования

Представляет интерес случай, когда получаемая строка не содержит ни одного нулевого элемента. Это бывает тогда, когда, например, каждый узел исследуемой схемы содержит реактивный элемент. В принципе, все схемы можно рассматривать с этих позиций. Тогда выбор шага позволяет сделать реактивную составляющую сколь угодно большой. В результате мы будем иметь дело с формулой (18).

Действительно, рассмотрим выражение

$$R = \left(\frac{1}{h_i} \frac{\partial F_i}{\partial \dot{x}} + \frac{\partial F_i}{\partial x} \right)^{-1} \frac{1}{h_i} \frac{\partial F_i}{\partial \dot{x}}, \quad (21)$$

для которого можно записать

$$\left(\frac{1}{h_i} \frac{\partial F_i}{\partial \dot{x}} + \frac{\partial F_i}{\partial x} \right) \cdot R = \frac{1}{h_i} \frac{\partial F_i}{\partial \dot{x}}. \quad (22)$$

Умножим правую и левую часть на h_i :

$$\left(\frac{\partial F_i}{\partial \dot{x}} + h_i \frac{\partial F_i}{\partial x} \right) \cdot R = \frac{\partial F_i}{\partial \dot{x}}, \quad (23)$$

откуда

$$R = \left(\frac{\partial F_i}{\partial \dot{x}} + h_i \frac{\partial F_i}{\partial x} \right)^{-1} \frac{\partial F_i}{\partial \dot{x}}. \quad (24)$$

Перейдем к пределу, при $h_i \rightarrow 0$:

$$\lim_{h_i \rightarrow 0} R = \left(\frac{\partial F_i}{\partial \dot{x}} \right)^{-1} \frac{\partial F_i}{\partial \dot{x}} = E. \quad (25)$$

После подстановки в (15) получим (18). Естественно, необходимо, чтобы существовала обратная матрица

$$\left(\frac{\partial F_i}{\partial \dot{x}} \right)^{-1}.$$

8. По-видимому, в общем случае, имеет смысл разбиения задачи определения функций чувствительности на подзадачи, одна из которых позволяет получать приемлемый результат с использованием выражения (18), вторая даст удовлетворительный результат по формуле (20), а для оставшихся функций чувствительности использовать метод моделей.

РЕЗУЛЬТАТЫ

Отмечена актуальность поиска методов определения функций чувствительности по многим варьируемым параметрам при исследовании динамических процессов в радиоэлектронных схемах. Получены выражения (12,14, 18,20), которые могут быть использованы при решении

данной задачи, избежав интегрирования в обратном масштабе времени, которое присуще методу присоединенных схем, и в то же время определять за один дополнительный просчет функцию чувствительности одной выходной переменной по всем варьируемым параметрам.

ВЫВОДЫ

Поиск путей определения функций чувствительности по многим параметрам в динамических режимах в общем случае представляет собой трудоемкую задачу. Однако учет особенностей исследуемых радиоэлектронных схем (схемы, содержащие только реактивные элементы, либо схемы, в которых реактивные элементы отсутствуют) может в некоторых случаях существенно снизить затраты и сделать сравнимыми с расчетами в статических режимах (12, 14, 18, 20).

Так как во многих случаях требования к погрешности определения значений функций чувствительности до-

вольно низки, то может быть использована приближенная формула (19), полученная из предположения, что основной вклад в значение функции чувствительности вносит ближайшая временная точка.

В общем случае, целесообразно поставить задачу разбиения функций чувствительности на подмножества, элементы которых целесообразно рассчитывать по одному из приведенных выражений и только некоторые из функций чувствительности определять методом моделей.

ПЕРЕЧЕНЬ ССЫЛОК

1. Петренко А.И. Основы автоматизации проектирования. - К.: Техніка, 1982. - 295 с.
2. Петренко А.И., Семенов О.И. Основы построения систем автоматизированного проектирования - К.: Вища школа. Головное изд-во, 1984. - 296 с.
3. Петренко А.И., Власов А.И., Тимченко А.П. Табличные методы моделирования электронных схем на ЭЦВМ - К.: Вища школа, 1977. - 192 с.
4. Вершина А.И., Кузьмина Л.В. Определение функций чувствительности в статике // "Радиоэлектроника, информатика, управління". - Запоріжжя: ЗДТУ. - 2000. - №1. - с.9-12.

УДК 621.391.26

ВЕСОВЫЕ ФУНКЦИИ ДЛЯ МАЛОТОЧЕЧНЫХ ДПФ В СИСТЕМАХ ПРИЕМА И ОБРАБОТКИ РАДИОЛОКАЦИОННЫХ СИГНАЛОВ

Ю.Л.Мейстер, Д.М.Пиза

Проанализированы характеристики фильтров радиолокационных сигналов, построенных на основе малоточечных ДПФ, при взвешивании временной выборки некоторыми известными "классическими" функциями окон. Выявлена неэффективность этих функций при малоточечных ДПФ. Показано, что специально сконструированные функции окон для малоточечных ДПФ обладают значительно лучшими характеристиками.

Проаналізовано характеристики фільтрів радіолокаційних сигналів, побудованих на основі малоточкових ДПФ, за зважування часової вибірки деякими відомими "класичними" функціями вікон. Виявлено неефективність цих функцій для малоточкових ДПФ. Показано, що спеціально сконструйовані функції вікон для малоточкових ДПФ мають значно кращі характеристики.

The radar signals filters characteristics built on the basis of the small-numbered DFTs have been analyzed when weighing a temporal sample by some known "classic" functions of windows. There has been revealed an ineffectiveness of these functions with the small-numbered point DFTs. It's shown that the small-numbered point DFTs possess considerably better characteristics.

ВВЕДЕНИЕ

Оптимальная процедура обработки радиолокационных сигналов включает в себя определение корреляционного интеграла [1], т.е. вычисление взаимокорреляцион-

ной функции между принятой смесью полезного сигнала и помех и "портретом" (образом) излученного сигнала. Причем в силу неизвестности параметров сигнала обработка проводится для каждого элемента дальности (эквивалентно скользящему стробированию приемника) в каждом из набора доплеровских фильтров.

Вычислитель корреляционного интеграла для одиночного радиоимпульса, при аналогово-цифровом преобразовании на радио- или промежуточной частоте, формирующий одновременно и квадратурные составляющие сигнала, рассмотрен в [2]. Для оптимальной обработки последовательно с ним должен быть включен когерентный накопитель пачечного сигнала, как показано на рисунке 1, где S и C - квадратурные составляющие сигнала. Штриховой линией показан фильтр сжатия, входящий в состав оптимального приемника в случае применения сложных сигналов с внутримпульсной модуляцией.

В связи с неизвестностью фазового сдвига, когерентный накопитель пачечного сигнала строится в виде набора фильтров, перекрывающих диапазон межпериодных фазовых сдвигов (доплеровских частот). Набор доплеровских фильтров может быть реализован на основе дискретного преобразования Фурье (ДПФ) как, например, в [3].

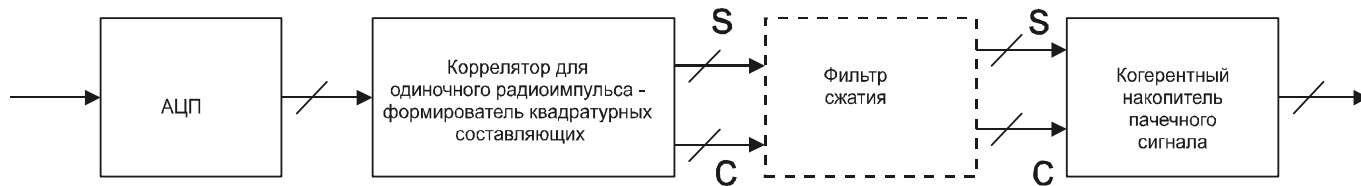


Рисунок 1 - Оптимальный приемник радиолокационных сигналов

Если $X(n)$ - последовательность из N эквидистантных временных выборок входного сигнала, то ДПФ создает последовательность из N эквидистантных частотных выборок $X(k)$.

$$X(k) = \sum_{n=0}^{N-1} X(n)e^{-j\frac{2\pi kn}{N}}, \quad (1)$$

где $e^{-j\frac{2\pi kn}{N}} = \cos\left(\frac{2\pi kn}{N}\right) - j\sin\left(\frac{2\pi kn}{N}\right)$; $k = 0, 1, \dots, N-1$.

Таким образом, ДПФ можно рассматривать как корреляцию входного сигнала с набором из N синусоид, расположенных с интервалом $1/NT$ на оси частот, где T -период выборки, т. е. период повторения зондирующих сигналов.

Размер выборки N в радиолокации ограничен временем облучения объекта антенной РЛС, и при возрастающих требованиях к скорости обзора, может составлять не более нескольких десятков. Кроме того, для борьбы с эффектом "слепых скоростей" и снижения неравномерности скоростной характеристики, применяют "пачечную" вобуляцию частоты повторения, при которой, зондирующий сигнал состоит из повторяющихся пачек из нескольких (обычно 2...4 частот повторения), причем внутри каждой пачки, частота повторения $1/T$ остается неизменной. В обзорных РЛС число импульсов в такой пачке часто находится в пределах 7...16 (лишь при слежении или сопровождении объектов, применяется большее число импульсов в пачке (32 и более)). Таким образом, малое число импульсов в пачке, характерной для РЛС в режиме обзора, обуславливает применение "малоточечных ДПФ" (соответственно величина выборки в выражении (1) $N=7...16$).

Пачечный характер сигналов (с конечным числом импульсов в пачке) эквивалентен умножению сигнала на прямоугольное окно. Вследствие такого умножения во временной области спектральная характеристика имеет характеристику вида $\sin(x)/x$ с главным лепестком и набором боковых лепестков с максимальным уровнем около -13дБ. Такой уровень боковых лепестков является недопустимо большим для построения доплеровских фильтров, применяемых в радиолокации, где спектральные составляющие помех могут значительно превышать

спектральные составляющие полезных сигналов.

Для снижения уровня боковых лепестков необходимо изменить функцию окна, в результате чего уровень боковых лепестков уменьшается, что ослабляет "просачивание" спектральных составляющих из одного фильтра в другой.

Применение весовых функций, кроме снижения уровня боковых лепестков, к сожалению, расширяет главный лепесток и, поэтому, в практике применяют различные весовые функции (окна), при которых достигается определенный компромисс между уровнем боковых лепестков и шириной главного лепестка. Однако известные результаты не позволяют найти приемлемый компромисс в радиолокационной практике при малом числе импульсов в пачке.

ПОСТАНОВКА ЗАДАЧИ

Задачей данной работы является исследование весовых функций при малом количестве точек ДПФ, а также синтезирование новых весовых функций, позволяющих оптимизировать характеристики доплеровской фильтрации радиолокационных сигналов. В частности, потребуем, чтобы синтезируемые функции обеспечивали уровень боковых лепестков ниже известных, описанных в литературе функций, а ширина главного лепестка должна быть, как минимум, не шире чем у известных функций.

РЕЗУЛЬТАТЫ ИССЛЕДОВАНИЙ

Исследования проведены для малоточечных ДПФ при использовании наиболее применяемых "классических" окон. В частности:

- окно Хэмминга [$W(n)_{X43}$] с заявленным уровнем боковых лепестков - 43дБ:

$$W(n)_{X43} = a_0 - (1 - a_0) \cos[2\pi n/N]; \quad (2)$$

- окно Блэкмана-Хэрриса [$W(n)_{BX67}$] с заявленным уровнем боковых лепестков - 67дБ:

$$W(n)_{BX67} = a_0 - a_1 \cos[2\pi n/N] + a_2 \cos[2\pi 2n/N]; \quad (3)$$

- окно Блэкмана-Хэрриса [$W(n)_{BX92}$] с заявленным

уровнем боковых лепестков -92дБ.

$$W(n)_{\text{БХ92}} = a_0 - a_1 \cos[2\pi n/N] + a_2 \cos[2\pi 2n/N] + a_3 \cos[2\pi 3n/N]. \quad (4)$$

Значения коэффициентов a_n приведены в таблице 1.

На рисунке 2 приведены характеристики фильтров ДПФ для указанных выше весовых функций восьми-точечного ДПФ.

В таблице 2 приведена ширина главного лепестка фильтров ДПФ по уровням 3дБ, 6дБ и по "нолям". Причем эти параметры оказались практически независимыми от числа точек ДПФ. В первой строке для сравнения приведены параметры прямоугольного окна $W(n)_{\text{П}}$, а в последней строке - параметры специально сконструированной весовой функции W_c . Сущность подхода, использованного при этом, заключалась в следующем. Если не требовать от оконных функций простоты их математического описания и вычисления коэффициентов, т.е. перейти к индивидуально сконструированным окнам, оптимальным для конкретных задач, то можно создать оконные функции W_c для малоточечных ДПФ, основным достоинством которых является малый уровень боковых лепестков. В процессе работы сконструирован ряд окон и проведено исследование их характеристик.

Уровни боковых лепестков для исследуемых функций для числа точек от 7 до 64 приведены в таблицах 3, 4, 5.

Таблица 1 - Значения коэффициентов a_n

| a_n | Значение коэффициентов a_n для окон | | |
|-------|---------------------------------------|----------------------|----------------------|
| | $W(n)_{\text{Х43}}$ | $W(n)_{\text{БХ67}}$ | $W(n)_{\text{БХ92}}$ |
| a_0 | 0.54348 | 0.42323 | 0.35875 |
| a_1 | - | 0.49755 | 0.48829 |
| a_2 | - | 0.07922 | 0.14128 |
| a_3 | - | - | 0.01168 |

Таблица 2 - Ширина главного лепестка фильтров ДПФ

| окно | Ширина главного лепестка по уровням в бинах (единицах $2\pi/N$) | | |
|----------------------|------------------------------------------------------------------|------|------------|
| | -3дБ | -6дБ | По "нолям" |
| $W(n)_{\text{П}}$ | 0.89 | 1.21 | 2.0 |
| $W(n)_{\text{Х43}}$ | 1.3 | 1.81 | 5.2 |
| $W(n)_{\text{БХ67}}$ | 1.66 | 2.38 | 6.0 |
| $W(n)_{\text{БХ92}}$ | 1.9 | 2.72 | 8.7 |
| $W(n)_c$ | 1.26 | 1.75 | 4.0 |

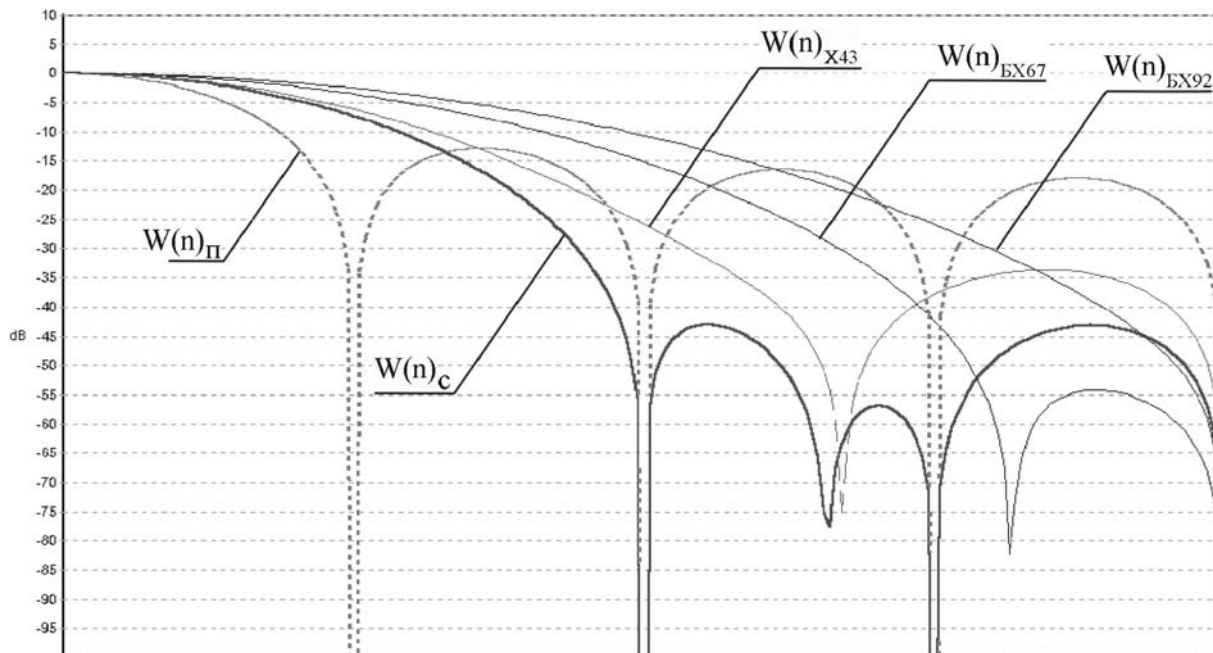


Рисунок 2 - Характеристики фильтров при 8-точечном ДПФ

Таблица 3 - Уровень боковых лепестков фильтров ДПФ для окна Хэмминга $W(n)_{X43}$

| Количество точек | Уровень боковых лепестков, дБ |
|------------------|-------------------------------|
| 7 | -32.3 |
| 8 | -33.6 |
| 9 | -34.9 |
| 10 | -35.9 |
| 11 | -36.7 |
| 12 | -37.4 |
| 13 | -38.0 |
| 14 | -38.5 |
| 15 | -38.9 |
| 16 | -39.3 |
| 32 | -41.6 |
| 64 | -43.0 |

Таблица 4 - Уровень боковых лепестков фильтров ДПФ для окна Блэкмана-Хэрриса $W(n)_{BX67}$

| Количество точек | Уровень боковых лепестков, дБ |
|------------------|-------------------------------|
| 7 | -54.0 |
| 8 | -55.5 |
| 9 | -56.5 |
| 10 | -57.3 |
| 11 | -58.0 |
| 12 | -58.5 |
| 13 | -59.0 |
| 14 | -59.5 |
| 15 | -60.0 |
| 16 | -61.0 |
| 32 | -65.0 |
| 64 | -67.0 |

Приведенные результаты исследований показывают, что рассмотренные выше "классические" окна не обеспечивают заявленные [4] уровни боковых лепестков при количестве точек ДПФ менее 30-50. При количестве точек 7...9 расхождение составляет 8...13дБ. При количестве точек 12...16 расхождение составляет 6...8дБ (в сторону ухудшения). Таким образом, оказалось, что "классические" окна при малом количестве точек (7...16) имеют значительно худшие характеристики по сравнению с многоточечными окнами (более 30...50).

На рисунке 2 приведены спектральные характеристики фильтров ДПФ при использовании сконструирован-

ных окон $W(n)_c$ для 8-точечных ДПФ, а в таблице 6 приведены уровни боковых лепестков фильтров ДПФ при использовании сконструированных окон W_c для 7...16 точечных ДПФ.

Таблица 5 - Уровень боковых лепестков фильтров ДПФ для окна Блэкмана-Хэрриса $W(n)_{BX92}$

| Количество точек | Уровень боковых лепестков, дБ |
|------------------|-------------------------------|
| 12 | -86.0 |
| 13 | -86.5 |
| 14 | -87.0 |
| 15 | -88.0 |
| 16 | -88.5 |
| 32 | -91.5 |
| 64 | -92.0 |

Из рисунка 2 и таблицы 6 видно, что уровень боковых лепестков сконструированных окон практически не зависит от количества точек ДПФ и значительно меньше уровня боковых лепестков окна Хэмминга, даже при несколько более узком главном лепестке (см. таблицу 2). Кроме того, сконструированные окна оптимизированы по расположению нулей. Как видно из рисунка 2, нули сконструированных функций расположены в точках $2\pi/N$, т.е. на месте максимумов всех N фильтров из набора, что при ширине главного лепестка по нулям $4\pi/N$ минимизирует спектральное просачивание, особенно существенное для соседних фильтров, а также минимизирует корреляцию перекрывающихся участков через один фильтр, что важно для некоторых алгоритмов обработки сигналов.

Таблица 6 - Уровни боковых лепестков сконструированных окон

| Количество точек | Уровень боковых лепестков, дБ |
|------------------|-------------------------------|
| 7 | -43.6 |
| 8 | -42.5 |
| 9 | -42.0 |
| 10 | -41.9 |
| 11 | -41.8 |
| 12 | -41.7 |
| 13 | -41.8 |
| 14 | -41.9 |
| 15 | -42.0 |
| 16 | -42.1 |

На рисунке 3, проиллюстрированы зависимости уровня боковых лепестков АЧХ доплеровских фильтров в зависимости от количества "точек" ДПФ, при взвешивании входной выборки оконными функциями Хэмминга $W(n)_{Х43}$, Блэкмана-Хэрриса $W(n)_{БХ67}$, а также специально сконструированной функцией $W(n)_c$.

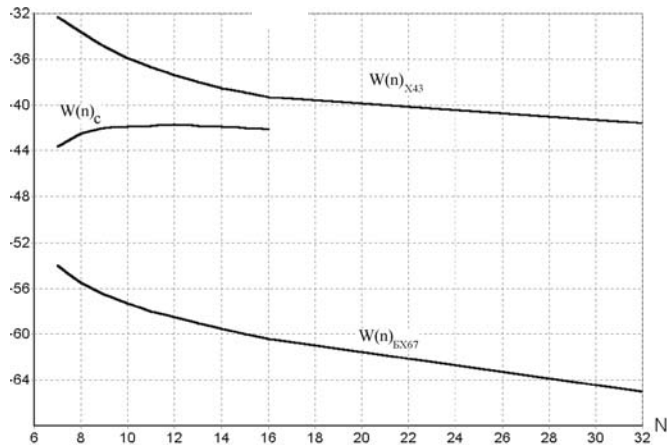


Рисунок 3 - Уровни боковых лепестков фильтров ДПФ в зависимости от точности ДПФ

УДК 621.373.5: 629.439

УСТРОЙСТВО ДЛЯ ОПРЕДЕЛЕНИЯ РАССТОЯНИЯ МЕЖДУ ТРАНСПОРТНЫМ СРЕДСТВОМ И АКТИВНОЙ ПУТЕВОЙ СТРУКТУРОЙ В СИСТЕМЕ УПРАВЛЕНИЯ МАГНИТОЛЕВИТИРУЮЩИМИ ТРАНСПОРТНЫМИ СРЕДСТВАМИ

В.Н.Привалов, С.В.Плаксин, И.И.Соколовский, А.А.Яшин

Исследован автогенераторный датчик измерения расстояний и их девиаций между подвижными объектами, построенный на основе двух СВЧ-генераторов, один из которых является опорным. Частота излучения второго генератора, являющегося одновременно смесителем, связана с изменением фазового расстояния между объектами. Разностная частота является информационной. Определены критерии достижения максимальной чувствительности датчика.

Досліджено автогенераторний датчик виміру відстаней і їх девіацій між рухомими об'єктами, побудований на основі двох НВЧ-генераторів, один із яких є опорним. Частота випромінювання другого генератора, що є одночасно змішу-

ВЫВОДЫ

Проведено исследование спектральных характеристик фильтров ДПФ при взвешивании входной временной выборки некоторыми известными "классическими" оконными функциями для малоточечных ДПФ. Результаты приведены в виде графиков и таблиц, из которых следует:

1. Известные "классические" окна имеют значительно более высокий уровень боковых лепестков, при малом количестве точек (7...16) по сравнению с заявленным уровнем, приведенным в публикациях.

2. Специально сконструированные для малоточечных ДПФ весовые функции могут обеспечить меньший уровень боковых лепестков, практически не зависящий от числа точек ДПФ, даже при несколько более узком главном лепестке. Эти функции уменьшают спектральное просачивание из соседних фильтров и минимизируют корреляцию перекрывающихся участков "через один" фильтр, что важно для некоторых алгоритмов обработки сигналов.

ПЕРЕЧЕНЬ ССЫЛОК

1. Радиотехнические системы. // Под ред. Ю.М. Казаринова, М., "Сов. радио", 1968, 496 с.
2. Д.М. Пиза, Ю.Л. Мейстер. Цифровые методы формирования квадратурных составляющих в системах приема и обработки радиолокационных сигналов. // "Радиоэлектроника, информатика, управління". - Запорожье, ЗГТУ, №1, 1999, с. 81-84.
3. Мейстер Ю.Л., Зайцев Е.И., Колпаков В.П., Кукольников А.Ф. / а.с. СССР №133322.
4. Ф.Дж.Хэррис. Использование окон при гармоническом анализе методом дискретного преобразования Фурье. / Пер. с англ., М., "Мир", ТИИЭР, т.66, №1, 1978, с. 60-96.

вачем, зв'язана зі зміною фазової відстані між об'єктами. Різнісна частота є інформаційною. Визначено критерії досягнення максимальної чутливості датчика.

Explored autogenerator-sensor of distance measurement and their deviations between moving objects, built on the base two SHF-generators (cm-wave band), one of which is stable (supporting). Frequency of radiating a second generator, being simultaneously blender, connected with changing a phase distance between objects. The difference frequency is information. Determined criterion of achievement of maximum sensitivity of sensor.

Определение пространственного положения магнитолевитирующего транспортного средства (МТС) - одна из важнейших задач при создании системы управления МТС и включает определение расстояния между МТС и путевой структурой, а также между МТС и боковыми направляющими. В силу специфики условий, в которых работают устройства для определения расстояний между указанными объектами, разработка указанных устройств должна базироваться на бесконтактных методах измерения. Устройства должны быть безинерционными и отслеживать малые изменения расстояний между объектами и при этом быть помехоустойчивыми.

Как нами уже подчеркивалось [1], указанным требованиям удовлетворяют радиоволновые методы измерений, использующие электромагнитные волны диапазона сверхвысоких частот (СВЧ). Основанный на этих принципах датчик малых перемещений, разработанный нами ранее [2], и базирующийся на использовании двухконтурного диодного СВЧ-автогенератора, соответствовал указанным выше условиям, но не был достаточно надежным из-за большесигнальности низкой частоты автоколебаний, являющейся информативной, и перегружающей генераторный полупроводниковый диод. Поэтому задачей настоящего исследования являлось обеспечение более высокой надежности функционирования радиоволнового устройства для определения расстояний с выдачей информации об изменениях указанных расстояний по радиоканалу со свободным выбором частоты информационного сигнала с тем, чтобы указанные значения частоты лежали в полосе минимальных значений помеховых сигналов.

Функционирование разработанного устройства основано на явлении затягивания частоты полупроводникового СВЧ-автогенератора рассогласованной СВЧ-нагрузкой, в качестве которой выступает путевая структура, а само устройство устанавливается на корпусе МТС. При этом уход частоты генерации, вызванный изменениями расстояния между ТС и частично отражающей путевой структурой, являются информативными. Информационный сигнал выделяется на контуре низкой частоты и является продуктом смещения частоты автогенератора, излучение которого взаимодействует с путевой структурой, и частоты второго СВЧ-генератора, развязанного с первым автогенератором с помощью соответствующих вентильных устройств.

На рис.1 представлена функциональная схема устройства, на рис.2 - общий вид. Полупроводниковый диод 1, размещенный в двухвходовом СВЧ генераторе 2, подключен к источнику питания 3 через фильтр нижних частот 4. К источнику питания 3 через фильтр нижних частот 5 подключен также полупроводниковый диод 6, включенный в высокостабильный СВЧ генератор 7, который через вентиль 8 и управляемый аттенуатор 9 подключен ко второму входу двухвходового СВЧ генератора 2, первый вход которого через фазовращатель 10

подключен к направленному излучателю (антенне) 11, направленному на отражающий объект I-II. Между источником питания 3 и индикатором 12 с помощью разделительных конденсаторов 13 и 14 включен контур низкой частоты 15, образованный индуктивностью 16, резистором 17 и конденсатором 18. Излучатель 11 выполнен в виде пирамидального рупора с соотношением сторон 1,23-1,25, что обеспечивает максимальное поступление отраженного от контролируемого объекта I-II СВЧ сигнала. Двухвходовый СВЧ-генератор выполнен в соответствии с [3], высокостабильный одновходовый генератор аналогичен представленному в [4].

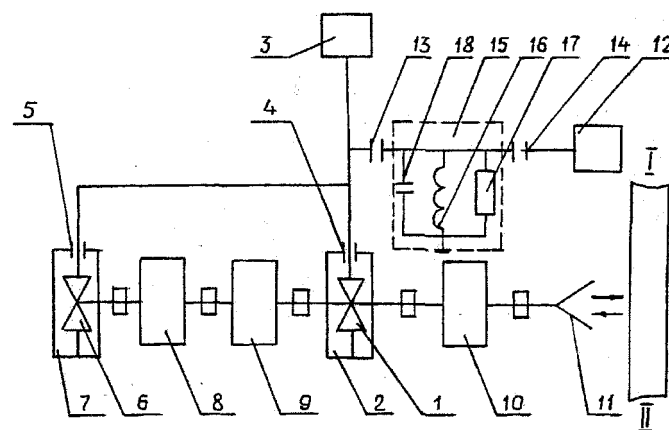


Рисунок 1 - Функциональная схема радиоволнового датчика расстояний между подвижными объектами

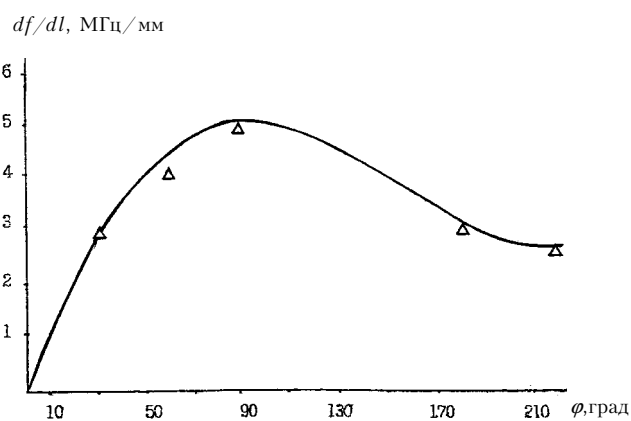


Рисунок 2 - Экспериментальная зависимость крутизны частотного сдвига при различных значениях фазового расстояния между датчиком и перемещающимся отражающим объектом ($|\rho_H| = 0,58$; $f = 9440$ МГц)

Устройство работает следующим образом. При поступлении от источника питания 3 на полупроводниковые генераторные диоды 1 и 6 номинального напряжения двухвходовый СВЧ генератор 2 и высокостабильный

СВЧ генератор 7 генерируют СВЧ сигналы частотой ω_0 и ω_g соответственно, при этом сигнал с частотой ω_0 через фазовращатель 10 и направленный излучатель 11 поступает на опорную отражающую плоскость (контролируемый объект), например, на полотно дороги или боковые направляющие магнитолевитирующего транспортного средства, которые являются высокочастотной нагрузкой двухвходового СВЧ генератора 2, и вследствие эффекта "затягивания частоты" изменение расстояния между устройствами и отражающей плоскостью приводит к изменению генерируемой частоты двухвходовым СВЧ генератором.

Следуя [5], запишем условие устойчивости автоколебательного режима генератора:

$$\frac{g+jb}{G_L\omega_0} = j\left(\frac{\omega}{\omega_0} - \frac{\omega_0}{\omega}\right) + \frac{1}{Q_0} + \frac{G_L+jB_L}{Q_{вн}}, \quad (1)$$

где ω_0 , ω , Q_0 , $Q_{вн}$ - резонансное и текущее значение частоты, собственная и внешняя добротность резонатора, g , b - активная и реактивная компоненты проводимости генераторного полупроводникового диода, G_L , B_L - компоненты проводимости нагрузки. В соответствии с [6], выражение для текущего значения нормированной к волновой проводимости в общем случае может быть записано в виде:

$$y_1 = \frac{1 - \rho_{н}\left(e - j\frac{2\omega}{c}\right)[l + v(t - t_0)]}{1 + \rho_{н}\left(e - j\frac{2\omega}{c}\right)[l + v(t - t_0)]}, \quad (2)$$

где l - расстояние между полупроводниковым диодом 1 и первоначальным положением опорной плоскости ($t = t_0$), $c = 3 \cdot 10^{10}$ см \cdot с $^{-1}$, $\rho_{н}$ - коэффициент отражения волны от опорной отражающей плоскости контролируемого объекта. Решая совместно уравнения (1) и (2), при $Q_0 \gg Q_{вн}$, $|\rho_{н}| \ll 1$ (что чаще всего близко к реальной ситуации) с использованием уравнения Эйлера, $e^{j\varphi} = \cos\varphi + j\sin\varphi$, нетрудно получить аналитические выражения, связывающие изменения частоты генерации и тока через генераторный диод, при вариациях положения нагрузки [7]:

$$\frac{\Delta\omega}{\omega'_0} \approx \frac{\omega - \omega'_0}{\omega'_0} = \frac{|\rho_{н}| \sin\varphi}{Q_{вн}}, \quad (3)$$

где $\varphi = \arg\rho_{н} - 2\omega/c[1 + v(t - t_0)]$, $\arg\rho_{н}$ - аргумент коэффициента отражения.

При неизменном значении модуля коэффициента от-

ражения активной путевой структуры, т.е. $|\rho_{н}| = \text{const}$, и при изменении расстояния между датчиком и путевой структурой, значения частоты генерируемого двухвходовым генератором сигнала являются функцией фазового расстояния между МТС и путевой структурой в соответствии с выражением (3).

Согласно функциональной схеме на полупроводниковый диод 1 через вентиль 8 и управляемый аттенюатор 9 поступает сигнал ω_g , генерируемый высокостабильным СВЧ генератором 7, так что в результате смещения на диоде 1 сигналов с частотами $\omega'_0(\varphi)$ и ω_g на контуре низкой частоты 15 выделяется сигнал разностной частоты $\omega_{н}(\varphi) = \omega'_0(\varphi) - \omega_g$, численное значение которой является функцией фазового расстояния φ между диодом 1 и опорной отражающей плоскостью контролируемого объекта I-II через периодическую функцию $\sin\varphi$, крутизна которой максимальна при $\varphi = n\pi/2$, где $n = 0, 1, 2, \dots$ (см. рис. 2). Для обеспечения высокой чувствительности устройства, т.е. получения максимальных изменений частоты $\omega_{н}(\varphi)$, являющейся информативной, при малых изменениях l - расстояния между диодом и опорной отражающей плоскостью контролируемого объекта, фазовращателем 10 устанавливаются соответствующие первоначальные значения фазового расстояния между устройством и опорной плоскостью.

Для того, чтобы изменения частоты двухвходового СВЧ генератора 2 надежно отслеживали изменения расстояния между подвижными объектами, т.е. для обеспечения надежного функционирования устройства, необходимо избежать явления принудительной синхронизации, т.е. необходимо выполнение условия, чтобы

$$\sqrt{P_1/P_2} \geq \frac{\omega_{н}(\varphi)}{\omega'_0(\varphi)} Q_{вн}, \quad \text{где } P_1, P_2 - \text{мощности сигнала,}$$

генерируемого двухвходовым СВЧ генератором 2, и сигнала, поступающего на диод 1 от высокостабильного СВЧ генератора 7. Соответствующий этим условиям уровень сигнала от высокостабильного СВЧ генератора 7, поступающего на диод 1, регулируется переменным аттенюатором 9, а вентиль 8 предотвращает попадание сигнала от двухвходового генератора 2 в высокостабильный СВЧ генератор 7 и обеспечивает этим стабильность его работы. Так как оба СВЧ генератора 2 и 7 запитаны от общего источника питания 3, то уходы напряжения источника питания дадут примерно равные уходы частот ω'_0 и ω_g , так что значения частоты $\omega_{н}(\varphi)$ при неизменных значениях $|\rho_{н}|$ будут зависеть только от фазового расстояния между устройством и контролируемым объектом, т.е. возможно точное измерение расстояний между подвижными объектами и малых изменений этих расстояний.

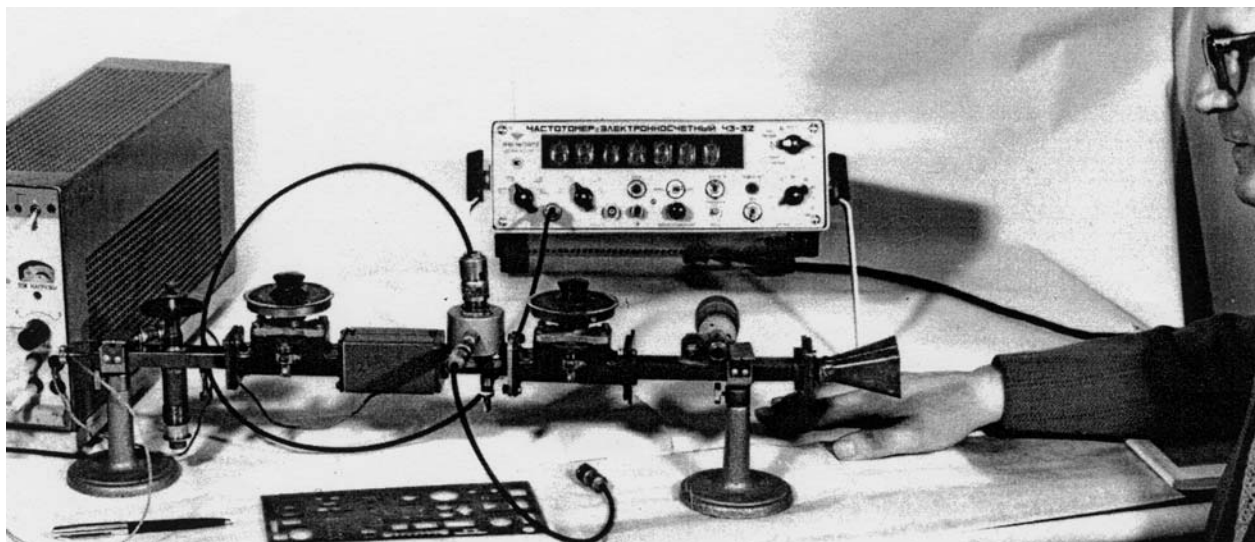


Рисунок 3 - Устройство, общий вид

Так как низкочастотный контур 15 подключен к генераторному диоду 1, то в принципе в этом контуре возможно возникновение автоколебаний [2]. Чтобы предотвратить возбуждение автоколебаний в этом контуре за счет внешней отрицательной проводимости диода, значения параметров диода и указанного контура должны удовлетворять соотношению $1/R_k < g < R_k/\rho_k^2$, где ρ_k - волновое (характеристическое) сопротивление контура, R_k - его омическое сопротивление, $|g|$ - отрицательная проводимость генераторного диода.

Юстировку устройства осуществляют следующим образом. Устанавливают некоторое приемлемое (с точки зрения имеющегося индикатора) значение разностной частоты $\omega_n(\varphi)$ при нормированном (первоначальном) расстоянии между устройством и объектом, регулировкой фазовращателя 10 обеспечивают то положение его органа регулировки, при котором минимальные изменения градуировки фазовращателя обеспечивают максимальное изменение частоты при изменении расстояния.

Информация о перемещении объекта может быть передана по радиоканалу и индикатором может служить частотомер либо традиционный амплитудно-частотный преобразователь, переводящий изменения разностной частоты $\omega_n(\varphi)$ в изменения напряжения, которые затем используются в качестве информативного сигнала в системе управления транспортным средством.

Макет устройства представлен на рис.3 (общий вид). Его испытания были проведены в лабораторных условиях, где в качестве подвижной отражающей плоскости была использована грудь человека (на рис.3 - справа). Осциллограмма соответствующих вариаций расстояния представлена на рис.4: левый фронт кривой на плос-

кости амплитуда-время соответствует вдоху, задний фронт - выдоху; видны все детали этого процесса, в том числе и фаза задержки дыхания. Это означает, в частности, что устройство обладает высоким быстродействием, устойчивостью в работе, что предполагает возможность его использования не только в транспортных системах, для которых оно целенаправленно разрабатывалось, но в ряде технологических процессов, а также в медико-биологических исследованиях.

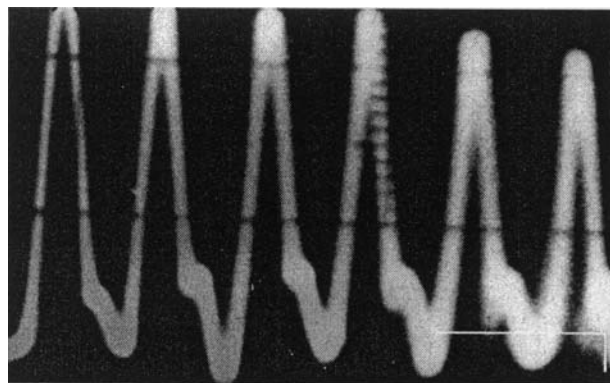


Рисунок 4 - Осциллограмма периодического изменения расстояния между датчиком и подвижным объектом

ПЕРЕЧЕНЬ ССЫЛОК

1. Привалов В.Н. Концепция построения системы управления магнитолевитирующими транспортными средствами / Вісник Дніпропетровського університету. Фізика. Радіоелектроніка, 2000, вип.7, С.85-89
2. Привалов В.Н., Погорелая Л.М., Соколовский И.И. Радиоволновый датчик вертикального зазора для высокоскоростного транспорта на магнитном подвесе / Современные методы и средства электромагнитного контроля и их применение в промышленности. Материалы VI Межвузовской конференции стран СНГ, Могилев, 26-28 сентября 1995, С.10

3. А.с. 1226602 СССР, МКИ4 Н 03 В 7/14 Генератор СВЧ / В.И.Гершун, В.С.Козлов, И.И.Соколовский, С.Д.Шулика. - Оpubл. 23.04.86 Бюл.№15
4. А.с. 1363422 СССР, МКИ 4 Н03 В 7/14 Генератор СВЧ / И.И.Соколовский, В.Я. Крысь, С.В.Плаксин.- Оpubл.30.12.87 Бюл. №48
5. Слэтер Дж. Электроника сверхвысоких частот. - М.: Сов. радио, 1965.- 336 с.
6. Лебедев И.В. Техника и приборы СВЧ - М.: Высшая школа, 1970.- 421 с.
7. Патент України по з-ці № 99063502 від 22.06.99, МПК-6: G01S 13/00, G01P 3/42 Пристрій для вимірювання малих переміщень / В.М.Привалов, І.І.Соколовський, М.Ф.Загуральський, В.В. Коломовець, О.В. Кравченко, О.Ю. Палагін.

УДК 621.396.962

ЭКСПЕРИМЕНТАЛЬНЫЕ ИССЛЕДОВАНИЯ АДАПТИВНОЙ КОМПОЗИЦИОННОЙ СИСТЕМЫ ПРОСТРАНСТВЕННО-ВРЕМЕННОЙ ОБРАБОТКИ СИГНАЛОВ И ТРАДИЦИОННОЙ АДАПТИВНОЙ СИСТЕМЫ ПАРАЛЛЕЛЬНОГО ВЕСОВОГО СУММИРОВАНИЯ

В.П.Прокофьев, Ю.П.Чиняев

Приведены результаты экспериментальной проверки коэффициента подавления активных шумовых помех, быстродействия и чувствительности к аппаратурным ошибкам многоканальных адаптивных систем с параллельной и параллельно-последовательной обработкой сигналов. Полученные результаты подтверждают теоретические выводы о преимуществах систем с параллельно-последовательной обработкой сигналов (композиционных адаптивных систем).

Приведено результати експериментальної перевірки коефіцієнта придушення активних шумових завад, швидкодії і чутливості до апаратурних помилок багатоканальних адаптивних систем з паралельною і паралельно-послідовною обробкою сигналів. Отримані результати підтверджують теоретичні висновки про переваги систем з паралельно-послідовною обробкою сигналів (композиційних адаптивних систем).

The results of experimental evaluation of interference cancellation factor, transient response and sensitivity to intrasystem errors of multi-channel adaptive system with parallel and parallel-serial processing of signals are presented. Obtained results confirm theoretical inferences on advantages of systems with parallel-serial signals processing (composition adaptive systems).

Благодаря широкой области применения, адаптивные методы компенсации помеховых сигналов до сих пор являются предметом интенсивных научных исследований. Применительно к области радиолокации среди систем адаптивной компенсации помех наибольшее распространение, в силу относительной простоты технической реализации, получили адаптивные системы параллельного весового суммирования (системы с параллельной обработкой сигналов), реализующие градиентную процедуру поиска весовых коэффициентов в соответствии с критерием минимизации мощности помехи на выходе [1].

Однако опыт разработки и эксплуатации таких систем свидетельствует о том, что их практическая реализация сопровождается проявлением целого ряда существенных недостатков, присущих всем системам параллельного весового суммирования, что, в конечном итоге, особенно по мере увеличения числа приемных каналов и

независимых источников помех, приводит к резкому ухудшению реальной эффективности.

Основная причина такого отличия реальных показателей качества от потенциально достижимых заложена с самой структуре построения систем параллельного весового суммирования, характеризующейся наличием глобальной обратной связи и реализацией параллельной обработки сигналов (ПОС), что и приводит на практике к проявлению таких недостатков систем с ПОС как низкая устойчивость, высокая чувствительность к различного рода внутрисистемным ошибкам (которые всегда присутствуют в реальной аппаратуре), недостаточно высокое быстродействие и пр.

Одним из способов, позволяющих существенно ослабить недостатки, присущие традиционным системам, является переход к новому методу построения адаптивных систем, названному композиционным [1,2].

Реализованные на базе композиционного метода адаптивные системы с параллельно-последовательной обработкой сигналов (ППОС) отличаются высокой степенью унификации и миниатюризации, повышенной надежностью, простотой структурной реконфигурации и многофункциональностью. При этом следует отметить, что если теоретическим исследованиям таких систем посвящено достаточно большое количество публикаций, то вопросы экспериментальной оценки эффективности их функционирования практически не нашли отражения в научной литературе.

В этой связи представляет несомненный интерес проведение сравнительных экспериментальных исследований адаптивной композиционной системы с параллельно-последовательной обработкой сигналов и традиционной адаптивной системы параллельного весового суммирования.

Исследуемые многоканальные адаптивные системы с параллельной и параллельно-последовательной обработ-

кой сигналов были реализованы на базе однотипных одноканальных модулей квадратурных автокомпенсаторов. Помеховая обстановка задавалась от специально разработанного имитатора активных шумовых помех, имитирующего до пяти независимых источников активных шумовых помех (АШП) с различной пространственной ориентацией и интенсивностью.

На рис.1 приведены зависимости коэффициента подавления АШП ($K_{П}$) четырехканальными ($N = 4$) адаптивной системой с ППОС (сплошные линии) и системой с ПОС (штриховые линии) от количества включенных компенсационных каналов ($N = 1, 2, 3, 4$) при воздействии одного, двух, трех и четырех независимых постановщиков помех (кривые 1,2,3,4 соответственно).

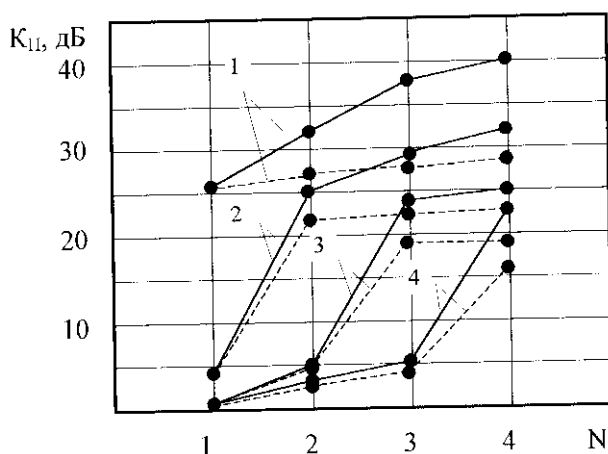


Рисунок 1

Анализ приведенных результатов свидетельствует о том, что в ситуации, когда число включенных каналов компенсации соответствует количеству воздействующих источников помех ($N = M$), выигрыш в коэффициенте подавления помех адаптивной композиционной системы по сравнению с традиционной системой параллельного весового суммирования, по мере увеличения числа каналов компенсации, составляет:

при $N = M = 2$ $\Delta K_{П} = 4$ дБ;

при $N = M = 3$ $\Delta K_{П} = 5$ дБ;

при $N = M = 4$ $\Delta K_{П} = 6$ дБ.

Выигрыш в коэффициенте подавления помех в ситуации структурной избыточности каналов компенсации ($N > M$) по сравнению с линейной ситуацией ($N = M$) у четырехканальной ($N = 4$) композиционной системы составляет:

при $M = 1$ $\Delta K_{П} = 14$ дБ;

при $M = 2$ $\Delta K_{П} = 7$ дБ;

при $M = 3$ $\Delta K_{П} = 2$ дБ,

в то время как традиционная адаптивная система с ПОС выигрыша в коэффициенте подавления практически не имеет. Таким образом, в ситуации структурной избыточности коэффициент подавления помех четырехканальной адаптивной композиционной системы с ППОС на 7...12 дБ выше, чем у существующих в настоящее время адаптивных систем параллельного весового суммирования.

Как отмечалось выше, при практической реализации, а также в процессе функционирования любых сложных систем, неизбежным является наличие различного рода внутрисистемных (аппаратурных) ошибок, которые в той или иной мере могут повлиять на работоспособность системы в целом. Наличие аппаратурных ошибок в адаптивных системах пространственно-временной обработки сигналов приводит к уменьшению практически достижимого коэффициента подавления помех.

При проведении экспериментальных исследований введение ошибок имитировалось путем искусственного ухудшения коэффициента подавления одноканальных модулей автокомпенсации, входящих в структуры исследуемых адаптивных систем.

На рис.2 приведены результаты исследования чувствительности к аппаратурным ошибкам четырехканальной ($N = 4$) адаптивной системы с ПОС (кривая 1) и двух-, трех- и четырехканальной адаптивной композиционной системы с ППОС (соответственно кривые 2,3,4), при воздействии одного ($M = 1$) источника помех (на рис.2 обозначено: $K_{П}$ - коэффициент подавления АШП; $\Delta K_{П}$ - величина, на которую был искусственно ухудшен $K_{П}$ одноканальных модулей автокомпенсации, входящих в структуры исследуемых многоканальных систем).

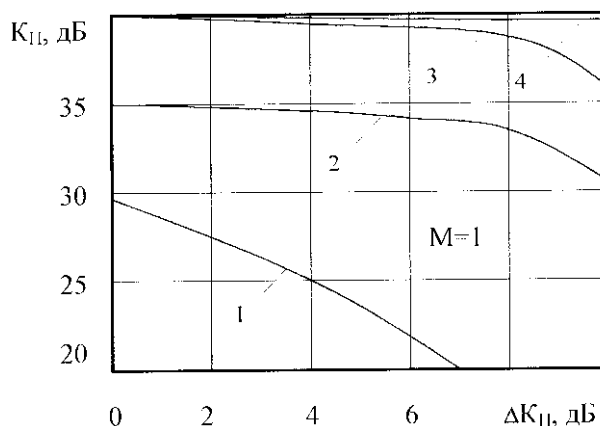


Рисунок 2

Анализ полученных результатов свидетельствует о том, что композиционная система с ППОС в значитель-

ной мере менее критична к наличию аппаратурных ошибок по сравнению с системой параллельного весового суммирования.

Для оценки скорости сходимости переходных процессов на входы исследуемых адаптивных систем подавались гармонические колебания, модулированные прямоугольными импульсами. Длительность переходных процессов определялась с помощью осциллографа как интервал времени от начала переходного процесса до момента, когда отклонение выходного сигнала от его нового установившегося значения становилась меньше $0,1U(t=0)$ (где $U(t=0)$ - амплитуда сигнала в начальный момент переходного процесса).

Результаты исследований, приведенные в табл.1, свидетельствуют о том, что четырехканальная композиционная система с ППОС при воздействии одного источника обеспечивает выигрыш по быстродействию относительно традиционной системы с ПОС примерно в 3 раза.

Таблица 1

| Исследуемая адаптивная система | Длительность переходного процесса, мкс | | | |
|--------------------------------|----------------------------------------|-----|-----|-----|
| | N=1 | N=2 | N=3 | N=4 |
| АС с ПОС | 75 | 60 | 50 | 45 |
| АС с ППОС | 75 | 50 | 30 | 17 |

Динамическая устойчивость исследуемых адаптивных систем исследовалась по их способности сохранять работоспособность при увеличении отношения мощности помехи на входе к мощности собственных шумов. В рамках

проведенных испытаний четырехканальная композиционная система с ППОС обеспечивала устойчивую работу во всех исследованных ситуациях, что в полной мере подтверждает теоретические выводы. Что касается реальной эффективности четырехканальной адаптивной системы с ПОС, то при отношении мощности помехи к мощности собственных шумов равном 60 дБ (что соответствует границе зоны устойчивой работы), коэффициент подавления активных шумовых помех существенно снижается (\approx на 30 %). При дальнейшем увеличении мощности помехи адаптивная система с параллельной обработкой сигналов теряет устойчивость (возбуждается) и становится неработоспособной.

Таким образом, проведенные экспериментальные исследования свидетельствуют о несомненных преимуществах композиционных систем обработки сигналов по сравнению с традиционными системами. Использование композиционного метода при разработке сложных радиоэлектронных средств позволит реализовать на практике многоканальные адаптивные системы различного класса и назначения с такими высокими эксплуатационными характеристиками, которые не достижимы в ранее известных решениях.

ПЕРЕЧЕНЬ ССЫЛОК

1. Прокофьев В.П., Неволько А.И., Чиняев Ю.П. Синтез адаптивной системы компенсации помех с параллельно-последовательной обработкой сигналов // Проблемы управления и информатики. - 1997. - № 6. - С. 66-69.
2. Прокофьев В.П., Чиняев Ю.П. Адаптивные радиоэлектронные устройства. Состояние проблемы, динамика развития / Арсенал XXI века. - 2000. - № 1. - С.37-39.

УДК 621.37

ВЛИЯНИЕ РАЗБРОСА ТЕПЛОФИЗИЧЕСКИХ ПАРАМЕТРОВ ТРАНЗИСТОРОВ НА ХАРАКТЕРИСТИКИ СИММЕТРИЧНЫХ ТРАНЗИСТОРНЫХ СХЕМ

В.А.Сергеев

Исследуется изменение положения рабочей точки транзисторов, вызванное разностью тепловых импедансов прибора в дифференциальном и двухтактном каскадах.

Досліджується зміна положення робочої точки транзисторів, викликана відмінністю теплових імпедансів приладів у диференційному та двотактному каскадах.

The variation of transistor point position based on the difference between the device thermal impedances in differential and two-tact cascades is investigated.

ВВЕДЕНИЕ

Большая группа транзисторных схем характеризуется симметричным включением основных активных элементов, которые по определению должны работать в одинаковых электрических режимах. Такие схемы принято называть симметричными. К ним относятся, например, дифференциальный каскад, симметричный мультивибратор, двухтактный (балансный) каскады усилителей мощности, схема токового зеркала и др. [1,2].

К транзисторам, которые используются в подобных

схемах, как правило, предъявляют достаточно жесткие требования по идентичности электрических параметров, таких как β , I_{K0} , R_{BX} и др. При расчетах характеристик таких симметричных схем электрические параметры симметрично включенных транзисторов всегда полагаются одинаковыми. Однако, как показывает приведенный ниже анализ, не менее важным является требование идентичности теплофизических параметров, поскольку даже небольшая (в несколько Кельвин) разность температур активных областей симметричных транзисторов может свести на нет идентичность электрических параметров.

Детальный анализ влияния разброса теплофизических параметров транзисторов на характеристики всех возможных вариантов симметричных схем представляет предмет отдельного серьезного исследования. Здесь мы ограничимся оценкой изменения положения рабочей точки транзисторов, вызванного разностью тепловых импедансов приборов в дифференциальном каскаде и в двухтактном каскаде класса В.

Следует отметить, что разброс теплофизических параметров в значительно меньшей степени отражается на характеристиках импульсных симметричных схем, т.к. рабочие точки находятся либо в области отсечки, либо в области насыщения, а тепловая обратная связь слабо проявляется в ключевом режиме [3].

1. ДИФФЕРЕНЦИАЛЬНЫЕ КАСКАДЫ

На рис. 1 приведена схема дифференциального каскада (ДК).

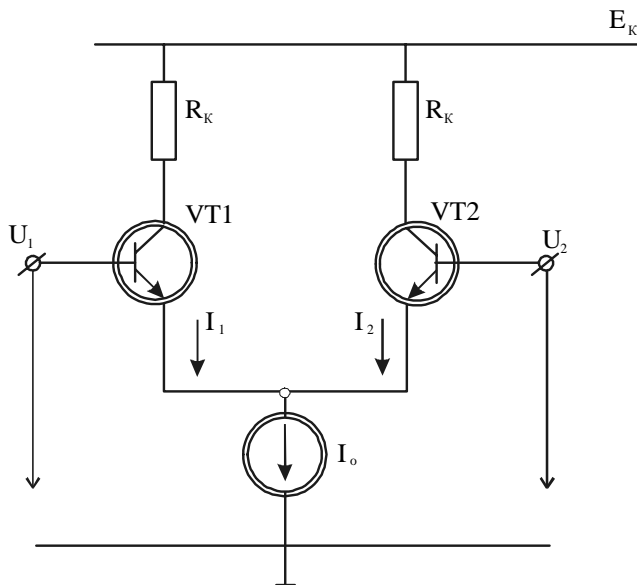


Рисунок 1 - Схема дифференциального каскада

Выходной сигнал в ДК может сниматься между коллекторами транзисторов VT1 и VT2 (ДК с симметричным выходом) или с коллектора одного из транзисторов VT1 или VT2 (ДК с несимметричным выходом) [4].

В [5] показано, что разность температур $\Delta T = T_1 - T_2$ связана с разностью токов $\Delta I = I_1 - I_2$ соотношением

$$\Delta T_{1-2} = \frac{2 \operatorname{arctg} \delta}{\varepsilon_T}, \quad (1)$$

где $\delta = \Delta I / I_0$ - относительная разность токов, $\varepsilon_T = (E_g - qU) / kT_1 T_2$, E_g - ширина запрещенной зоны полупроводника, q - заряд электрона, k - постоянная Больцмана.

При небольших разностях токов ($\delta < 0,4$), пренебрегая в разложении $\operatorname{arctg} \delta = \delta + \delta^3/3 + \delta^5/5 + \dots$ членами выше первой степени малости, с погрешностью не более 5% (1) можно записать в виде

$$\Delta T_{1-2} = \frac{2kT_1 T_2}{E_g - qU} \delta \approx \xi_T \delta, \quad (2)$$

где $\xi_T^{[k]}$ - температурный коэффициент, практически постоянный в достаточно широком диапазоне температур $\xi_T \approx 30 \div 40 \text{ K}$. Соответственно для приращений температур можно записать:

$$\Delta T_1 = R_{T1} (E_k - I_1 R_K) I_1, \quad (3.1)$$

$$\Delta T_2 = R_{T2} (E_k - I_2 R_K) I_2. \quad (3.2)$$

Выражая I_1 и I_2 через δ : $I_1 = \frac{I_0}{2}(1 + \delta)$ и $I_2 = \frac{I_0}{2}(1 - \delta)$; и подставляя эти выражения в уравнения (3.1) и (3.2), а затем, вычитая (3.2) из (3.1) после несложных преобразований получим, квадратное уравнение относительно δ :

$$\delta^2 + \frac{R_{T1} + R_{T2}}{R_{T1} - R_{T2}} \left[\frac{\xi_T}{(R_{T1} + R_{T2}) R_K I_0^2} - \theta \right] \delta - (\theta + 1) = 0, \quad (4)$$

где $\theta = 2E_k / R_K I_0 - 2$.

Физический смысл имеет корень уравнения, удовлетворяющий условию $\delta < 1$. Поскольку в большинстве практических случаев тепловые сопротивления транзисторов удовлетворяют условию $(R_{T1} - R_{T2}) \ll R_{T1} + R_{T2}$, то решение (4) существенно упрощается и в линейном приближении имеет вид:

$$\delta = \frac{(R_{T1} - R_{T2})(\theta + 1) R_K I_0^2}{\xi_T - \frac{R_K I_0^2}{4} (R_{T1} + R_{T2}) \theta}. \quad (5)$$

Для реальных ДК обычно $\theta \ll 1$, так что для практических расчетов можно записать

$$\frac{\Delta I}{I_0} = \delta = \frac{(R_{T1} - R_{T2}) \left(\frac{R_K I_0^2}{4} \right)}{\xi_T}. \quad (6)$$

Таким образом, разбаланс ДК в статическом режиме пропорционален разности тепловых сопротивлений транзисторов.

При подаче на вход ДК переменного сигнала мгновенная рассеиваемая транзисторами мощность будет содержать две составляющих: постоянную составляющую, которая будет отличаться от рассеиваемой мощности в статическом режиме, и переменную составляющую мощности, которая будет изменяться по закону изменения входного сигнала. Действительно, если положить входной сигнал гармоничным $\Delta U = U_1 - U_2 = U_m \sin \omega t$, то соответственно, токи I_1 и I_2 будут иметь гармонические противофазные составляющие:

$$I_1 = I_{10} + I_m \sin \omega t; \quad (7.1)$$

$$I_2 = I_{20} - I_m \sin \omega t. \quad (7.2)$$

Тогда мощности, рассеиваемые на коллекторном $p-n$ переходе транзисторов, будут равны:

$$P_{K1} = P_{10} - P_{1m} \sin \omega t - P'_{1m} \cos \omega t; \quad (8.1)$$

$$P_{K2} = P_{20} - P_{2m} \sin \omega t - P_{2m} \cos \omega t, \quad (8.2)$$

где

$$P_{i0} = \left(I_i E_K - R_K I_{i0}^2 - \frac{R_K I_m^2}{2} \right), \quad (9)$$

$$P_{im} = (-1)^i (E_K - 2R_K I_{i0}) I_m, \quad (10)$$

$$P'_{im} = \frac{R_K I_m^2}{2}, \quad (11)$$

а $i = 1, 2$ - номер транзистора.

Переменная составляющая мощности P_{im} приводит к появлению переменной составляющей температуры, амплитуда которой при $\omega \ll \tau_{T1,2}$ ($\tau_{T1,2}$ - тепловые постоянные времени переход-корпус транзисторов) определяется выражением:

$$T_{im} = (-1)^i (E_K - 2R_K I_i) R_{Ti}. \quad (12)$$

В результате действия тепловой обратной связи с коэффициентом $K_T \approx -2, 2$ мВ/К на входах транзистора появляется дифференциальный сигнал

$$\Delta U_m = K_T I_m [(E_K - 2R_K I_{10}) R_{T1} - (E_K - 2R_K I_{20}) R_{T2}] = K_T I_m [(E_K - R_K I_0) (R_{T1} - R_{T2}) + 2R_K \Delta I (R_{T1} + R_{T2})], \quad (13)$$

который является в общем случае нежелательным, так как приводит к дополнительным линейным искажениям.

Подставляя в (13) выражение для ΔI из (6), получим

$$\Delta U_m = K_T I_m \left[(E_K - R_K I_0) + R_K I_0 \frac{R_{T1} + R_{T2} R_K I_0^2}{\xi_T} \right] (R_{T1} - R_{T2}). \quad (14)$$

Соответственно, коэффициент линейных искажений, возникающих в дифференциальном каскаде из-за различия тепловых сопротивлений на низкой частоте, будет равен:

$$K_{ЛТ} = \frac{\Delta U_m}{U_m} = \frac{K_T K_Y}{R_K} \left[(E_K - R_K I_0) + \frac{R_K I_0}{2(\xi_T - \theta)} \right] (R_{T1} - R_{T2}), \quad (15)$$

где K_Y - коэффициент усиления дифференциального сигнала.

На высоких частотах $\omega \gg \tau_{T1}$ линейные искажения этой природы будут исчезать в результате инерционности тепловых процессов. Указанный эффект может быть использован для оценки разности тепловых сопротивлений транзисторов путем их включения по дифференциальной схеме.

Переменные составляющие мощности $P'_{1m} \cos 2\omega t$ и $P'_{2m} \cos 2\omega t$ создают синфазные дополнительные сигналы на выходе усилителя, которые в ДК с симметричным выходом подавляются дифференциальной схемой.

2. ДВУХТАКТНЫЕ УСИЛИТЕЛЬНЫЕ КАСКАДЫ

В двухтактных усилительных каскадах класса В независимо от схемы реализации в состоянии покоя (т.е. в отсутствие сигнала) рабочая точка обоих транзисторов находится в области отсечки и рассеиваемая мощность очень мала и не может привести к сколь-нибудь заметному изменению положения рабочей точки.

Наиболее распространенными схемами усилительных каскадов класса В являются трансформаторная схема с дополнительным смещением (рис.2), а также бестрансформаторная схема каскада на комплементарных транзисторах (рис.3).

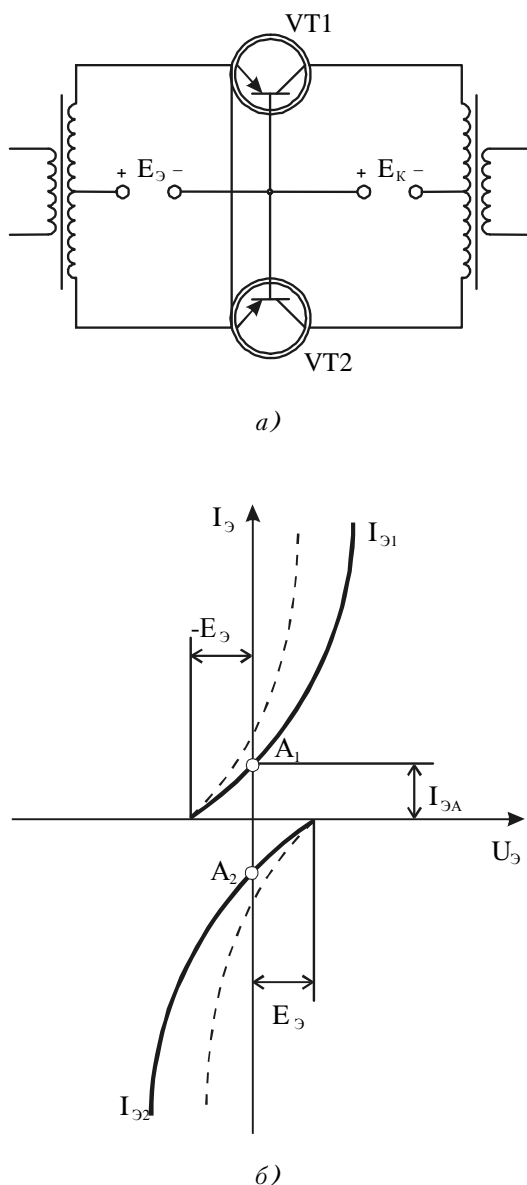


Рисунок 2 - Трансформаторная схема каскада класса В (а) и суммарная входная характеристика (б)

Входные вольтамперные характеристики для обеих схем будут иметь одинаковый вид, показанный на рисунке 2. Подсмещение в схеме на рисунке 3 обеспечивается резистором R_2 , вместо которого иногда включают диоды.

При подаче на вход двухтактного каскада гармонического сигнала каждый из транзисторов будет поочередно усиливать соответствующую полуволну сигнала. Для мгновенной мощности рассеяния на коллекторном переходе каждого из транзисторов можно записать следующее выражение:

$$P_K(t) = E_K(1 - \sin \omega t)I_m \sin \omega t = E_K I_m \sin \omega t - \frac{E_K I_m}{2} + \frac{E_K I_m}{2} \cos \omega t, \quad (16)$$

где I_m - амплитуда переменной составляющей коллекторного тока транзисторов.

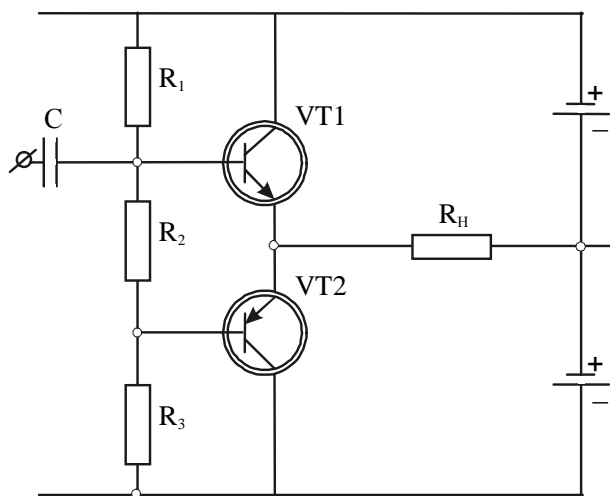


Рисунок 3 - Двухтактный каскад класса В с дополнительной симметрией

На низких частотах $2\omega < \tau_{T_i}$ температура кристалла транзисторов будет изменяться по закону изменения греющей мощности. Максимальное значение температуры $p-n$ перехода i -го транзистора будет равно:

$$T_{ni} = T_0 + R_{T_i} E_K I_m. \quad (17)$$

Разность пиковых температур переходов соответственно будет равна

$$\Delta T = E_K I_m (R_{T1} - R_{T2}). \quad (18)$$

На высоких частотах после усреднения (16) для разности средних температур переходов получим

$$\overline{\Delta T} = \frac{E_K I_m}{\pi} (R_{T1} - R_{T2}). \quad (19)$$

Разность мгновенных температур переходов будет еще более значительной, так как температура транзисторных структур изменяется в противофазе. Однако более существенной является разность средних температур $p-n$ переходов.

Разность средних температур приводит к различному смещению рабочих точек A_1 и A_2 на рисунке 2 и соответственно к различию крутизны сквозных входных характеристик S_1 и S_2 для различных полуволн. Различие крутизны будет давать дополнительные нелинейные искажения сигнала, величина которых определяется разностью тепловых сопротивлений транзисторов. Оценить эти искажения можно методом трех ординат [6], соглас-

но которого амплитуды первой I_{m1} и второй гармоники I_{m2} выходного тока вычисляются по формулам:

$$I_{m1} = \frac{I_{\max} - I_{\min}}{2}; I_{m2} = \frac{I_{\max} + I_{\min} - 2I_0}{4}, \quad (20)$$

где $I_{\max} \approx S_1 U_{\text{ВХ}}$ - амплитуда выходного тока первого транзистора, $I_{\min} \approx -S_2 U_{\text{ВХ}}$ - амплитуда выходного тока второго транзистора. Учитывая, что $I_0 \approx 0$ из (20) получим выражение для коэффициента гармоник:

$$K_{\Gamma} = \frac{I_{m2}}{I_{m1}} = \frac{S_1 - S_2}{2(S_1 + S_2)}. \quad (21)$$

Если положить, что крутизна меняется пропорционально температуре, т.е. $S_i = S_0 + K_S \Delta T_i$, то коэффициент гармоник будет равен

$$K_{\Gamma} = \frac{K_S E_K I_m (R_{T1} - R_{T2})}{4\pi S_0}, \quad (22)$$

где K_S - температурный коэффициент крутизны сквозной характеристики.

Таким образом, разность тепловых сопротивлений транзисторов в выходных каскадах класса В является причиной дополнительных нелинейных искажений сигналов на низкой частоте.

УДК 658.512.2

МОДЕЛЮВАННЯ ПРОФІЛІВ РОЗПОДІЛУ ДОМІШКОВИХ ІОНІВ, ВАКАНСІЙ ТА МІЖВУЗЛОВИХ АТОМІВ У НАПІВПРОВІДНИКОВОМУ МАТЕРІАЛІ ПРИ ІОННОМУ ЛЕГУВАННІ

В.М.Теслюк, О.Р.Корбецький

В работе предложена математическая модель ионной имплантации, которая основывается на методе Монте-Карло и позволяет увеличить точность моделирования, учитывая основные законы взаимодействия иона с атомами многокомпонентной мишени.

В роботі запропоновано математичну модель іонної імплантації, яка базується на методі Монте-Карло, що дозволяє підвищувати точність моделювання враховуючи основні закони взаємодії іона з атомами багатоконпонентної мішені.

The mathematical model of ion implantation is proposed in this paper, which is based on Monte-Carlo method and allows to increase an accuracy of simulation with taking into account the main laws of interaction between ion and atoms of multicomponent target.

ВЫВОДЫ

Проведенный выше анализ показывает, что разброс тепловых сопротивлений транзисторов приводит к появлению дополнительных линейных и нелинейных искажений в симметричных транзисторных схемах. Для оценки величины этих искажений в дифференциальном каскаде класса В можно использовать полученные соотношения. Учитывая, что тепловые сопротивления реальных мощных транзисторов могут различаться в несколько раз [7], при разработке и изготовлении прецизионной радиоэлектронной аппаратуры с использованием симметричных транзисторных каскадов следует предусматривать специальный контроль и подбор транзисторов по величине теплового сопротивления.

ПЕРЕЧЕНЬ ССЫЛОК

1. Степаненко И.П., Основы теории транзисторов и транзисторных схем. - Изд.4-е, перераб. и доп. М.: Энергия, 1977 - 672 с.
2. Хоровиц П., Хилл У. Искусство схемотехники. В 3-х томах: Т.1/Пер. с англ.-М.: Мир, 1993.- 413 с.
3. Сергеев В.А. Тепловые переходные процессы в ключевых транзисторных схемах // Проблемы и решения современной технологии: Сборник научных трудов ПТИС, выпуск 6, часть II.- Тольятти, изд-во ПТИС, 2000.- С.34 - 39.
4. Волгин Л.И. Линейные электрические преобразования для измерительных приборов и систем. - М.: Советское радио, 1971.- 334 с.
5. А.С. СССР 1247796 Устройство для отбраковки мощных транзисторов/В.А. Сергеев, П.А. Голенкин.-Бюл. изобр., 1986, №28.
6. Головин О.В., Кубицкий А.А. Электронные усилители: Учебник для техникумов связи.-М.: Радио и связь, 1983 - 320 с.
7. Сергеев В.А., Контроль качества и отбраковка мощных транзисторов по теплофизическим параметрам.- Научно-технический калейдоскоп. №1, 2000.-С.37 - 43.

ВСТУП

Технологія виробництва великих інтегральних схем (ВІС) щорічно вдосконалюється та розробляються нові підходи при виготовленні інтегральних приладів.

На сьогоднішній день проектний розмір виготовлення мікропроцесорів досягнув значення 0.35-0.13мкм. При таких мінімальних проектних нормах існуючі математичні моделі технологічних процесів не задовольняють вимогам, які поставлені до точності та адекватності вихідних результатів техпроцесу. Перш за все, це пов'язано з впливом ефектів тунелювання, гарячих електронів [1], співмірністю області просторового заряду біля р-п пере-

ходу з мінімальним проектним розміром та ін.

Враховувати відповідні особливості виробництва інтегральних схем, можна шляхом використання методу Монте-Карло [2] та врахуванням взаємодії кожного атома домішки з кристалічною ґраткою напівпровідника.

Легування напівпровідникових підкладок можна проводити в основному з допомогою дифузії [3] чи іонної імплантації [4]. Якщо порівняти ці два методи, то іонна імплантація має ряд суттєвих переваг [5]. По-перше, вона є низькотемпературною технооперацією і при проведенні її не відбувається перерозподілу домішок, які вже присутні в напівпровідниковому кристалі. По-друге, при легуванні напівпровідника з допомогою дифузії концентрація домішки в підкладці не може перевищити рівноважної розчинності дифузанта при температурі процесу (900-1200°C). Дане обмеження відсутнє для іонної імплантації. По-третє, іонне легування, у порівнянні з дифузєю, є більш керованим техпроцесом і дозволяє впровадити в кристал задану кількість домішки з вищою точністю.

МАТЕМАТИЧНА МОДЕЛЬ ІОННОЇ ІМПЛАНТАЦІЇ

Запропонована в роботі модель базується на класичній теорії зіткнень між частинками [6]. Кут повороту іона Φ_0 (рис.1) визначається з наступної формули:

$$\Phi_0 = \int_{r_{\min}}^{\infty} \frac{\rho dr}{r^2 \sqrt{1 - \frac{\rho^2}{r^2} - \frac{V(r)}{E_{\text{відн}}}}} , \quad (1)$$

де r - міжатомна відстань, ρ - прицільний параметр; $E_{\text{відн}}$ - відносна енергія [6]; $V(r)$ - потенціал міжатомної взаємодії; r_{\min} - відстань максимального зближення частинок.

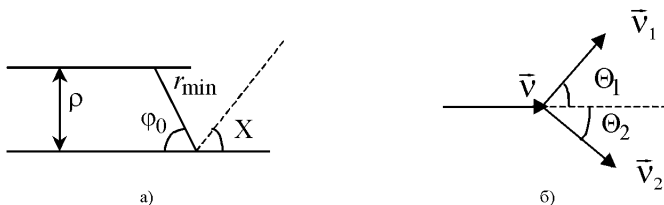


Рисунок 1 - Розсіяні частинки в системі: а) центра мас; б) в лабораторній системі координат

Відомо, що для $r < 1A^0$ реальний потенціал добре описується кулонівським екранованим потенціалом з функцією екранування, що розрахована на моделі Томсона-Фермі:

$$V(r) = \frac{z_1 \cdot z_2 \cdot e^2}{r} \cdot X\left(\frac{r}{a}\right), \quad a = \frac{0,885a_B}{(z_1^{2/3} + z_2^{2/3})^{1/2}}, \quad (2)$$

де a - масштабний фактор; a_B - радіус першої борівської орбіти $a_B = 0,523A^0$, z_1 - заряд ядра іона, z_2 - заряд атома мішені. Виявляється, що для $2 \leq x \leq 10$, $X(x)$ досить точно апроксимується наступним виразом $X_{\text{ann}}(x) = \frac{0,576}{x} - (0,037)$, де $x = r/a$.

Тоді для V_{ann} отримаємо:

$$V'_{\text{ann}}(r) = \frac{\alpha_1}{r} + \frac{\alpha_2}{r^2}, \quad (3)$$

де $\alpha_1 = -0,037z_1z_2e^2$; $\alpha_2 = 0,576z_1z_2e^2a$; (4)

На невеликих відстанях ($2a_B \leq r \leq 3a_B$) потенціал Томсона-Фермі (2) дає завищені значення. Найкращу відповідність з експериментальними даними дає потенціал Томсона-Фермі-Дірака, розрахований з врахуванням обмінної взаємодії. На великих відстанях він добре апроксимується потенціалом Борна-Майєра:

$$V(r) = Ae^{r/a}. \quad (5)$$

Для констант A і a були отримані значення:

$$A = 154(z_1 \cdot z_2)^{3/4}, \quad a = 0,195A^0. \quad (6)$$

З цим потенціалом ми обчислюємо повний переріз розсіювань. Для обчислення інтегралу (1) використовується апроксимація:

$$V'_{\text{ann}}(r) = \frac{\alpha_1}{r} + \frac{\alpha_2}{r^2} + \alpha_3. \quad (7)$$

Параметри α_1 , α_2 , α_3 визначаються виходячи з рівності (7) з відповідним потенціалом на активній ділянці траєкторії атомів, що зіштовхувались, тобто для r в інтервалі $r_{\min} \leq r \leq \rho_0$, де r_{\min} - відстань між атомами в точці повороту; r_{\min} визначається з умови:

$$1 - \frac{\rho_0^2}{r_{\min}^2} - \frac{V(r_{\min})}{E_{\text{відн}}} = 0, \quad (8)$$

де ρ_0 - радіус повного перерізу пружного розсіювання. Для Φ_0 в тому випадку отримуємо вираз:

$$\Phi_0 = \frac{\rho}{\sqrt{T_2}} \left(\frac{\pi}{2} - \arcsin \frac{T_1 + \frac{T_2}{\rho_0}}{\sqrt{T_2 T_3 + \pi^2}} \right) + \arcsin \frac{\rho}{\rho_0}, \quad (9)$$

$$\text{де } T_1 = \frac{\alpha_1}{2E_2}; T_2 = \rho^2 + \frac{\alpha_2}{E_2}; T_3 = 1 - \frac{\alpha_3}{E_2}; E_2 = \frac{E}{\alpha + 1};$$

$$\alpha_3 = Ae^{-1/2} - \frac{\alpha_1}{r} - \frac{\alpha_2}{r^2}.$$

Якщо відомий потенціал міжатомної взаємодії і можна взяти інтеграл (1), то кути розсіювання в лабораторній системі координат обчислюються за формулою:

$$\text{tg} \Theta_1 = \frac{\sin 2\Phi_0}{\alpha - \cos 2\Phi_0}, \quad \Theta = \Phi_0, \quad (10)$$

де $\alpha = m_1/m_2$, m_1 - маса іона, m_2 - маса атома мішені.

Як видно з (1) результат розсіювання залежить від потенціалу міжатомної взаємодії.

Середня довжина вільного пробігу при статистичному рівномірному розподілі атомів може бути визначена наступним чином:

$$\lambda(E) = \frac{1}{n_a \pi \rho_0^2(E)}, \quad (11)$$

де $\pi \rho_0^2$ - повний переріз пружних зіткнень. Радіус ρ_0 можна знайти як параметр, який розділяє далекі і близькі зіткнення. Далекі зіткнення можна врахувати як втрати енергії на безперервне гальмування. Вони були розраховані в імпульсному наближенні з потенціалом (2) і виявилось, що для енергій більших 1 KeV ці втрати - непружні.

Втрати на електронне гальмування можна розрахувати по формулі, отриманій Ліндхардом:

$$\frac{dE}{dx} = K \sqrt{E}, \quad (12)$$

$$\text{де } K = 2\pi a_B z_1^{7/6} z_2 n_a / (z_1^{2/3} + z_2^{2/3})^{3/2}.$$

При $E\rho_0$ можна знайти як такий параметр зіткнення, розв'язавши рівняння $Q(\rho_0) = E_d$.

$$\text{Отже, передана енергія } Q = \frac{4m_1 m_2}{(m_1 + m_2)^2} E \cos^2 \Phi_0,$$

де Φ_0 - кут розсіювання в системі центру мас, який визначається в імпульсному наближенні наступною формулою:

$$\Phi_0 = \frac{(\alpha + 1)\rho}{E} \int \frac{dV}{dr} \cdot \frac{dr}{\sqrt{r^2 - \rho^2}}. \quad (13)$$

Для потенціалу Борна-Майєра отримуємо:

$$\Phi_0 = \frac{\alpha + 1}{E} \cdot \rho \cdot \frac{A}{a} \cdot K_0\left(\frac{\rho}{a}\right), \quad (14)$$

де K_0 - функція МакДональдса нульового порядку. Для великих ρ/a функція $K_0(x) = \sqrt{\frac{\pi}{2x}} e^{-x}$.

Отже рівняння для визначення ρ_0 має вигляд:

$$\alpha = \frac{A^2}{E} \cdot \frac{\rho_0^2}{a^2} \cdot K_0\left(\frac{\rho_0}{a}\right) = E_d. \quad (15)$$

Звідси можна отримати ρ_0 методом послідовних наближень. В першому наближенні:

$$\rho_0 = \frac{a}{2} \ln \frac{\alpha A^2 c \pi}{r E E_d}. \quad (16)$$

При низьких енергіях ρ_0 можна знаходити з вимоги $\Theta_1(\rho) \leq \beta$, де Θ_1 - кут розсіювання в лабораторній системі координат. Кут β має малі значення. Ми покладемо $\beta = 3^0$. Це приводить до наступної умови: $Q(\rho) \leq \alpha E \beta^2$, звідки отримуємо:

$$\rho_0 = \frac{a}{2} \ln \frac{\pi A^2 c}{2 E^2 \beta^2}.$$

Знаючи повний переріз пружних зіткнень, можна розіграти вільний пробіг у відповідності з імовірністю $P(l) = e^{-l/\lambda} \cdot d \cdot 1/\lambda$. Однак слід мати на увазі, що рух іона відбувається в гальмуючому середовищі. Тому довжина вільного пробігу

$$l = \int \frac{d\varepsilon}{d\varepsilon/dx}. \quad (17)$$

Оскільки $\lambda = \frac{1}{n_a \pi \rho_0^2(E)}$, то можна записати:

$$\frac{1}{\lambda} = \int_E^{E'} \frac{d\varepsilon}{(d\varepsilon/dx) \lambda(\varepsilon)}. \quad (18)$$

Після підстановки (18) ми отримуємо ймовірність того, що в кінці вільного пробігу атом буде мати енергію $E = E'$, якщо на початку його енергія була E :

$$P(E, E')dE' = C \exp \left(- \int_E^{E'} \frac{d\varepsilon}{(d\varepsilon/dx)\lambda(\varepsilon)} \right) \frac{dE'}{(dE'/dx)\lambda(E')} \quad (19)$$

Знаючи $P(E, E')$, можна визначити у відповідності з узгодженою раніше процедурою. В результаті отримується рівняння $\int_E^{E'} \frac{d\varepsilon}{(d\varepsilon/dx)\lambda(\varepsilon)} = -\ln \gamma$. Розв'язки рівнянь мають вигляд:

$$E' = \left[\sqrt{E} + \frac{2k \ln \gamma}{n_a \pi a^2 \left[\left(\ln \frac{\pi A^2 \alpha \rho_0}{2EE_0} + 2 \right)^2 + 4 \right]} \right] \quad (20)$$

Розігравши E' , можна обчислити шлях між зіткнення

$$S_i = \int_{E_i}^{E'_i} \frac{d\varepsilon}{d\varepsilon/dx}.$$

Для визначення положення атома в просторі потрібно знати направляючі косинуси вектора швидкості \vec{v}_i після i -го зіткнення. Якщо розглянути дві системи координат - систему пов'язану з мішенню і систему з рухомими атомами (вісь x співпадає з \vec{v}_i), то, провівши перетворення від системи \vec{v}_i до системи \vec{v}_{i+1} , можна знайти рекурентні співвідношення, які пов'язують направляючі косинуси \vec{v}_{i+1} по відношенню до нерухомої системи з направляючими косинусами \vec{v}_i :

$$\vec{v}_i = \cos \alpha_i \vec{x}_0 + \cos \beta_i \vec{y}_0 + \cos \gamma_i \vec{z}_0;$$

$$\vec{v}_{i+1} = \cos \alpha_{i+1} \vec{x}_0 + \cos \beta_{i+1} \vec{y}_0 + \cos \gamma_{i+1} \vec{z}_0.$$

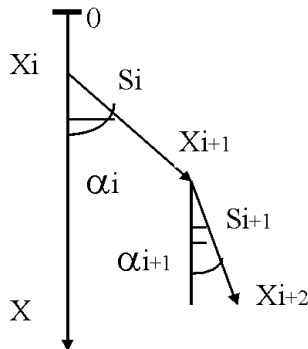


Рисунок 2 - Схема руху іона в мішені

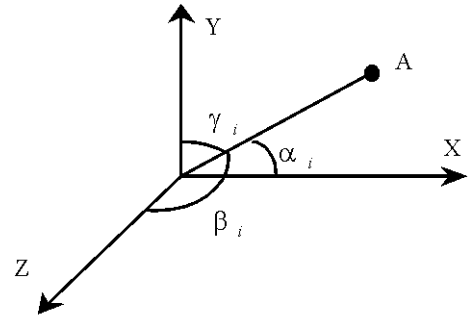


Рисунок 3 - Система координат

Отримаємо [7]:

$$\begin{aligned} \cos \gamma_{i+1} &= \\ &= \cos \gamma_i \cos \Theta_i + \cos \alpha_i \sin \Theta_i \cos \varphi v_2 - \sin \Theta_i \sin \varphi v_1, \end{aligned}$$

$$\begin{aligned} \cos \beta_{i+1} &= \\ &= \cos \beta_i \cos \Theta_i + \cos \alpha_i \sin \Theta_i \cos \varphi v_1 - \cos \Theta_i \sin \varphi v_2, \end{aligned} \quad (21)$$

$$\cos \alpha_{i+1} = \cos \alpha_i \cos \Theta_i - \sin \alpha_i \cos \Theta_i \cos \varphi.$$

Тут Θ_i - кут розсіювання в лабораторній системі координат після i -го зіткнення, φ - кут повороту площини зіткнення, який розігрується за формулою:

$$\varphi = 2\pi\gamma; \quad v_1 = \frac{\cos \beta_i}{\sin \alpha_i}; \quad v_2 = \frac{\cos \gamma_i}{\sin \alpha_i}.$$

Тоді координати атома після i -го зіткнення[7]:

$$x_{i+1} = x_i + S_i \cos \alpha_i,$$

$$y_{i+1} = y_i + S_i \cos \beta_i, \quad (22)$$

$$z_{i+1} = z_i + S_i \cos \gamma_i.$$

АЛГОРИТМ РОЗВ'ЯЗКУ ЗАДАЧІ

При розрахунку пробігів іонів у напівпровідниковій мішені методом Монте-Карло розігрують параметри зіткнень, пробіги між двома послідовними зіткненнями, кут повороту площини зіткнення, сорт атома, з яким відбудеться зіткнення (для багатокомпонентної моделі). При цьому використовується наближення парних зіткнень і вважається, що атоми статистично розподілені з деякою середньою густиною n_0 . В кожному досліді знаходять повний траєкторний пробіг $R = \sum l_i$, $x = \sum l_i \cdot \cos \alpha_i$, $y = \sum l_i \cdot \cos \beta_i$, $z = \sum l_i \cdot \cos \gamma_i$ (x, y, z - координати точки, де іон зупинився). Провівши N випробувань можна побудувати гістограми розподілу іонів у мішені

(рис.6,9), а також обчислити [8] \bar{x} , \bar{y} , \bar{z} , \bar{x}^2 , \bar{y}^2 , \bar{z}^2 і т. п. При зіткненні швидкого іона з атомами ґратки останні можуть бути зміщені атоми мішені і вони можуть отримати достатню енергію для того, щоб вибивати з вузлів ґратки атоми. В результаті утворюються каскади атомних зміщень [6](рис.5).

В процесі розвитку каскаду утворюються точкові дефекти [9],[10]:

- Вакансії (V) (розподіл вакансій наведений на рис.7);
- Міжвузлові атоми (I) (розподіл міжвузлових атомів на рис.8).

При бомбардуванні напівпровідникової мішені іонами відбувається її розпилення[11]. Для моделювання даного ефекту і визначення товщини стравленого шару необхідно визначити коефіцієнт розпилення, який, в даному випадку, розраховується, як відношення кількості вибитих іонів домішки (координата $x < 0$, рис.7) та атомів мішені до загальної кількості іонів N [12].

Точність розрахунку залежить від кількості випробувань (похибка $\approx 1/\sqrt{N}$), рис.4). Точність запропонованої математичної моделі залежить від виразу, який апроксимує потенціал міжатомної взаємодії, коректного врахування непружних втрат енергії іонів та ефектів на поверхні напівпровідникової пластини (ефектів відбивання та розпилення).

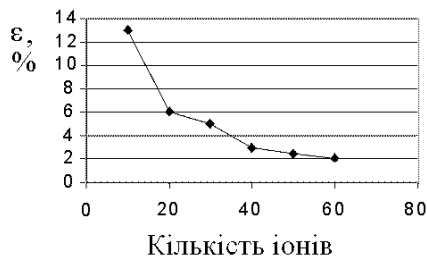


Рисунок 4 - Залежність відносної похибки від кількості випробувань

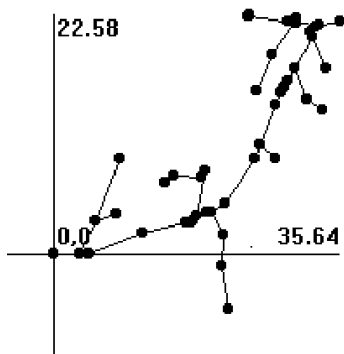


Рисунок 5 - Траєкторія руху іона фосфору з каскадами зіткнень в Si (енергія 2KeV)

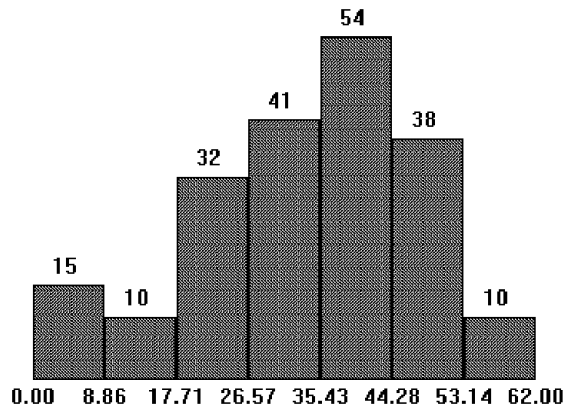


Рисунок 6 - Розподіл берилію в арсеніді галію

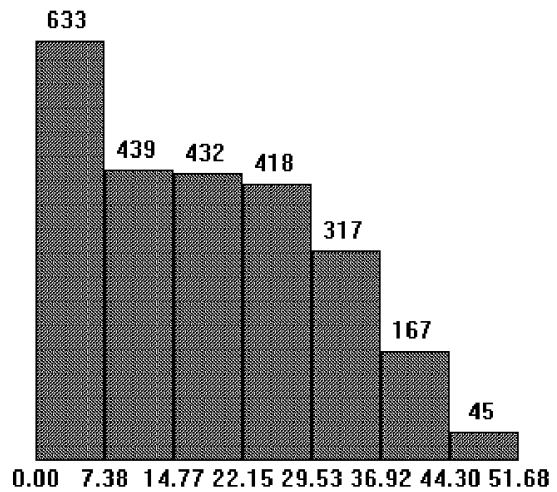


Рисунок 7 - Розподіл вибитих з вузлів атомів мішені (Дірки)

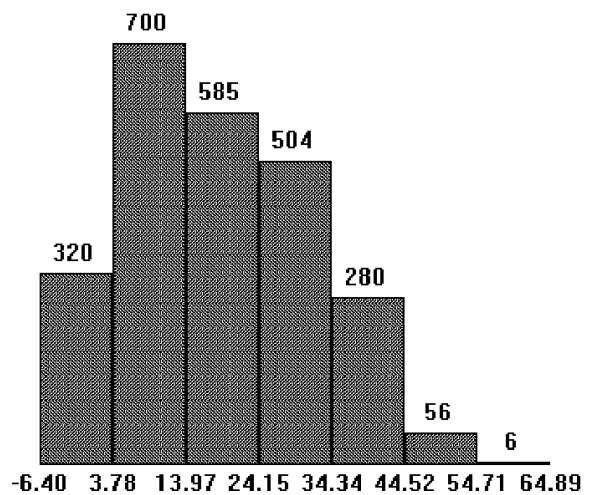


Рисунок 8 - Розподіл міжвузлових атомів мішені (Міжвузля)

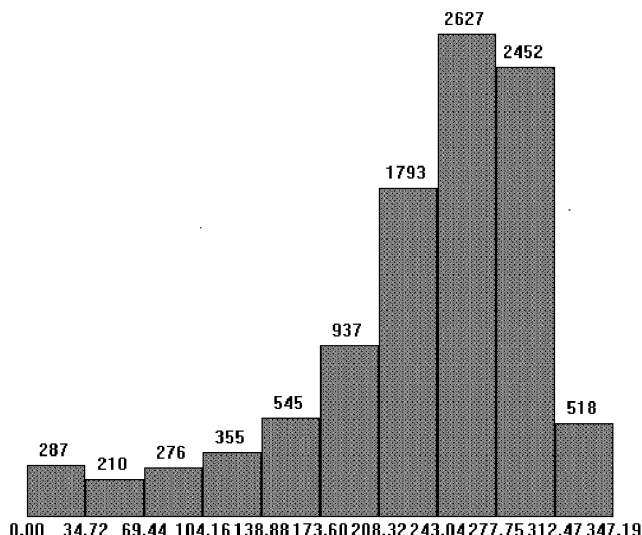


Рисунок 9 - Розподіл бору в кремнії ($E=50\text{KeV}$, кількість випробувань = 10000)

РЕЗУЛЬТАТИ

З допомогою запропонованої моделі були проведені розрахунки розподілів іонів фосфору, бору, арсену та сурьми в кремнії, берилію, цинку, селену, магнію, кад-

мію, вуглецю, телуру, свинцю та ін. в арсеніді галію, та ряду іонів домішок в підкладках типу GaP, InP, InAs, InSb, ZnSe, CdS, AsGa(1-x)Alx (рис. 5-9). Окрім того, розроблена модель дозволяє проводити впровадження іонів у багатокомпонентні та багат шарові мішені і похибка розрахунку, при цьому, не перевищує 5-10%.

ПЕРЕЛІК ПОСИЛАНЬ

1. Аваев Н.А., Наумов Ю.Е. Элементы сверхбольших интегральных схем. - М.: Радио и связь, 1986. - 168 с.
2. Соболев И.М. Метод Монте-Карло. - М.: Наука. Главная редакция физико-математической литературы, 1985. - 80 с.
3. Бубенников А.Н. Моделирование интегральных микротехнологий, приборов и схем. -М.: Высшая школа, 1989. - 320 с.
4. Dutton R.W., Yu Z. Technology CAD. Computer simulation of IC processes and devices. - Boston/Dordrecht/London, Kluwer Academic Publishers, 1993. - 373 p.
5. Росадо Р. Физическая электроника и микроэлектроника М.: Высш. шк., 1991. - 351 с.
6. Комаров Ф.Ф., Новиков А.П., Соловьев В.С., Ширяев С.Ю. Дефекты структуры в ионноимплантированном кремнии. - Мн.: Университетское, 1990. - 322 с.
7. Малви Т., Скотт В.Д. и др. Количественный электронно-зондовый микроанализ Пер. с англ. - М.: Мир, 1986. - 352 с.
8. Ермаков С.М. Метод Монте-Карло и смежные вопросы. -М.: Наука. Главная редакция физико-математической литературы, 1975. - 472 с.
9. Матаре Г. Электроника дефектов в полупроводниках. Пер.с англ.-М.: Мир, 1974. - 464 с.
10. Майер Дж., Эрикссон Л., Девис Дж. Ионное легирование полупроводников.(Кремний и германий). М.: Мир, 1973. - 296 с.
11. Фелдман Л., Майер Д. Основы анализа поверхности и тонких пленок: Пер. с англ. - М.: Мир, 1989. - 344 с.
12. Данилин Б.С., Киреев В.Ю. Ионное травление микроструктур. - М.: Сов. Радио, 1979. - 104 с.

УДК 621.314.12.001.5

ВРАХУВАННЯ ПАРАМЕТРІВ ДЖЕРЕЛА ТА НАВАНТАЖЕННЯ ПРИ РОЗРАХУНКАХ ХАРАКТЕРИСТИК ДВОПОЛЯРНИХ ІМПУЛЬСНИХ ПЕРЕТВОРЮВАЧІВ

В.Д.Флора

В радиоэлектронных устройствах как вторичные источники питания постоянным током применяют двухполярные импульсные преобразователи. Получены соотношения для расчета характеристик преобразователя с учетом параметров источника и нагрузки в относительных единицах. Приведен расчет характеристик преобразователя на примере двигателя, питающегося от аккумуляторной батареи.

У радиоэлектронных приборах как вторичные джерела живлення постійним струмом застосовують двополярні імпульсні перетворювачі. Одержані співвідношення для розрахунку характеристик перетворювача з урахуванням параметрів джерела та навантаження у відносних одиницях. Наведено розрахунок характеристик перетворювача на прикладі двигуна, який живиться від акумуляторної батареї.

The two-polar pulse converters are used in radioelectronic devices as the direct current secondary power supplies. The ratio for account of the converter characteristics are received in view of source parameters and loading relative units. The calculation of the converter characteristics as an example of the engine working from the storage battery is given.

В радіоелектронних пристроях як вторинні джерела живлення досить часто застосовуються двополярні імпульсні перетворювачі [1]. У загальному випадку з ідеальними прямокутними формами імпульсів часова діаграма напруги навантаження може бути подана рис.1.

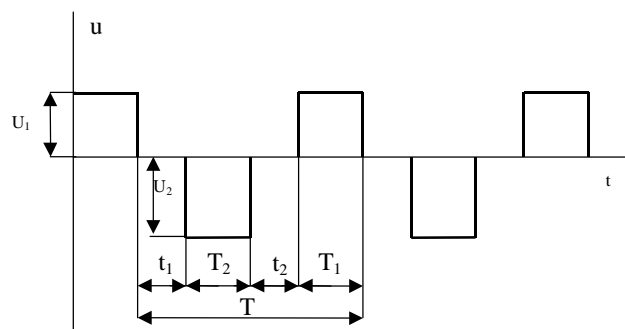


Рисунок 1 - Часова діаграма напруг двополярного імпульсного живлення навантаження

Відносні тривалості характерних проміжків часу позначимо

$$\left. \begin{aligned} \alpha_1 &= \frac{T_1}{T}; & \alpha_2 &= \frac{T_2}{T}; \\ \alpha_{01} &= \frac{t_1}{T}; & \alpha_{02} &= \frac{t_2}{T}, \end{aligned} \right\} \quad (1)$$

де T_1, T_2 - тривалості позитивного та негативного імпульсів напруги;

t_1, t_2 - тривалості пауз після позитивного та негативного імпульсів.

Період визначається, як

$$T = T_1 + T_2 + t_1 + t_2. \quad (2)$$

Якщо співвідношення (2) розділити на T , то

$$\alpha_1 + \alpha_2 + \alpha_{01} + \alpha_{02} = 1, \quad (3)$$

або, вводячи позначення

$$\lambda = \alpha_{01} + \alpha_{02}, \quad (4)$$

одержимо

$$\alpha_1 + \alpha_2 + \lambda = 1. \quad (5)$$

Звичайно $\lambda \neq 0$ при точному регулюванні середніх та діючих напруг навантаження, які становлять частки або одиниці відсотків від напруги джерела.

Для живлення електродвигунів слідкуючих електроприводів, застосовуваних у радіолокації, радіонавігації, радіоавтоматиці і т. і., звичайно такий широкий діапазон малих вихідних напруг перетворювача не потрібен, тим більше, що такий спосіб регулювання напруги дещо ускладнює систему керування перетворювачем. Тому в інженерній практиці в таких випадках застосовують спосіб регулювання, при якому $\lambda = 0$, тобто

$$\alpha_1 + \alpha_2 = 1. \quad (6)$$

У подальшому приділимо увагу саме цьому випадкові.

Використовуючи методику аналізу, викладену в [1], і метод рівності пульсацій, описаний в [2], та, враховуючи результати аналізу однополярного імпульсного перетворювача [3], одержимо для схеми з двома джерелами живлення [1]

$$U_{\text{ср}} = \frac{U_1 \alpha_1 - I_{\text{н ср}} \alpha_1^3 R_{\text{дж 1}}}{1 + (1 - \alpha_1) \frac{L_{\text{дж 1}}}{L_{\text{н}}}} - \frac{U_2 (1 - \alpha_1) - I_{\text{н ср}} (1 - \alpha_1)^3 R_{\text{дж 2}}}{1 + \alpha_1 \frac{L_{\text{дж 2}}}{L_{\text{н}}}} + I_{\text{н ср}} R_{\text{н}} (1 - 2\alpha_1), \quad (7)$$

де $U_{\text{ср}}, I_{\text{н ср}}$ - середні напруга та струм навантаження;

$R_{\text{дж 1}}, R_{\text{дж 2}}, L_{\text{дж 1}}, L_{\text{дж 2}}$ - відповідно внутрішні активні опори та індуктивності першого й другого джерел живлення;

U_1, U_2 - напруги холостого ходу джерел;

$L_{\text{н}}$ - індуктивність навантаження.

Розмах пульсацій струму навантаження для цього випадку

$$\Delta I_{\text{н}} =$$

$$= \frac{1}{L_{\text{н}} f} \left\{ \left[\frac{U_1 \alpha_1}{1 + (1 - \alpha_1) \frac{L_{\text{дж 1}}}{L_{\text{н}}}} - I_{\text{н ср}} \frac{\alpha_1^3 R_{\text{дж 1}}}{1 + (1 - \alpha_1) \frac{L_{\text{дж 1}}}{L_{\text{н}}}} \right] (1 - \alpha_1) + \left[\frac{U_2 (1 - \alpha_1)}{1 + \alpha_1 \frac{L_{\text{дж 2}}}{L_{\text{н}}}} - I_{\text{н ср}} \frac{(1 - \alpha_1)^3 R_{\text{дж 2}}}{1 + \alpha_1 \frac{L_{\text{дж 2}}}{L_{\text{н}}}} \right] \alpha_1 \right\}, \quad (8)$$

де частота комутації

$$f = \frac{1}{T}. \quad (9)$$

Потужність, споживана перетворювачем від двох джерел,

$$P_1 = P_{11} + P_{12}, \quad (10)$$

де P_{11}, P_{12} - відповідні потужності, споживані від першого та другого джерела.

Потужності P_{1i} можливо підрахувати за формулою

$$P_{1i} = U_{i \text{д вх}} I_{i \text{д дж}}, \quad (11)$$

де $U_{i \text{д вх}}, I_{i \text{д дж}}$ - відповідно діючі напруга та струм на вході перетворювача для i -го джерела.

Згідно з [1, 4],

$$U_{i \text{д вх}} = U_i, \quad (12)$$

$$I_{i \text{д дж}} = I_{\text{н ср}} \sqrt{\alpha_i}. \quad (13)$$

З урахуванням (12), (13) з формули (11):

$$P_{1i} = U_i I_{\text{н ср}} \sqrt{\alpha_i}. \quad (14)$$

Тоді зі співвідношень (10), (14):

$$P_1 = k_I U_1 I_{н\text{ ср}N} (\sqrt{\alpha_1} + k_U \sqrt{1 - \alpha_1}), \quad (15)$$

$$U_{д} = \sqrt{U_{д1}^2 + U_{д2}^2}, \quad (21)$$

де

$$k_I = \frac{I_{н\text{ ср}}}{I_{н\text{ ср}N}}; \quad (16)$$

$I_{н\text{ ср}N}$ - номінальний середній струм навантаження;

$$k_U = \frac{U_2}{U_1}. \quad (17)$$

Якщо базовою потужністю вважати

$$P_6 = U_1 I_{н\text{ ср}N}, \quad (18)$$

то зі співвідношень (15), (18) у відносних одиницях

$$p_{1*} = P_1/P_6 = k_I (\sqrt{\alpha_1} + k_U \sqrt{1 - \alpha_1}). \quad (19)$$

Потужність на виході перетворювача

$$P_2 = U_{д} I_{нд}, \quad (20)$$

де $U_{д}$, $I_{нд}$ - діючі напруга та струм навантаження.

Введемо позначення:

$$\begin{aligned} u_{*1} &= \frac{U_{ср1}}{U_1}; \quad u_{*2} = \frac{U_{ср2}}{U_2}; \quad R_{не1} = \frac{U_1}{I_{н\text{ ср}N}}; \\ X_{LнN} &= 2\pi f L_{нN}; \quad K_L = \frac{L_{н}}{L_{нN}}; \quad X_{Lн} = X_{LнN} K_L; \\ m_{Lн2} &= \frac{R_{не2}}{X_{Lн}}; \quad m_{Lн1} = \frac{R_{не1}}{X_{Lн}}; \quad R_{не1} = \frac{U_1}{I_{н\text{ ср}N}}; \\ R_{не2} &= \frac{U_2}{I_{н\text{ ср}N}}; \quad m_{R_{31}} = \frac{R_{дж1}}{R_{н}}; \quad m_{R_{32}} = \frac{R_{дж2}}{R_{н}}; \\ L_{дж1*} &= \frac{L_{дж1}}{L_{н}}; \quad L_{дж2*} = \frac{L_{дж2}}{L_{н}}; \\ \Delta u_{*1N} &= \frac{I_{н\text{ ср}N} R_{дж1}}{U_1}; \quad \Delta u_{*2N} = \frac{I_{н\text{ ср}N} R_{дж2}}{U_2}; \\ \Delta u_{*1} &= k_I \Delta u_{*1N}; \quad \Delta u_{*2} = k_I \Delta u_{*2N}; \\ I_{61} &= \frac{U_1}{X_{LнN}}; \quad I_{62} = \frac{U_2}{X_{LнN}}; \end{aligned}$$

$U_{ср1}$, $U_{ср2}$ - середні напруги позитивного та негативного імпульсів;

$X_{LнN}$ - номінальний індуктивний опір навантаження для основної гармоніки змінної складової струму.

Діюча напруга навантаження:

де $U_{д1}$, $U_{д2}$ - діючі напруги позитивних та негативних імпульсів.

Користуючись методикою [3, 4], одержимо

$$U_{д} = U_1 \{u_{*1}^2 + \alpha_1(1 - \alpha_1) + k_U^2 [u_{*2}^2 + \alpha_1(1 - \alpha_1)]\}^{0,5}, \quad (22)$$

де

$$u_{*1} = \alpha_1 - k_I \Delta u_{*1N} (1 + \alpha_1^3 m_{R_{31}}); \quad (23)$$

$$u_{*2} = 1 - \alpha_1 - k_I \Delta u_{*2N} [1 + (1 - \alpha_1)^3 m_{R_{32}}]. \quad (24)$$

Діючий струм навантаження [4]:

$$I_{нд} = k_I I_{н\text{ ср}N} \left(1 + \frac{\Delta I_{н}^2}{12 k_I^2 I_{н\text{ ср}N}^2} \right)^{0,5}. \quad (25)$$

З формули (25), користуючись співвідношенням (8) та введеними позначеннями, після перетворень

$$I_{нд} = k_I I_{н\text{ ср}N} \left(1 + \frac{m_{Lн1}^2 k_L^2 \Delta i_{*}^2}{12 k_I^2} \right)^{0,5}, \quad (26)$$

де розмах пульсацій струму у відносних одиницях

$$\begin{aligned} \Delta i_{н*} &= \frac{\Delta I_{н}}{I_{61}} = \\ &= \frac{2\pi}{K_L} \alpha_1 (1 - \alpha_1) \times \\ &\times \left\{ \frac{1 - k_I \Delta u_{*1N} \alpha_1^2}{1 + (1 - \alpha_1) L_{дж1*}} + \frac{k_U [1 - k_I \Delta u_{*2N} (1 - \alpha_1)^2]}{1 + \alpha_1 L_{дж2}} \right\}. \quad (27) \end{aligned}$$

Підставляючи з формул (22), (26) у (20), після перетворень

$$\begin{aligned} P_2 &= k_I U_1 I_{н\text{ ср}N} \left\{ [u_{*1}^2 + \alpha_1(1 - \alpha_1) + \right. \\ &\left. + k_U^2 (u_{*2}^2 + \alpha_1(1 - \alpha_1))] \times \left(1 + \frac{m_{Lн1}^2 K_L^2 \Delta i_{н*}^2}{12 k_I^2} \right) \right\}^{0,5}. \quad (28) \end{aligned}$$

Або, після перетворень, у відносних одиницях

$$\begin{aligned} p_{2*} &= k_I \left\{ [u_{*1}^2 + \alpha_1(1 - \alpha_1) + \right. \\ &\left. + k_U^2 (u_{*2}^2 + \alpha_1(1 - \alpha_1))] \times \left(1 + \frac{m_{Lн1}^2 K_L^2 \Delta i_{н*}^2}{12 k_I^2} \right) \right\}^{0,5}. \quad (29) \end{aligned}$$

ККД перетворювача дорівнює:

$$\eta = \frac{P_2}{P_1} = \frac{P_{2*}}{P_{1*}} = \frac{\left[u_{*1}^2 + \alpha_1(1-\alpha_1) + k_I^2(u_{*2}^2 + \alpha_1(1-\alpha_1)) \right] \times \left(1 + \frac{m_{LH1}^2 K_L^2}{12k_I^2} \Delta i_{H*} \right)^{0,5}}{\sqrt{\alpha_1} + \sqrt{1-\alpha_1}}. \quad (30)$$

Розглянемо вплив параметрів джерела та навантаження на характеристики імпульсного перетворювача для більш простого випадку, коли навантаження (якір двигуна постійного струму) живиться від однієї акумуляторної батареї. В такому разі можливо вважати $U_1 = U_2 = U$, $L_{дж1} = L_{дж2} \cong 0$, $R_{дж1} = R_{дж2} = R_{дж}$.

Тоді з формул (7) та (8):

$$U_{ср} = U(2\alpha_1 - 1) - I_{н ср} R_{дж} [\alpha_1^3 - (1 - \alpha_1)^3] + I_{н ср} R_H (1 - 2\alpha); \quad (31)$$

$$\Delta I_H = \frac{\alpha_1(1 - \alpha_1)[2U - I_{н ср} R_{дж}(2\alpha_1 - 1)]}{L_H f}. \quad (32)$$

Тоді введені позначення змінюються таким чином:

$$k_U = 1; R_{не1} = R_{не2} = R_{не} = \frac{U}{I_{н ср} N};$$

$$m_{LH1} = m_{LH2} = m_{LH} = \frac{R_{не}}{X_{LHN}};$$

$$m_{R31} = m_{R32} = m_{R3} = \frac{R_{дж}}{R_H};$$

$$\Delta u_{*1N} = \Delta u_{*2N} = \Delta u_{*N} = \frac{I_{н ср} N R_{дж}}{U};$$

$$\Delta u_{*1} = \Delta u_{*2} = k_I \Delta u_{*N}.$$

З формул (23), (24):

$$u_{*1} = \alpha_1 - k_I \Delta u_{*N} (1 + \alpha_1^3 m_{R3}); \quad (33)$$

$$u_{*2} = 1 - \alpha_1 - k_I \Delta u_{*N} [1 + (1 - \alpha_1)^3 m_{R3}]. \quad (34)$$

Крім того, для спрощення розрахунків вважаємо $K_L = 1$.

В такому разі зовнішня та регульовальна характеристики перетворювача у відносних одиницях подаються формулою

$$u_* = 2\alpha_1 - 1 - k_I \Delta u_{*N} \{ [\alpha_1^3 - (1 - \alpha_1)^3] - 1 + 2\alpha_1 \}. \quad (35)$$

З формули (19):

$$P_{1*} = k_I (\sqrt{\alpha_1} + \sqrt{1 - \alpha_1}). \quad (36)$$

Потужність P_{2*} з формули (29):

$$P_{2*} = k_I \left\{ [u_{*1}^2 + 2\alpha_1(1 - \alpha_1) + u_{*2}^2] \times \left(1 + \frac{m_{LH1}^2}{12k_I^2} \Delta i_{H*} \right)^{0,5} \right\}. \quad (37)$$

Формула для ККД зі співвідношення (30):

$$\eta = \frac{\left\{ [u_{*1}^2 + 2\alpha_1(1 - \alpha_1) + u_{*2}^2] \times \left(1 + \frac{m_{LH1}^2}{12k_I^2} \Delta i_{H*} \right)^{0,5} \right\}}{\sqrt{\alpha_1} + \sqrt{1 - \alpha_1}}, \quad (38)$$

де з виразу (27):

$$\Delta i_{H*} = 2\pi\alpha_1(1 - \alpha_1) \{ 2 - k_I \Delta u_{*N} [\alpha_1^2 - (1 - \alpha_1)^2] \}. \quad (39)$$

В практичних розрахунках найчастіше

$$1 + \frac{m_{LH1}^2}{12k_I^2} \Delta i_{H*} \cong 1.$$

Тому зі співвідношення (38) можна одержати наближену формулу

$$\eta = \frac{[u_{*1}^2 + 2\alpha_1(1 - \alpha_1) + u_{*2}^2]^{0,5}}{\sqrt{\alpha_1} + \sqrt{1 - \alpha_1}} = \frac{u_{*д}}{\sqrt{\alpha_1} + \sqrt{1 - \alpha_1}}, \quad (40)$$

де відносна діюча напруга навантаження з формули (22):

$$u_{*д} = \frac{U_{д}}{U} = [u_{*1}^2 + 2\alpha_1(1 - \alpha_1) + u_{*2}^2]^{0,5}. \quad (41)$$

Проаналізуємо вплив параметрів на характеристики перетворювача (35), (40) на прикладі імпульсного живлення двигуна постійного струму типу МТ-2 від акумуляторної батареї типу ТЖН-250. Основні дані джерела та навантаження, визначені за паспортними даними та експериментально, такі: $U = 32$ В; $I_{н ср} N = 80$ А;

$$L_{HN} = 1,22 \cdot 10^{-3} \text{ Гн}; R_{HN} = 0,043 \text{ Ом}; R_{дж} = 0,064 \text{ Ом};$$

$L_{дж} \cong 0$ при розрахунковій температурі $\theta = 75^\circ \text{C}$.

Комутатор побудовано на польових транзисторах MOSFET, а частота комутації $f = 4000$ Гц.

Зовнішні та регульовальні характеристики розраховані за формулою (35). Результати розрахунків наведено у табл.1.

Середня напруга навантаження, таким чином, змінюється в процесі регулювання за величиною та знаком. Діюча напруга, розрахована за формулою (41), змінюється за даними табл.2.

Таблиця 1 - Результати розрахунків зовнішніх та регулювальних характеристик $u_*(k_f, \alpha_1)$

| k_f α_1 | 0 | 0,2 | 0,4 | 0,6 | 0,8 | 1,0 |
|---------------------|------|--------|--------|--------|--------|--------|
| 0 | -1 | -0,947 | -0,893 | -0,84 | -0,786 | -0,733 |
| 0,1 | -0,8 | -0,759 | -0,719 | -0,678 | -0,638 | -0,597 |
| 0,2 | -0,6 | -0,571 | -0,542 | -0,513 | -0,484 | -0,455 |
| 0,3 | -0,4 | -0,381 | -0,363 | -0,344 | -0,325 | -0,306 |
| 0,4 | -0,2 | -0,191 | -0,182 | -0,173 | -0,163 | -0,154 |
| 0,5 | 0 | 0 | 0 | 0 | 0 | 0 |
| 0,6 | 0,2 | 0,191 | 0,182 | 0,173 | 0,163 | 0,154 |
| 0,7 | 0,4 | 0,381 | 0,363 | 0,344 | 0,325 | 0,306 |
| 0,8 | 0,6 | 0,571 | 0,542 | 0,513 | 0,484 | 0,455 |
| 0,9 | 0,8 | 0,759 | 0,719 | 0,678 | 0,638 | 0,597 |
| 1,0 | 1,0 | 0,947 | 0,893 | 0,84 | 0,786 | 0,733 |

Таблиця 2 - Залежність $u_{\partial*}(k_f, \alpha_1)$

| k_f α_1 | 0 | 0,2 | 0,4 | 0,6 | 0,8 | 1,0 |
|---------------------|-----|-------|-------|-------|-------|-------|
| 0 | 1,0 | 0,947 | 0,894 | 0,842 | 0,791 | 0,74 |
| 0,1 | 1,0 | 0,958 | 0,917 | 0,876 | 0,836 | 0,798 |
| 0,2 | 1,0 | 0,965 | 0,931 | 0,899 | 0,867 | 0,836 |
| 0,3 | 1,0 | 0,971 | 0,888 | 0,837 | 0,789 | 0,744 |
| 0,4 | 1,0 | 0,948 | 0,9 | 0,853 | 0,81 | 0,769 |
| 0,6 | 1,0 | 0,948 | 0,9 | 0,853 | 0,81 | 0,769 |
| 0,7 | 1,0 | 0,971 | 0,888 | 0,837 | 0,789 | 0,744 |
| 0,8 | 1,0 | 0,965 | 0,931 | 0,899 | 0,867 | 0,836 |
| 0,9 | 1,0 | 0,958 | 0,917 | 0,876 | 0,836 | 0,798 |
| 1,0 | 1,0 | 0,947 | 0,894 | 0,842 | 0,791 | 0,74 |

При $\alpha_1 = 0,5$ через навантаження проходить чисто змінний струм. Тому в даному випадку $I_{н ср} = 0$, $k_f = 0$, $u_{\partial*} = 1$.

Залежність $\eta(k_f, \alpha_1)$ наведено у табл.3.

При $\alpha_1 = 0,5$ виникають лише втрати потужності на активному опорі навантаження. Тому в цьому випадку $k_f = 0$; $\eta = 0,707$.

З наведених досліджень можливо зробити такі висновки.

1. Одержані загальні співвідношення у відносних одиницях для розрахунку напруг, струмів, потужностей та ККД в залежності від коефіцієнта α_1 та відносного середнього стуму k_f з урахуванням параметрів джерела та навантаження.

Таблиця 3 - Результати розрахунків залежності $\eta(k_f, \alpha_1)$

| k_f α_1 | 0 | 0,2 | 0,4 | 0,6 | 0,8 | 1,0 |
|---------------------|-------|-------|-------|-------|-------|-------|
| 0 | 1,0 | 0,947 | 0,894 | 0,842 | 0,791 | 0,74 |
| 0,1 | 0,79 | 0,757 | 0,725 | 0,693 | 0,661 | 0,631 |
| 0,2 | 0,745 | 0,719 | 0,694 | 0,67 | 0,646 | 0,623 |
| 0,3 | 0,722 | 0,701 | 0,642 | 0,605 | 0,57 | 0,538 |
| 0,4 | 0,711 | 0,674 | 0,639 | 0,607 | 0,576 | 0,547 |
| 0,6 | 0,711 | 0,674 | 0,639 | 0,607 | 0,576 | 0,547 |
| 0,7 | 0,722 | 0,701 | 0,642 | 0,605 | 0,57 | 0,538 |
| 0,8 | 0,745 | 0,719 | 0,694 | 0,67 | 0,646 | 0,623 |
| 0,9 | 0,79 | 0,757 | 0,725 | 0,693 | 0,661 | 0,631 |
| 1,0 | 1,0 | 0,947 | 0,894 | 0,842 | 0,791 | 0,74 |

2. Середня відносна напруга навантаження змінюється в межах від "-1" до "+1" із змінням величини та знаку.

3. Діюча напруга навантаження в процесі регулювання середньої напруги змінюється в межах $(74 \div 100)\%$ зі змінням α_1 та k_f .

4. При $\alpha_1 = 0,5$ навантаження перебуває під дією змінною напругою, яка дорівнює напрузі джерела. При цьому $k_f = 0$.

5. При $\alpha_1 = 0,5$ від джерела споживається потужність втрат на активному опорі навантаження від змінного несинусоїдального струму.

6. Залежність $\eta(\alpha_1)$ однакова при фіксованому k_f для $0,5 \leq \alpha_1 \leq 1$ та $0 \leq \alpha_1 \leq 0,5$.

ПЕРЕЛІК ПОСИЛАНЬ

1. Флора В.Д. Дослідження двополярних імпульсних регуляторів. // Радіоелектроніка, інформатика, управління. - 1999. - №2. - С.48-50.
2. Флора В.Д. Метод рівності розмахів пульсацій періодично змінних величин в інтервалах часу накопичення та витрати енергії. // Електротехніка та електроенергетика. - 2000. - №2. - С.45-46.
3. Флора В.Д. Вплив параметрів джерела та навантаження на характеристики імпульсного перетворювача для зниження напруги. // Радіоелектроніка, інформатика, управління. 2000. - №2. - С. 26-30.
4. Флора В.Д. Апроксимація графіка не синусоїдної функції за допомогою найпростіших геометричних фігур та елементарних функцій. // Електротехніка та електроенергетика. - 1999. - №1. - С. 24-25.

ОПЕРАТОРИ ІНТЕРВАЛЬНОЇ АРИФМЕТИКИ ДЛЯ РОЗРАХУНКУ ДОПУСКІВ

Г.М.Шило

Введени поняття полудопускового, полууправляемого и полного допускового множества решений. Для их определения предложены операторы прямых и обратных арифметических операций.

Введено понятия напівдопускової, напівкерованої та повної допускової множини розв'язків. Для їх визначення запропоновані оператори прямих і зворотних арифметичних операцій.

Notations of semitolerable, semicontrollable and full tolerable solution sets are introduced. To determine them the operator of direct and inverse arithmetical operation are suggested.

ВСТУП

В умовах конкуренції актуальною задачею є підвищення якості продукції. Одним із напрямків розв'язання цієї задачі є визначення оптимальних номінальних значень параметрів елементів і їх допустимих відхилень, що забезпечує зменшення собівартості продукції і гарантує її якість і надійність при зовнішніх впливах. Зручним математичним апаратом для синтезу і аналізу допусків є інтервальна математика [1-5].

У інтервальної математиці використовуються правильні інтервали

$$\mathbf{a} = [a; \bar{a}], \quad a \leq \bar{a}$$

і неправильні інтервали [3,5]

$$\mathbf{a} = [a; \bar{a}], \quad a > \bar{a},$$

де a і \bar{a} - нижня і верхня межа інтервалу \mathbf{a} .

Арифметичні операції в найбільш повній інтервальній арифметиці, відомої як арифметика Каухера, визначаються за допомогою співвідношення:

$$\mathbf{x} * \mathbf{y} = \bigcup_{x \in \text{рго } \mathbf{x}} \bigcup_{y \in \text{рго } \mathbf{y}} (\Omega^x * y), \quad (1)$$

де $*$ $\in \{+, -, \cdot, / \}$; Ω^x і Ω^y - умовні решітчасті операції взяття екстремуму по включенню щодо інтервалів \mathbf{x} і \mathbf{y} ; $\text{рго } \mathbf{x}$ і $\text{рго } \mathbf{y}$ - правильні проєкції інтервалів \mathbf{x} і \mathbf{y} .

Оператори Ω і рго визначаються за допомогою виразів:

$$\Omega^x: = \begin{cases} \vee, & \mathbf{x} \text{ - правильний;} \\ \wedge, & \text{в інших випадках,} \end{cases}$$

$$\text{рго } \mathbf{x}: = \begin{cases} \mathbf{x}, & \mathbf{x} \text{ - правильний;} \\ \text{dual } \mathbf{x}, & \text{в інших випадках,} \end{cases}$$

де $\text{dual } \mathbf{x} = [\bar{x}; x]$ - інтервал дуальний інтервалу $\mathbf{x} = [x; \bar{x}]$.

Арифметика Каухера успішно виконує свої функції при аналізі рівнянь:

$$\mathbf{a} \cdot \mathbf{x} = \mathbf{b}, \quad (0 \notin \mathbf{a}; 0 \notin \mathbf{b}). \quad (2)$$

У цьому випадку для рівняння (2) можуть бути визначені три основних види розв'язків [5]:

- об'єднана множина розв'язків

$$\mathbf{X}_u(\mathbf{a}, \mathbf{b}) = \{x \in \mathbf{R} | (\exists a \in \mathbf{a})(\exists b \in \mathbf{b})(a \cdot x = b)\}; \quad (3)$$

- допускова множина розв'язків

$$\mathbf{X}_t(\mathbf{a}, \mathbf{b}) = \{x \in \mathbf{R} | (\forall a \in \mathbf{a})(\exists b \in \mathbf{b})(a \cdot x = b)\}; \quad (4)$$

- керована множина розв'язків

$$\mathbf{X}_c(\mathbf{a}, \mathbf{b}) = \{x \in \mathbf{R} | (\forall b \in \mathbf{b})(\exists a \in \mathbf{a})(a \cdot x = b)\}. \quad (5)$$

Розв'язкам (3) і (5) рівняння (2) відповідають неправильні інтервали, а розв'язку (4) - правильні.

Ситуація змінюється для нульмістивних інтервалів. Арифметика Каухера припускає нетривіальний дільник нуля при множенні правильних і неправильних інтервалів, наприклад, $[-1; 3][2; -1] = 0$. Це означає, що розв'язок (3) у цьому випадку може бути записане у виді:

$$\mathbf{X}_u(\mathbf{a}, \mathbf{b}) = \{x \in \mathbf{R} | (\exists a=0)(\exists b=0)(a \cdot x = 0)\},$$

а керована множина розв'язків (5) узагалі не може бути визначено. Графічний розв'язок рівняння (2) показує, що розв'язок (5) існує і для нульмістивних інтервалів (рис.1). Керована множина розв'язків у цьому випадку складається з двох областей:

$$\mathbf{X}_c(\mathbf{a}, \mathbf{b}) = \mathbf{X}'_c(\mathbf{a}, \mathbf{b}) \cup \mathbf{X}''_c(\mathbf{a}, \mathbf{b}). \quad (6)$$

ЗАУВАЖЕННЯ. Цікавою особливістю рівнянь з нульмістивними інтервалами є наявність напівдопускової множини

$$\mathbf{X}_{st}(\mathbf{a}_{st}, \mathbf{b}_{st}) = \{x \in \mathbf{R} | (\forall a \in \mathbf{a}_{st})(\exists b \in \mathbf{b}_{st})(a \cdot x = b)\}; \quad (7)$$

і напівкерованої множини розв'язків

$$\mathbf{X}_{sc}(\mathbf{a}_{sc}, \mathbf{b}_{sc}) = \{x \in \mathbf{R} | (\forall b \in \mathbf{b}_{sc})(\exists a \in \mathbf{a}_{sc})(a \cdot x = b)\}. \quad (8)$$

Для рис.1 $\mathbf{a}_{st} = [a; 0]$, $\mathbf{b}_{st} = [0; \bar{b}]$, $\mathbf{a}_{sc} = [0; \bar{a}]$, $\mathbf{b}_{sc} = [b; 0]$. Межа повної допускової множини розв'язків

$$X_{ft}(\mathbf{a}, \mathbf{b}) = X_t(\mathbf{a}, \mathbf{b}) \cup X_{st}(\mathbf{a}_{st}, \mathbf{b}_{st}). \quad (9)$$

очевидно, є внутрішніми межами керованої множини розв'язків (6).

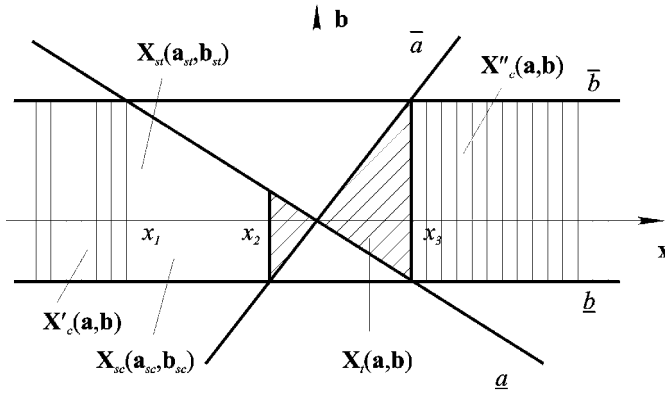


Рисунок 1 - Види множин розв'язків

У даній роботі арифметика Каухера доповнюється можливістю множення правильних і неправильних нульмістивних інтервалів із метою визначення меж повної допускової множини розв'язків (9). Для цього виявилось необхідним видозмінити оператори інтервальної арифметики.

1. ОПЕРАТОРИ ПРЯМИХ ІНТЕРВАЛЬНИХ АРИФМЕТИЧНИХ ОПЕРАЦІЙ

Будемо вважати, що межі результуючого інтервалу арифметичної операції вибираються з множини сполучень результатів арифметичних операцій над межами інтервалів. Вибір елементів множини здійснюється за допомогою подвійного застосування операторів пошуку екстремальної підмножини. При кожному пошуку упорядкована множина розділяється на дві підмножини з рівною кількістю елементів.

ОЗНАЧЕННЯ. Елементи упорядкованої множини $Q = \{q_1, q_2, q_3, q_4\}$, $q_1 < q_2 < q_3 < q_4$ визначаються за допомогою співвідношень:

$$q_1 = \min \min Q = \text{Min} Q; \quad (10)$$

$$q_2 = \max \min Q; \quad (11)$$

$$q_3 = \min \max Q; \quad (12)$$

$$q_4 = \max \max Q = \text{Max} Q. \quad (13)$$

ТЕОРЕМА 1 (пряма операторна). Результуючий інтервал прямої арифметичної операції визначається виразами:

$$\mathbf{b} = \mathbf{a} * \mathbf{x} = \begin{cases} [\text{Min} \mathbf{F}; \text{Max} \mathbf{F}], (\mathbf{a}, \mathbf{x} - \text{правильні}) & (14) \\ [\max \min \mathbf{F}; \min \max \mathbf{F}], (\mathbf{a} \text{ або } \mathbf{x} - \text{правильні,} & (15) \\ & \mathbf{b} - \text{правильний)} \\ [\min \max \mathbf{F}; \max \min \mathbf{F}], (\mathbf{a} \text{ або } \mathbf{x} - \text{неправильні,} & (16) \\ & \mathbf{b} - \text{неправильний)} \\ [\text{Max} \mathbf{F}; \text{Min} \mathbf{F}], (\mathbf{a}, \mathbf{x} - \text{неправильні}) & (17) \end{cases}$$

де $* \in \{+, -, \cdot, /\}$; $\mathbf{F} = \{\underline{a} * \underline{x}; \underline{a} * \bar{x}; \bar{a} * \underline{x}; \bar{a} * \bar{x}\}$.

Доказ. Результуючий інтервал при операції додавання визначається виразами:

$$\mathbf{b} = \mathbf{a} + \mathbf{x} = [\underline{a} + \underline{x}; \bar{a} + \bar{x}]. \quad (18)$$

Ранжування елементів множини \mathbf{F} призводить до таких послідовностей:

$$\underline{a} + \underline{x} < \underline{a} + \bar{x}, \bar{a} + \underline{x} < \bar{a} + \bar{x}, (\mathbf{a}, \mathbf{x} - \text{правильні}) \quad (19)$$

$$\underline{a} + \bar{x} < \underline{a} + \underline{x} < \bar{a} + \bar{x} < \bar{a} + \underline{x}, (\mathbf{a} - \text{правильний,} & (20) \\ & \mathbf{x} - \text{неправильний,} \\ & \mathbf{b} - \text{правильний)}$$

$$\underline{a} + \bar{x} < \bar{a} + \bar{x} < \underline{a} + \underline{x} < \bar{a} + \underline{x}, (\mathbf{a} - \text{правильний,} & (21) \\ & \mathbf{x} - \text{неправильний,} \\ & \mathbf{b} - \text{неправильний)}$$

$$\bar{a} + \bar{x} < \underline{a} + \bar{x}, \bar{a} + \underline{x} < \underline{a} + \underline{x}, (\mathbf{a}, \mathbf{x} - \text{неправильні}). \quad (22)$$

Вибір елементів послідовностей (18)-(22) за допомогою співвідношень (14)-(17) завжди призводить до виразу (18). Аналогічно доводиться придатність співвідношень (15) і (16) для неправильних \mathbf{a} і правильних \mathbf{x} інтервалів.

Результуючий інтервал при множенні визначається виразами:

$$\mathbf{b} = \mathbf{a} \cdot \mathbf{x} = [\underline{a} \cdot \underline{x}; \bar{a} \cdot \bar{x}], (\underline{a}, \bar{a}, \underline{x}, \bar{x}) > 0. \quad (23)$$

Ранжування елементів множини для цього випадку множення інтервалів призводить до послідовностей:

$$\underline{a} \cdot \underline{x} < \underline{a} \cdot \bar{x}, \bar{a} \cdot \underline{x} < \bar{a} \cdot \bar{x}, (\mathbf{a}, \mathbf{x} - \text{правильні}) \quad (24)$$

$$\underline{a} \cdot \bar{x} < \underline{a} \cdot \underline{x} < \bar{a} \cdot \bar{x} < \bar{a} \cdot \underline{x}, (\mathbf{a} - \text{правильний,} & (25) \\ & \mathbf{x} - \text{неправильний,} \\ & \mathbf{b} - \text{правильний)}$$

$$\underline{a} \cdot \bar{x} < \bar{a} \cdot \bar{x} < \underline{a} \cdot \underline{x} < \bar{a} \cdot \underline{x}, (\mathbf{a} - \text{правильний,} & (26) \\ & \mathbf{x} - \text{неправильний,} \\ & \mathbf{b} - \text{неправильний)}$$

$$\bar{a} \cdot \bar{x} < \underline{a} \cdot \bar{x}, \bar{a} \cdot \underline{x} < \underline{a} \cdot \underline{x}, (\mathbf{a}, \mathbf{x} - \text{неправильні}). \quad (27)$$

Вибір елементів послідовностей (24)-(27) за допомогою співвідношень (14)-(17) призводить до виразу (23). Аналогічно доводиться придатність операторів (14)-(27) для інших сполучень нульмістивних інтервалів.

Вибір операторів при множенні нульмістивних інтервалів розглянемо за допомогою графічного розв'язку на рис.1. Множенню правильних інтервалів відповідає випадок $\underline{x} = x_2$, $\bar{x} = x_3$. Якщо виконуються умови:

$$\underline{a} \cdot \bar{x} < \bar{a} \cdot \underline{x}, \underline{a} \cdot \underline{x} < \bar{a} \cdot \bar{x},$$

то множення призводить до інтервалу

$$\mathbf{b} = \mathbf{a} \cdot \mathbf{x} = [\underline{a} \cdot \bar{x}; \bar{a} \cdot \bar{x}]. \quad (28)$$

Ранжована послідовність

$$\underline{a} \cdot \bar{x} < \bar{a} \cdot \underline{x} < \underline{a} \cdot \underline{x} < \bar{a} \cdot \bar{x}$$

відповідає застосуванню операторів (14). Інші випадки множення правильних нульмістивних інтервалів також призводять до використання цих операторів.

Множення правильного \mathbf{a} і неправильного \mathbf{x} інтервалів відповідають межах $\underline{x} = x_3$, $\bar{x} = x_1$. Якщо виконуються умови:

$$\mathbf{a} \cdot \bar{x} > \bar{\mathbf{a}} \cdot \underline{x}, \quad \mathbf{a} \cdot \underline{x} > \bar{\mathbf{a}} \cdot \bar{x},$$

то множення призводить до інтервалу:

$$\mathbf{b} = \mathbf{a} \cdot \mathbf{x} = [\underline{\mathbf{a}} \cdot \underline{x}; \bar{\mathbf{a}} \cdot \bar{x}]. \quad (29)$$

Ранжована послідовність

$$\bar{\mathbf{a}} \cdot \bar{x} < \underline{\mathbf{a}} \cdot \underline{x} < \bar{\mathbf{a}} \cdot \underline{x} < \underline{\mathbf{a}} \cdot \bar{x}$$

відповідає застосуванню операторів (15). Аналогічно доводиться застосування операторів (15)-(17) для інших співвідношень між елементами множини \mathbf{F} й інших сполучень правильних і неправильних інтервалів. Оскільки операції віднімання і ділення можуть бути перетворені до операцій додавання і множення, то оператори (14)-(17) можуть використовуватися й у цих випадках.

2. ОПЕРАТОРИ ЗВОРОТНИХ ІНТЕРВАЛЬНИХ АРИФМЕТИЧНИХ ОПЕРАЦІЙ

У інтервальній арифметиці звичайні для дійсних чисел відповідності між прямими й зворотними арифметичними операціями набувають інтервального характеру, що призводить до подвоєння числа арифметичних операцій у інтервальній арифметиці (табл.1).

Таблиця 1 - Відповідність між прямими й зворотними операціями в інтервальній арифметиці

| Операція | Позначення | | | |
|----------|------------|---|---|---|
| Пряма | + | - | . | / |
| Зворотна | ⊖ | ⊕ | ⊘ | ⊙ |

ПРОПОЗИЦІЯ. Будемо вважати, що результат прямої арифметичної операції існує. Це дозволяє здійснювати зворотне ділення двох нульмістивних інтервалів. Ця процедура може розглядатися як розкриття невизначеності виду 0/0.

Очевидно, що результуючий інтервал зворотної арифметичної операції також визначається за допомогою операторів (10)-(13). Умови їхнього застосування визначаються теоремою 2.

ТЕОРЕМА 2 (зворотна операторна). Результуючий інтервал зворотної арифметичної операції визначається виразами:

$$\mathbf{x} = \mathbf{b}^{**} \mathbf{a} = \begin{cases} [\text{Min } \mathbf{G}; \text{Max } \mathbf{G}], & (\mathbf{b} \text{ - правильний, } \mathbf{a} \text{ - неправильний}) & (30) \\ [\max \min \mathbf{G}; \min \max \mathbf{G}], & (\mathbf{b}, \mathbf{a} \text{ - правильні, або неправильні, } \mathbf{x} \text{ - правильний}) & (31) \\ [\min \max \mathbf{G}; \max \min \mathbf{G}], & (\mathbf{b}, \mathbf{a} \text{ - правильні, або неправильні, } \mathbf{x} \text{ - неправильний}) & (32) \\ [\text{Max } \mathbf{G}; \text{Min } \mathbf{G}], & (\mathbf{b} \text{ - неправильний, } \mathbf{a} \text{ - правильний}) & (33) \end{cases}$$

крім випадків

$$\mathbf{x} = \mathbf{b} \oslash \mathbf{a} = \begin{cases} [\text{Max } \mathbf{G}; \text{Min } \mathbf{G}] & (\mathbf{b}, \mathbf{a} \text{ - правильні; } 0 \in \mathbf{b}, \mathbf{a}; \mathbf{x} \text{ - неправильний}) & (34) \\ [\text{Min } \mathbf{G}; \text{Max } \mathbf{G}] & (\mathbf{b}, \mathbf{a} \text{ - неправильні; } 0 \in \mathbf{b}, \mathbf{a}; \mathbf{x} \text{ - правильний}) & (35) \end{cases}$$

де $** \in \{\oplus, \ominus, \odot, \oslash\}$; $\mathbf{G} = \{\underline{b}^* \underline{a}; \underline{b}^* \bar{a}; \bar{b}^* \underline{a}; \bar{b}^* \bar{a}\}$; $* \in \{+, -, \cdot, /\}$.

Доказ. Зворотні арифметичні операції можуть бути перетворені до прямих шляхом застосування оператора dual до правого операнду [5]:

$$\mathbf{x} = \mathbf{b}^{**} \mathbf{a} = \mathbf{b}^* \text{dual}(\mathbf{a}). \quad (36)$$

Співвідношення (36) перетворить умови виразів (30)-(33) до умов (14)-(17).

Вираз (34) відповідає пошуку внутрішніх меж керованої множини розв'язків (6). Для зображеного на рис.1 випадку вони призводять до інтервалу:

$$\mathbf{x} = \left[\frac{\bar{b}}{\bar{a}}; \frac{\underline{b}}{\underline{a}} \right]. \quad (37)$$

Ранжована послідовність елементів множини \mathbf{G} для цього випадку має вид:

$$\frac{\bar{b}}{\underline{a}} < \frac{\underline{b}}{\bar{a}} < \frac{\underline{b}}{\underline{a}} \leq \frac{\bar{b}}{\bar{a}}$$

і відповідає застосуванню виразу (34). Аналогічно доводиться застосування виразу (34) для інших співвідношень між елементами множини \mathbf{G} . Доказ застосування виразу (35) аналогічно.

ВИСНОВКИ

Таким чином, показано, що крім відомих допускових множин розв'язків існують напівдопускові, які разом утворюють повну допускову множину. Це розширює можливості інтервальної математики і дає змогу проводити аналіз і синтез допусків при зовнішніх впливах. Обчислення результатів інтервальних арифметичних операцій відбувається за допомогою подвійного застосування операторів пошуку екстремальної підмножини до множини сполучень результатів арифметичних операцій над межами інтервалів.

ПЕРЕЛІК ПОСИЛАНЬ

1. Алефельд Г., Херцбергер Ю. Введение в интервальные вычисления. - М.: Мир, 1987.
2. Калмыков С.А., Шокин Ю.И., Юдашев З.Х. Методы интервального анализа. - Новосибирск: Наука, 1986.
3. Markov, S.M. Extended interval arithmetic.-C.R. Acad. Bulgare Sci., 1977, v.30, p.1239 - 1242.
4. Вошинин А.П., Сотиров Г.Р. Оптимизация в условиях неопределенности. Москва, МЭИ 1989.
5. Shary, S.P. Algebraic Approach to the Interval Linear Static Identification, Tolerance and Control Problems, or One More Application of Kaucher Arithmetic. Reliable Computing 2 (1) (1996), p.3-33.

II. ИНФОРМАТИКА

УДК 519.873 + 519.7.24/25

LINEAR ALGEBRA AND PETRI NETS

A.Bourjij, M.Boutayeb

This contribution is devoted to fast computation of a basis of invariants (P-invariants or T-invariants) in Petri Nets with extension to large-scale systems. Unlike the previous techniques, our approach consists first, in transforming the homogeneous equation into a reduced form and next, with the aid of performed simple rules, a basis of invariants is deduced. In the second part of this note, we propose a useful decentralized algorithm for computing invariants in large-scale interconnected Petri Nets. The main feature of the proposed technique lies in elaborating each subsystem's decision by using only the local incidence matrix and by the aid of an adjustment procedure we deduce the global solutions taking the constraints interconnection into account. Furthermore, it is shown that computational requirements are reduced considerably in comparison with the global approach. Accuracy and performances of the proposed approach are illustrated through numerical examples.

INTRODUCTION

Petri Net has been one of the most frequently used tools for modeling and analysis of discrete event systems since its origins, about thirty years ago. It has been particularly used for representing computer systems to describe concurrency, conflicts synchronization of processes etc. Furthermore, as a graphical tool, Petri Net is well adapted to supervise dynamical systems in real time and thus to improve performances, such as the reliability and productivity of processes. It is not intended here to give a total overview and summary of the theory and applications of Petri Nets; for more details the reader is referred to T. Murata's [1] paper and the references mentioned inside.

In this note, we are interested in the particular case of determining a basis of invariants, which constitutes a powerful tool for studying the structural properties of Petri Nets[2]-[5], [15]. Several methods have been developed in literature to deal with this problem, see for example [11]-[14]. The basic idea of these techniques is to generate the space solutions by several linear transformations of the incidence matrix [7].

The method we put forward in this paper has an important advantage in that computational requirements are reduced in comparison with the existing methods. This property is particularly relevant when large-scale Petri Nets, or when time varying incidence matrices, are considered. Indeed, many dynamical processes, such as distributed computer systems, are described by a time varying incidence matrix, which may be due to changes in the communication structure, or to changes in system configuration. Thus structural properties must be determined continually in order to enhance the per-

formances of the process. The proposed approach consists first, in transforming the homogeneous equation into a reduced form. We next establish simple rules to determine the whole space solutions. To show the efficiency of the proposed algorithm, a numerical example is provided and computational requirements are compared to those of the method in [13], which is based on the well-know Farkas algorithm[11].

In the second part, we propose a decentralized algorithm for computing invariants in large-scale interconnected Petri Nets, the latter may be obtained from systems that are composed of geographically distributed systems. The main feature of the proposed architecture lies in elaborating each subsystem's decision using only the local incidence matrix and with an adjustment procedure ensures the global solution taking the constraints interconnection into account. In the proposed architecture, we may use any method to determine invariants of the subsystems.

In the last section, we discuss some aspects of the decentralized algorithm implementation and the computational requirements. One advantage of the above structure is the possibility to implement the proposed algorithm on a multi-processor environment where each subsystem is treated by a local processor. What is more, all local solutions are transmitted to a co-ordination processor to deduce the global solution.

Furthermore, we give an idea of the computational savings in the global and decentralized implementations of the proposed method. To show performances of the proposed technique, a numerical example is provided.

BASIC TERMINOLOGY

A Generalized Petri Net is a 4-tuple (P, T, F, M_0) where:

$P = \{P_1, P_2, \dots, P_n\}$ is a finite set of places,

$T = \{T_1, T_2, \dots, T_m\}$ is a finite set of transitions,

$P \cap T = \emptyset$ and $P \cup T \neq \emptyset$,

$F \subseteq (P \times T) \cup (T \times P)$ is a set of arcs,

$M_0: P \rightarrow \mathbf{N}$ the initial marking is a n-vector where the i th entry $M_0(P_i)$ represents the number of tokens in place P_i , \mathbf{N} is the set of positive integer numbers (hereafter, \mathbf{Z} is used as the set of integer numbers, \mathbf{Q} as the set of rational numbers and \mathbf{R} as the set of real numbers).

The fundamental relation of Petri Nets is:

$$M = M_0 + C\bar{\sigma} \quad (1)$$

where $C \in \mathbf{Z}^{n \times m}$ is the incidence matrix, only generalized and capacity PNs are considered.

M is the marking reachable from M_0 through a firing sequence.

$\bar{\sigma} \in \mathbf{N}^m$ is called the firing count vector where the i th entry of $\bar{\sigma}$ denotes the number of times the transition T_i must fire.

T -invariants are positive integer solutions of the homogeneous equation:

$$C \cdot x = 0. \quad (2)$$

It corresponds to the firing sequence which can leads back to the initial marking and can help to improve the liveness of a system. P -invariants are solutions of the transposed incidence matrix:

$$C^T \cdot y = 0.$$

It proves that the token count in a directed circuit is invariant under any firing. The solution of this kind of systems give invariant relations that can be used for the diagnosis of the modeled system [6]. The invariants computing can be reduced to the resolution of a homogeneous system (2) and so, hereafter only T -invariants are considered. To obtain the P -invariants enough is to transpose the incidence matrix C .

In this case the space of the solutions can be extended to \mathbf{Z}^n or \mathbf{Q}^n (as it can be shown in section 3 the space dimension is the same in \mathbf{Q}^n and \mathbf{R}^n).

1. INVARIANT COMPUTING

Most of the classical techniques for P and T -invariants computing consist in starting with the matrix $[C|I]$ and by linear combinations of its columns transform it into an equivalent form $[0|D]$ where the rows of the matrix D are all the invariants. This is an attractive and simple method to deal with this problem, but for large scale Petri Nets computational requirements are fairly numerous.

In the sequel we present a simple algorithm to determine the invariants [20], which has the advantage of being easily implemented and at the same time going faster than the previous techniques.

1. Proposed method

We are going first to reduce the system (2) to an equivalent form and after we give theorems that permit the computation of the solution in different spaces.

Consider the incidence matrix $C \in \mathbf{Z}^{n \times m}$, by appropriate permutations of rows and columns, it can always be parti-

tioned into the following form:

$$C = \begin{pmatrix} C_1 & C_2 \\ C_3 & C_4 \end{pmatrix} \text{ with rank}(C) = m_1 \quad (3)$$

where block $C_1 \in \mathbf{Z}^{m_1 \times m_1}$ is a square full rank matrix, C_1 is then invertible, with $\text{rank}(C) = \text{rank}(C_1)$. C_2 is an m_1 by m_2 block matrix; C_3 and C_4 are block matrices of appropriate dimensions.

As the submatrix $\begin{pmatrix} C_3 & C_4 \end{pmatrix}$ is linearly dependent on $\begin{pmatrix} C_1 & C_2 \end{pmatrix}$, we have

$$\begin{pmatrix} C_3 & C_4 \end{pmatrix}_i = \sum_j b_{ij} \begin{pmatrix} C_1 & C_2 \end{pmatrix}_j, \forall i$$

$\begin{pmatrix} C_3 & C_4 \end{pmatrix}_i$ and $\begin{pmatrix} C_1 & C_2 \end{pmatrix}_j$ denote the i th and the j th rows of $\begin{pmatrix} C_3 & C_4 \end{pmatrix}$ and $\begin{pmatrix} C_1 & C_2 \end{pmatrix}$ respectively.

These constraints are redundant for the system resolution and thus the system (2) is equivalent to:

$$\begin{pmatrix} C_1 & C_2 \end{pmatrix} x = 0 \text{ or } \begin{pmatrix} C_1 & C_2 \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \end{pmatrix} = 0 \quad (4)$$

with $x = \begin{pmatrix} x_1 \\ x_2 \end{pmatrix}$, $x_1 \in \mathbf{N}^{m_1}$, $x_2 \in \mathbf{N}^{m_2}$,

and $m = m_1 + m_2$.

Thus, instead of resolving (2) in the full order, we are interested in finding vectors x so that equation (4) is fulfilled.

From (4) we obtain a reduced and equivalent form:

$$x_1 = Mx_2 \quad (5)$$

with $M = -C_1^{-1}C_2$ or equivalently $M = \frac{1}{\text{Det}(C_1)} \cdot (\text{com}(C_1))^T \cdot C_2$

where $\text{com}(C_1)$ is the cofactors matrix of C_1 .

Our goal now is to find out conditions on the matrix M so as to construct x_1 and x_2 . The vectors x_1 and x_2 are related by equation (5).

Remark:

While the elements of the matrix C_1 are integers, elements of $\text{com}(C_1)$ are integers too. But if $|\text{Det}(C_1)| \neq 1$,

then $M \in (\mathbf{Q})^{m_1 \times m_2}$ and the solutions x can be rational vectors. Thus, the space of the solutions depends on the matrix M , and by an appropriate choice of x_2 we obtain x_1 .

The rest of this section is dispatched into four parts that give the solutions to system (2) in \mathbf{R}^m , \mathbf{Q}^m , \mathbf{Z}^m , and finally \mathbf{N}^m .

Solutions in \mathbf{R}^m

Below, we recall two well-known theorems in system resolution [17].

Theorem 1: The system (2) of n equations and m unknowns has a solution if and only if:

$$\text{Rank}(C) < m \text{ and then } \text{Det}(C) = 0.$$

Theorem 2: Furthermore, if $r = \text{rank}(C) < m$, the system has exactly $m-r$ linear independent solutions.

Thus, the solutions are computed in the same way as in the next part of this section (see solutions in \mathbf{Q}^m), but if C is full column rank ($m_1 = m$) the system has no other solution than the trivial one: $x = 0$ (see theorem 1 above).

Solutions in \mathbf{Q}^m

Theorem 3: The space dimension of the solutions to system (2) is m_2 . By choosing vectors x_2 in the canonical basis or equivalently the columns I_j of an identity matrix, we obtain $x_{ij} = M \cdot I_j$ with $1 \leq j \leq m_2$, then a basis of the space of solutions to the system (2) is given by the columns of the following matrix S

$$S = \begin{bmatrix} M \\ I_{m_2} \end{bmatrix}$$

where M is given by (5) and I_{m_2} is the identity matrix of dimension m_2 .

Proof: Applying theorem 1 and 2 we obtain:

$$\text{Rank}(C) = m_1$$

$$m = m_1 + m_2$$

then the maximum number of linearly independent solutions to this system is m_2 .

The solutions are the columns of the matrix S and we have:

$$\text{Rank}(S) = \text{Rank}(I_{m_2}) = m_2$$

thus we obtain m_2 linearly independent solutions

Solutions in \mathbf{Z}^m

In the previous section we have shown that the matrix M can contain rational elements if $|\text{Det}(C_1)| \neq 1$.

Theorem 4: The integer solutions to the system (2) are obtained by multiplying all the elements of the matrix S by an integer $k = \text{Det}(C_1)$:

$$S = \begin{bmatrix} k \cdot M \\ k \cdot I_{m_2} \end{bmatrix} = \begin{bmatrix} \bar{M} \\ \bar{I}_{m_2} \end{bmatrix}.$$

Proof: (the same as in the previous section)

Solutions in \mathbf{N}^m

Below, two theorems are given to determine the whole space solutions of equation (5):

Theorem 5: If there exists a column M_j of M with only positive components, then the system (2) has exactly m_2 linearly independent solutions in \mathbf{N}^m .

Proof and solutions computation:

If M has a positive column j , we can always choose x_{2j} large enough to verify:

$$x_{1j} = x_{2j} \cdot M_{ij} + \sum_{k=1}^{j-1} x_{2k} \cdot M_{ik} + \sum_{k=j+1}^{m_2} x_{2k} \cdot M_{ik} \geq 0$$

for $1 \leq i \leq m_1$.

To obtain a basis of the solutions we choose: x_2 in the canonical basis of dimension m_2 and set $x_{2j} = \mathbf{a} = |\text{Min}(M_{kh})|$ $1 \leq k \leq m_1$, $1 \leq h \leq m_2$.

For example if $m_2 = 4$, the constructed vectors x_2 are:

$$\begin{pmatrix} 1 \\ \mathbf{a} \\ 0 \\ 0 \end{pmatrix}, \begin{pmatrix} 0 \\ \mathbf{a} \\ 0 \\ 0 \end{pmatrix}, \begin{pmatrix} 0 \\ \mathbf{a} \\ 1 \\ 0 \end{pmatrix}, \begin{pmatrix} 0 \\ \mathbf{a} \\ 0 \\ 1 \end{pmatrix} \leftarrow \text{row } j.$$

The chosen basis is canonical and all other solutions are linear combinations of this basis.

Then the columns of the following matrix S give a basis of the solutions:

$$S = \begin{bmatrix} \bar{M} \cdot R \\ R \cdot \bar{I}_{m_2} \end{bmatrix} \text{ with } R = \begin{pmatrix} 1 & 0 & \dots & 0 \\ \mathbf{a} & \dots & \mathbf{a} & \mathbf{a} \\ 0 & & 1 & 0 \\ 0 & \dots & 0 & 1 \end{pmatrix}.$$

R is an identity matrix with all components of row j set to value \mathbf{a} (where j is the number of the positive column of M).

Axiom 1: If all the components of \bar{M} verify:

$\bar{M}_{ij} \geq 0$ (and not necessarily $\bar{M}_{ij} \geq 0$) then the solutions are immediate and are given by theorem 4.

Lemma 1:

1 - if there exists a row \bar{M}_i of \bar{M} with only strictly negative components:

$$\bar{M}_i/\bar{M}_{ij} < 0, \quad 1 \leq j \leq m_2, \quad \bar{M}_{ij} \in \mathbf{Z}$$

the unique and trivial solution of (5) in \mathbf{N}^m is:

$$x = \begin{pmatrix} x_1 \\ x_2 \end{pmatrix}.$$

2 - if there exists a row \bar{M}_k of \bar{M} with negative or null components

$$\bar{M}_k/\bar{M}_{ij} \leq 0, \quad 1 \leq j \leq m_2, \quad \bar{M}_{ij} \in \mathbf{Z}$$

the space dimension of the solutions is $(m_2 - s)$ where s is the number of strictly negative components.

3 - if contains d ($d > 1$) negative rows (like M_k), we create a vector v :

$$v = \sum_{k=1}^a M_k$$

then the space dimension of the solutions is $(m_2 - s)$ where s is the number of strictly negative components of v . Proof:

1 - From (5) the i th component of x_1 is given by:

$$x_{i1} = \sum_{j=1}^{m_2} M_{ij} x_{j2} \quad (6)$$

as $M_{ij} < 0$ for $j = 1, \dots, m_2$ and with the condition $x_{i1} \geq 0$, the vector x_2 must be set to zero, so from (6) we deduce that also $x_1 = 0$.

2- and 3- to obtain m_2 linearly independent solutions, we choose vectors x_{2j} satisfying:

$$[x_{2j}] \text{ is an } m_2 \times m_2 \text{ matrix} \\ \text{if } \exists k/M_{ik} < 0 \text{ and } M_{i(j \leq 0)} \quad (7)$$

with $1 \leq j \leq m_2$ and $j \neq k$, then to be sure that x is a positive integer vector we must choose $x_{2k} = 0$ and therefore $\text{rank}([x_{2j}]) = m_2 - 1$.

If M has s rows verifying condition (7) then $\text{rank}([x_{2j}]) = m_2 - s$ because $[x_{2j}]$ has s null rows.

Lemma 2: If the theorem 5 and the lemma 1 cannot be applied, a necessary but not sufficient condition to obtain a positive column in matrix M is: each row of M must contain at least one strictly positive component.

Proof:

We can easily find positive integers (b_1, \dots, b_{m_2}) such that (5) is transformed into an equivalent form like:

$$x_1 = \bar{M} D D^{-1} x_2 \quad (8)$$

or
$$x_1 = \bar{M} \bar{x}_2 \quad (9)$$

where \bar{M} is with a positive column \bar{M}_j and the matrix D is in the form:

$$D = \begin{pmatrix} 1 & 0 & \dots & 0 & b_1 & 0 & \dots & 0 \\ 0 & 1 & \dots & 0 & b_2 & 0 & \dots & 0 \\ & & & & \cdot & & & \\ & & & & \cdot & & & \\ 0 & 0 & \dots & 0 & b_{m_2} & 0 & \dots & 1 \end{pmatrix}. \quad (10)$$

The vector $\begin{pmatrix} b_1 \\ b_2 \\ \cdot \\ \cdot \\ b_{m_2} \end{pmatrix}$ is the j th column of D with $b_i > 0$

with $\hat{\bar{M}} = \bar{M} D$ and $x_2 = D \bar{x}_2$. (11)

Then we apply theorem 5 on the obtained system (8), which gives the solutions.

Corollary 1: If the theorem 5 and the lemma 1 cannot be applied and it is not possible to transform the matrix M by use of lemma 2, then this approach cannot be used directly.

But, we can transform the system into an equivalent form, like this:

$$x_1 = M \cdot x_2 \\ \Leftrightarrow [M - 1] \begin{bmatrix} x_1 \\ x_2 \end{bmatrix} = 0$$

$$\text{rank}([x_{2j}]) = m_2.$$

We obtain a new reduced homogeneous system made of two blocks, where the first is compact and the second is sparse. The Farkas algorithm is very appropriate for this system because the columns of the block M contain positive and negative components that are necessary for the transformations of the system.

Finally, the method can be synthesized in the following algorithm:

Step 1:

Compute permutation of rows and columns of C to obtain the four blocs matrix (3) and memorize these permutations.

Step 2:

Compute $\bar{M} = |\text{Det}(C_1)| \cdot C_1^{-1} \cdot C_2$.

Step 3:

If (\bar{M} has a row < 0) Then

The system has no other solution than 0 in \mathbb{N}

Else

Find or construct a positive column of M

Compute the solutions

Endif

Step 4:

With use of memorized permutations made in step 1 rearrange the solutions x .

Step 5:

Extract the minimal support of invariants from the computed basis. To do this we have: first to divide each column by the highest common factor of this column, and after to subtract from each column, which is linear combination of other columns, these latter.

2. Numerical example

Let C be the incidence matrix of a Petri Net with nine places and ten transitions [18]:

$$C = \begin{pmatrix} -1 & 0 & 1 & 1 & 1 & 0 & 0 & 0 & 0 \\ 1 & -1 & -1 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 1 & 0 & -1 & -1 & 0 & 0 & 0 & 0 \\ 0 & 0 & -1 & 0 & -1 & 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & -1 & 0 & 0 & 1 & 0 & 0 \\ -1 & 0 & 0 & 0 & 0 & 0 & 1 & 1 & 1 \\ 1 & 0 & 0 & 0 & 0 & -1 & -1 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 1 & 0 & -1 & -1 \\ 0 & 1 & 0 & 0 & 0 & 0 & -1 & 0 & -1 \\ 0 & 0 & 1 & 0 & 0 & 0 & 0 & -1 & 0 \end{pmatrix}.$$

The step 1 of algorithm 2 gives the two blocks C_1 and C_2 , see (4)

$$C_1 = \begin{pmatrix} -1 & 0 & 1 & 1 & 1 & 0 \\ 1 & -1 & -1 & 0 & 0 & 0 \\ 0 & 0 & -1 & 0 & -1 & 1 \\ 0 & 0 & 0 & -1 & 0 & 0 \\ -1 & 0 & 0 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 & 0 & 0 \end{pmatrix}.$$

C_1 is always a square full rank matrix (then invertible).

$$C_2 = \begin{pmatrix} 0 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \\ 1 & 0 & 0 \\ 1 & 1 & 1 \\ -1 & 0 & -1 \end{pmatrix}$$

then the reduced form (5) can be computed by use of C_1 and C_2 :

$$\text{Det}(C_1) = 1 \Rightarrow M = \bar{M} = \begin{pmatrix} 1 & 1 & 1 \\ 1 & 0 & 1 \\ 0 & 1 & 0 \\ 1 & 0 & 0 \\ 0 & 0 & 1 \\ 0 & 1 & 1 \end{pmatrix}.$$

While all components of M are positive or null we can apply the axiom 1, then the space dimension of the solutions is $m_2 = 3$, and the solutions are given by the columns of the matrix S :

$$S = \begin{bmatrix} M \\ I_{m_2} \end{bmatrix}.$$

After permutations memorized in step 1 and the execution of step 5, T-semiflows are:

$$S_1 = [1 \ 1 \ 0 \ 1 \ 0 \ 0 \ 1 \ 0 \ 0], \quad S_2 = [1 \ 0 \ 1 \ 0 \ 0 \ 1 \ 0 \ 1 \ 0],$$

and $S_3 = [1 \ 1 \ 0 \ 0 \ 1 \ 1 \ 0 \ 0 \ 1].$

To compute P-semiflows we use C' , the transpose of the incidence matrix C . The main blocks C'_1 and C'_2 are:

$$C'_1 = \begin{pmatrix} -1 & 1 & 0 & 0 & -1 & 0 \\ 0 & -1 & 0 & 0 & 0 & 1 \\ 1 & -1 & -1 & 0 & 0 & 0 \\ 1 & 0 & 0 & -1 & 0 & 0 \\ 1 & 0 & -1 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 & 0 & 0 \end{pmatrix}, \quad C'_2 = \begin{pmatrix} 0 & 1 & 0 & 0 \\ 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 \\ -1 & 0 & 0 & 0 \\ -1 & 0 & 0 & 0 \\ 0 & -1 & 1 & 0 \end{pmatrix}$$

$$\text{then } M = \begin{pmatrix} 1 & 1 & -1 & 2 \\ 1 & 0 & 0 & 2 \\ 0 & 1 & -1 & 1 \\ 0 & 1 & -1 & 1 \\ 0 & 0 & 1 & 1 \\ 0 & 0 & 0 & 1 \end{pmatrix}.$$

M has a positive column, then by application of theorem 5, the space dimension of the solution is 4, and a basis of the solutions is the column of S :

$$S = \begin{pmatrix} 3 & 3 & 1 & 2 \\ 3 & 2 & 2 & 2 \\ 3 & 1 & 1 & 1 \\ 1 & 2 & 0 & 1 \\ 1 & 2 & 0 & 1 \\ 1 & 1 & 2 & 1 \\ 1 & 2 & 1 & 1 \\ 0 & 0 & 1 & 0 \\ 1 & 1 & 1 & 1 \\ 1 & 1 & 1 & 1 \end{pmatrix},$$

the step 5 of the algorithm gives the P -invariants:

$$\begin{aligned}
 F_1 &= [1 \ 1 \ 1 \ 0 \ 0 \ 0 \ 0 \ 0 \ 0 \ 0], \\
 F_2 &= [1 \ 0 \ 0 \ 1 \ 1 \ 0 \ 1 \ 0 \ 0 \ 0], \\
 F_3 &= [0 \ 1 \ 0 \ 0 \ 0 \ 2 \ 1 \ 1 \ 1 \ 1], \\
 F_4 &= [1 \ 1 \ 0 \ 1 \ 1 \ 1 \ 1 \ 0 \ 1 \ 1].
 \end{aligned}$$

3. Computational requirements

In this paragraph, we compare execution time of two kinds of algorithms for invariants computing. The algorithm 1 is the Farkas algorithm and the algorithm 2 is our approach. The two algorithms have been implemented as Matlab procedures in a Macintosh centris 650 and tested for different incidence matrices. The obtained results are given in the table 1.

Table 1

| Algorithm | Size of C | Execution time in Sec |
|-----------|----------------|-----------------------|
| 1 | 10×9 | 0.75 |
| 2 | 10×9 | 0.95 |
| 1 | 24×22 | 14.5 |
| 2 | 24×22 | 6.48 |
| 1 | 48×44 | 89.22 |
| 2 | 48×44 | 32.28 |

The incidence matrices are sparse and the use of the reduction procedure (steps 1 and 2) of algorithm 2 takes advantage of this property. Therefore, as can be seen from the previous table, the computational requirements of the proposed procedure are less numerous than for the first algorithm as the size of C (incidence matrix of the PN) increases.

II. DECENTRALIZED APPROACH

1. Problem formulation

For large-scale systems, incidence matrices are huge, thus invariants time computing can be excessive, and analyzing this kind of Petri Net is not easy [16]. To optimize the analysis of these Petri Net models, we propose to decompose huge incidence matrices into much smaller blocks using the following technique[8],[19].

2. Proposed approach

To perform an algebraic analyze of Petri Nets we can compute invariants by solving these two homogeneous systems:

$$C \cdot x = 0 \text{ and } C' \cdot x = 0 \tag{12}$$

An important remark is that C incidence matrices of large-scale systems are sparse. Thus it is always possible, by use of

group technology techniques [9], [10], or other mathematical tools to transform a sparse matrix into an equivalent diagonal block matrix as presented in the fig.1.

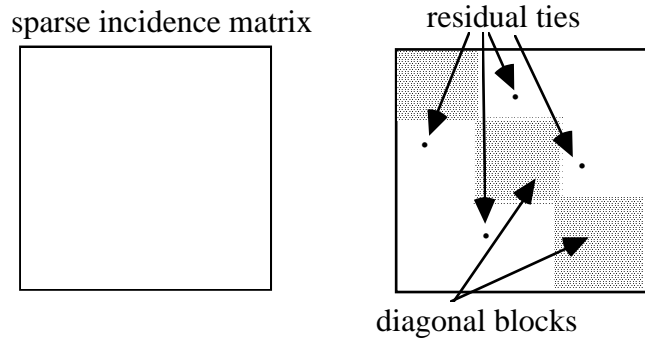


Figure 1

But the different blocks of the obtained matrix cannot be used separately to compute the invariants of the Petri Net due to the presence of residual ties. In order to do that, we apply a second transformation to the incidence matrix.

The columns containing residual ties are moved to the right of the matrix after the last column, as it can be shown in the fig.2.

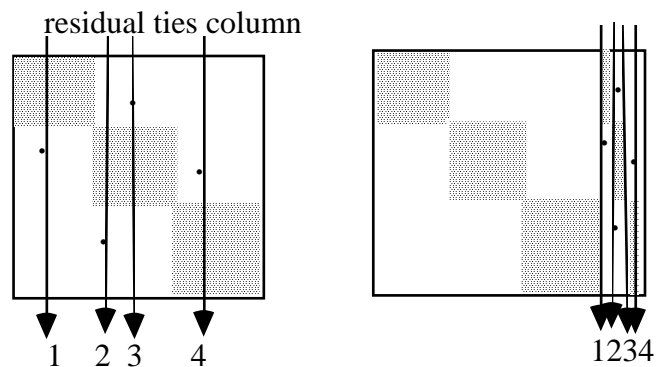


Figure 2

We may recall that all these transformations are regular and do not affect the solutions of the homogeneous system: $C \cdot x = 0$.

Finally, the incidence matrix of large scale Petri Nets is in the form:

$$\begin{pmatrix}
 C_1 & 0 & \dots & 0 & C_{e1} \\
 \dots & \dots & \dots & \dots & \dots \\
 0 & \dots & 0 & C_p & C_{cp}
 \end{pmatrix} \tag{13}$$

In order to analyze one huge system we can analyze p smaller subsystems represented by their incidence matrices:

$$\left(\begin{array}{c} C_i \\ C_{ci} \end{array} \right) \text{ for } i = 1, \dots, p$$

where the block C_{ci} represents the coupling constraints.

Thus invariants are the solutions of the following homogeneous system:

$$\begin{pmatrix} C_1 & 0 & \dots & 0 & C_{c1} \\ \dots & \dots & \dots & \dots & \dots \\ 0 & \dots & 0 & C_p & C_{cp} \end{pmatrix} \begin{pmatrix} x_1 \\ \dots \\ x_p \\ x_c \end{pmatrix} = 0 \quad (14)$$

$$\Leftrightarrow \begin{cases} C_1 x_1 + 0x_2 + \dots + 0x_p + C_{c1} x_c = 0 \\ 0x_1 + C_2 x_2 + \dots + 0x_p + C_{c2} x_c = 0 \\ \dots \\ 0x_1 + 0x_2 + \dots + C_p x_p + C_{cp} x_c = 0 \end{cases} \quad (15)$$

$$\Leftrightarrow \begin{cases} C_1 x_1 + C_{c1} x_c = 0 \\ C_2 x_2 + C_{c2} x_c = 0 \\ \dots \\ C_p x_p + C_{cp} x_c = 0 \end{cases} \quad (16)$$

$$\Leftrightarrow \left(\begin{array}{c} C_i \\ C_{ci} \end{array} \right) \begin{pmatrix} x_i \\ x_c \end{pmatrix} = 0 \text{ for } i = 1, \dots, p. \quad (17)$$

With this decomposition we can notice that the invariants (also called "local invariants" in the following text) of each subsystem (7) can be computed simultaneously in a parallel or distributed architecture computer. Let S_i be a generatrix basis of the solution space of the subsystem (i), $1 \leq i \leq p$:

$$\left(\begin{array}{c} C_i \\ C_{ci} \end{array} \right) \begin{pmatrix} x_i \\ x_{ci} \end{pmatrix} = 0$$

and S_g a generatrix basis of the solution space of the global system (4), we can deduce from the previous equation that:

$$S_g \subseteq \bigcup_{i=1}^p S_i. \quad (18)$$

Remark 1: By the proposed decomposition (7), local invariants computation can only gives a part of the global solution. To obtain the global solution we must concatenate the x_i , $1 \leq i \leq p$ components that have the same coupling part (see corollary below and its proof).

The question is now: what local solutions must be concatenated to made a global solution?

The resolution of each subsystem gives a part $\begin{pmatrix} x_i \\ x_{ci} \end{pmatrix}$

$1 \leq i \leq p$ of the global solution.

A simple procedure to obtain a basis of the solutions of the system (4), is to hold solutions from

$$\bigcup_{i=1}^p S_{ci} / \{0\} \quad (19)$$

where S_{ci} is the set made by the coupling part of the basis of the solution of the subsystem (i).

In other words, the local solutions that have the same coupling part are concatenated to give the global solution.

Corollary 1: A solution $\begin{pmatrix} x_i \\ x_{ci} \end{pmatrix}$ of the subsystem (i) and a

solution $\begin{pmatrix} x_j \\ x_{cj} \end{pmatrix}$ of the subsystem (j) are the same part of a

solution of the global system if and only it have the same coupling components x_{ci} . If $x_{ci} = 0$ the local solution

$$\begin{pmatrix} x_i \\ x_{ci} \end{pmatrix} \text{ is also a global solution } \begin{pmatrix} 0 \\ x_i \\ 0 \\ x_{ci} \end{pmatrix} = \begin{pmatrix} 0 \\ x_i \\ 0 \\ 0 \end{pmatrix}.$$

Proof: For a system made up by two subsystems, let

$\begin{pmatrix} x_1^i \\ x_{c1}^i \end{pmatrix}$ be the ith solution of the subsystem (1) and

$\begin{pmatrix} x_2^j \\ x_{c2}^j \end{pmatrix}$ the jth solution of the subsystem (2). These two

solutions are each a part of the same solution of the global system if and only if:

$$C. \begin{pmatrix} x_1^i \\ x_2^j \\ x_{c1}^i \end{pmatrix} = 0 \text{ and } C. \begin{pmatrix} x_1^i \\ x_2^j \\ x_{c2}^j \end{pmatrix} = 0$$

$$\Leftrightarrow \begin{cases} \left(\begin{array}{c} C_1 & 0 & C_{c1} \\ 0 & C_2 & C_{c2} \end{array} \right) \cdot \begin{pmatrix} x_1^i \\ x_2^j \\ x_{c1}^i \end{pmatrix} = 0 \\ \left(\begin{array}{c} C_1 & 0 & C_{c1} \\ 0 & C_2 & C_{c2} \end{array} \right) \cdot \begin{pmatrix} x_1^i \\ x_2^j \\ x_{c2}^j \end{pmatrix} = 0 \end{cases}$$

$$\Leftrightarrow \begin{cases} C_1 x_1^i + C_{c1} x_{c1}^i = 0 \\ C_2 x_2^j + C_{c2} x_{c1}^i = 0 \\ C_1 x_1^i + C_{c1} x_{c2}^j = 0 \\ C_2 x_2^j + C_{c2} x_{c1}^i = 0 \end{cases}$$

We can deduce: $C_{c1} x_{c1}^i = C_{c1} x_{c2}^j$ and $C_{c2} x_{c1}^i = C_{c2} x_{c2}^j$.

For a system made up by p subsystems the proof is the same.

If a solution $\begin{pmatrix} x_i \\ x_{ci} \end{pmatrix}$ of the subsystem (i) verify:

$$x_{ci} = 0, \text{ (solution with no coupling part)}$$

this local solution is also a global solution. Due to the block diagonal form of the incidence matrix C :

$$\begin{pmatrix} C_1 & \dots & C_{c1} \\ \dots & C_i & \dots \\ 0 & \dots & C_{cp} \end{pmatrix} \cdot \begin{pmatrix} 0 \\ x^i \\ 0 \\ 0 \end{pmatrix} = 0 \text{ if and only if } C_i x_i = 0.$$

Here is the approach summed up in a six-step algorithm:

Step 1:

Break up the incidence matrix of the Petri Net model into a diagonal block matrix (but residual ties appear).

Step 2:

Rearrange the incidence matrix to isolate residual ties.

Thus we obtain separate subsystems.

Step 3:

Compute invariants of each subsystem by appropriate method (the proposed one or the Farkas approach).

Step 4:

By the help of invariants retrieve the structural properties of each subsystem. If the model is inaccurate, it must be modified then repeat from step 1.

Step 5:

By means of corollary 1 assemble local invariants to obtain the invariants of the global system.

Step 6:

Deduce the main properties of the global solution.

3. Computational requirements

Due to the block diagonal decomposition the invariants of each subsystem can be computed separately. Thus, the required computing time τ will be the greatest computing time of each subsystem. If τ_i is the computing time of the subsystem (i), and by using a parallel or distributed computer:

$$\tau = \max_i (\tau_i)$$

we can neglect the necessary time to assemble the local solutions (step 5), but for huge Petri Nets, which are not

interconnected subsystems, the time needed by the step 1 can be significant.

4. Numerical example

The example proposed below is not a huge model, but the method is just the same. This Petri Net model consist of two subsystems coupled by two places P_4 and P_8 (see Fig.3):

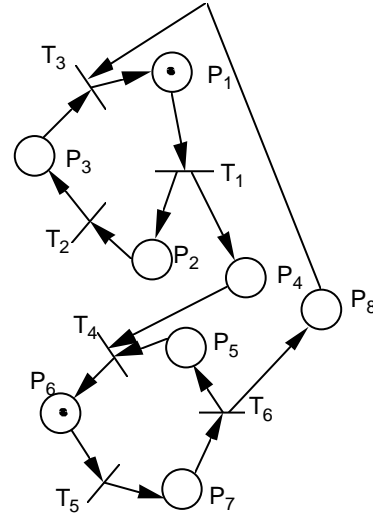


Figure 3 - Petri Net model

Step 1:

The associated incidence matrix is:

$$C^T = \begin{pmatrix} -1 & 1 & 0 & 1 & 0 & 0 & 0 & 0 \\ 0 & -1 & 1 & 0 & 0 & 0 & 0 & 0 \\ 1 & 0 & -1 & 0 & 0 & 0 & 0 & -1 \\ 0 & 0 & 0 & -1 & -1 & 1 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & -1 & 1 & 0 \\ 0 & 0 & 0 & 0 & 1 & 0 & -1 & 1 \end{pmatrix} \begin{matrix} \leftarrow C_1 \\ \leftarrow C_2 \\ \leftarrow C_{c1} \\ \leftarrow C_{c2} \end{matrix}$$

1 2 3 4 5 6 7 8 \leftarrow $\begin{cases} \text{column} \\ \text{numbers} \end{cases}$

Step 2:

This matrix is sparse and thus if we move columns 4 and 8 (these columns corresponds to the coupling part between the two subsystems) to the end of C , we obtain an equivalent matrix with two interconnected subsystems like the diagonal blocks matrix presented in (3):

$$C^T = \begin{pmatrix} -1 & 1 & 0 & 0 & 0 & 0 & 1 & 0 \\ 0 & -1 & 1 & 0 & 0 & 0 & 0 & 0 \\ 1 & 0 & -1 & 0 & 0 & 0 & 0 & -1 \\ 0 & 0 & 0 & -1 & 1 & 0 & -1 & 0 \\ 0 & 0 & 0 & 0 & -1 & 1 & 0 & 0 \\ 0 & 0 & 0 & 1 & 0 & -1 & 0 & 1 \end{pmatrix} \begin{matrix} \leftarrow C_1 \\ \leftarrow C_{c1} \\ \leftarrow C_2 \\ \leftarrow C_{c2} \end{matrix}$$

Step 3:

Now we can solve the system by the decentralized method.
The resolution of the subsystem (1):

$$\begin{pmatrix} C_1 & C_{c1} \end{pmatrix} \begin{pmatrix} x_1 \\ x_{c1} \end{pmatrix} = 0$$

gives two solutions:

$$X^1 = \begin{pmatrix} 1 \\ 1 \\ 1 \\ 0 \\ 0 \end{pmatrix} \text{ where } X_1^1 = \begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix} = \begin{pmatrix} 1 \\ 1 \\ 1 \end{pmatrix}$$

$$\text{and } X_{c1}^1 = \begin{pmatrix} x_4 \\ x_8 \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \end{pmatrix}.$$

$$X^2 = \begin{pmatrix} 1 \\ 1 \\ 1 \\ 0 \\ 0 \end{pmatrix} \text{ where } X_1^2 = \begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix} = \begin{pmatrix} 1 \\ 0 \\ 0 \end{pmatrix}$$

$$\text{and } X_{c1}^2 = \begin{pmatrix} x_4 \\ x_8 \end{pmatrix} = \begin{pmatrix} 1 \\ 1 \end{pmatrix}.$$

The resolution of the subsystem (2):

$$\begin{pmatrix} C_2 & C_{c2} \end{pmatrix} \begin{pmatrix} x_2 \\ x_{c2} \end{pmatrix} = 0$$

gives two solutions:

$$X^3 = \begin{pmatrix} 1 \\ 1 \\ 1 \\ 0 \\ 0 \end{pmatrix} \text{ where } X_2^3 = \begin{pmatrix} x_5 \\ x_6 \\ x_7 \end{pmatrix} = \begin{pmatrix} 1 \\ 1 \\ 1 \end{pmatrix}$$

$$\text{and } X_{c2}^3 = \begin{pmatrix} x_4 \\ x_8 \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \end{pmatrix}.$$

$$X^4 = \begin{pmatrix} 1 \\ 0 \\ 0 \\ 1 \\ 1 \end{pmatrix} \text{ where } X_2^4 = \begin{pmatrix} x_5 \\ x_6 \\ x_7 \end{pmatrix} = \begin{pmatrix} 1 \\ 0 \\ 0 \end{pmatrix}$$

$$\text{and } X_{c2}^4 = \begin{pmatrix} x_4 \\ x_8 \end{pmatrix} = \begin{pmatrix} 1 \\ 1 \end{pmatrix}.$$

Step 4:

With the computed local P -invariants we can proof that the subsystems (1) and (2) are bounded.

Step 5:

The common coupling blocks are given by:

$$\bigcap_{i=1}^p S_{c1} = \bigcap_{i=1}^2 S_{c1} = \left\{ \begin{pmatrix} 0 \\ 0 \end{pmatrix}, \begin{pmatrix} 1 \\ 1 \end{pmatrix} \right\}.$$

Thus the global solutions are obtained by concatenation of the local solutions which have the coupling part $\begin{pmatrix} 1 \\ 1 \end{pmatrix}$,

where local solutions X_1 and X_4 which have no coupling part are then global solutions:

$$S_g = \{(1, 1, 1, 0, 0, 0, 0, 0)^T, (0, 0, 0, 0, 1, 1, 1, 0)^T, (1, 0, 0, 1, 0, 1, 1, 1)^T\}.$$

Step 6:

These three P -invariants can now be used to prove that this Petri Net is bounded because all the places are in a P -invariant relation:

$$M(P_1) + M(P_2) + M(P_3) = 1,$$

$$M(P_5) + M(P_6) + M(P_7) = 1,$$

$$M(P_1) + M(P_4) + M(P_6) + M(P_8) = 1.$$

With the same method we can compute T -invariants and verify the reversibility of the Petri Net for the given initial marking of the figure 3.

III. CONCLUSION

In this paper, a new approach to determine invariants in Petri Nets has been presented. We have shown that the invariants may be obtained directly from the reduced form of the homogeneous equation. One major advantage of the approach presented above is the low computational requirements in the case of large scale Petri Nets. Besides the space dimension of the solution can be given directly (without major computations) and that space can be reduced from \mathbf{Q}^m to \mathbf{Z}^m and finally to \mathbf{N}^m . This has been shown by means of a numerical examples.

A decentralized method to analyze huge systems modeled by Petri Nets is developed. The proposed algorithm consists first in partitioning the incidence matrix and determining independently the invariants of each subsystem; this may be achieved by classical methods. We give a corollary for deducing the invariants of the global system. This approach has the

main advantage that it permits a structural analysis of all the subsystems that compose the system and a structural analysis of the global system. We have shown that the architecture presented reduces significantly the computational requirements, especially for a high number of interconnected subsystems and by using a multi-processor environment. We are now trying to evaluate the performances of our algorithms in terms of elementary operations (like addition, multiplication,...).

REFERENCES

- [1] T. Murata, "Petri Nets: properties, analysis and applications," Proceedings of IEEE Vol. 77, N° 4, pp. 541-580, April 1989.
- [2] R. David, H. Alla "Petri Nets for modeling of dynamic systems". Automatica, 2:175-202,1994.
- [3] G. Berthelot "Verification des Reseaux de Petri" These de doctorat de troisieme cycle, universite P. et M. Curie (PARIS IV), 12 January 1978.
- [4] K. Lauterbach, "Linear algebraic calculation of deadlocks and traps" {\it Concurrency and nets}, ed. K. Voss and H. J. Genrich, Springer-Verlag, 1987.
- [5] E. R. Boer and T. Murata, "Generating basis siphons and traps of Petri nets using sign incidence matrix" {\it IEEE Trans. Circuits and Systems}, Vol. 41, No 4, pp 266-271, April 1994.
- [6] A. Bourjij & al. "Algebraic Analysis of Discrete Event Systems Modelled by Generalized Petri Nets by use of Invariants" Mathematical Theory of Networks and Systems MTNS'96, Saint-Louis Missouri, June 1996.
- [7] A. Bourjij "Contribution a la surete de fonctionnement des processus industriels par les reseaux de Petri", PhD Thesis, University Henri Poincare of Nancy I, France, 5 Dec, 1994.
- [8] M. Boutayeb, A. Bourjij, M. Darouach, "A decentralized algorithm to determine invariants in Petri Nets" MATHMOD Vienna, IMACS, 5-7 February, 1997.
- [9] A. Kusiak, S. C. Wing, "Decomposition of manufacturing systems" IEEE Journal of Robotics and Automation, Vol. 4, N°5, Oct. 1988.
- [10] N. Dridi, J. M. Proth "Ordonnancement des taches: une methode basee sur la technologie de groupes" Second international Conference on Production Systems, INRIA, tome 1, pp. 67-74, April 1987.
- [11] J. Farkas, "Theorie der einfachen Ungleichungen" J.F.d. reine Angew Math., 124, 1902, 1-27.
- [12] J. Martinez and M. Silva, "A simple and fast algorithm to obtain all invariants of a generalized Petri Net," 2nd European workshop on application and theory of Petri Net, Bod-honneff, 1981.
- [13] H. Alaiwan and J.M. Toudic, "Recherche de semi-flots, de verrous et de trappes dans les reseaux de Petri," TSI Vol. 4, N° 1, pp. 104-112, 1985
- [14] M. Silva and J.M. Colom, "On the computation of structural synchronic invariants in P/T nets," Lecture Notes in Computer Science Vol. 340, pp. 386-417, Springer-Verlag, 1988.
- [15] T. Murata, B. Shenker and S.M. Shatz, "Detection of Ada deadlocks using Petri Net invariants" IEEE Trans. on Software Engineering Vol. 15, N° 3, pp. 314-326, march 1989.
- [16] D.C. Marinescu, M. Beaven and R. Stansifer "A parallel algorithm for computing invariants of Petri Net models" Proceedings of the fourth international workshop on PN and performance models, Melbourne, IEEE Computer Soc., pp. 136-43, 1991.
- [17] P. Naudin and C.Quitte "Algorithme Algebrique" Masson 1992.
- [18] K. Garg "An approach to performance specification of communication protocols using timed Petri Nets" IEEE Trans. on Software Engineering Vol. SE11, N° 10, pp. 1216-1224, 1985.
- [19] A. Bourjij & al. "A decentralized approach for computing invariants in Petri Nets". IEEE SMC'97 orlando pp 1741-1746.
- [20] A. Bourjij & al. "On generating a basis of invariants in Petri Nets". IEEE SMC'97 orlando pp 2228-2233.

УДК 681.32

BUILDING EUROPEAN COMPETENCE CENTRES TO SUPPORT KNOWLEDGE DISTRIBUTION AND LEARNING PROCESSES

I.Hamburg, S.Balanica

В статье представлены результаты сотрудничества между Германией и Румынией в рамках проекта EU-TEMPUS INCOT по разработке Web-центра. Этот центр поддерживает управление знаниями и европейскую сеть экспертов.

In this paper some results of a co-operation between the Germany and Romania within the EU-TEMPUS-project INCOT about the development of a Web-based competence centre are presented. This centre supports processes of knowledge management and telelearning and an European network of experts.

1. INTRODUCTION

The Internet with different forms of supported networking and knowledge management processes play an important role within the transition to the information society and knowledge-based economy.

On one hand, new information technologies (IT) e.g. the Internet and multimedia support knowledge management through services like information-transfer, navigation and access to different applications; on the other hand, the Internet causes an increasing flood of information and a decreasing validity period of it. So (potential) users of information

should be able to identify relevant one, to transform it in goal-oriented skills (knowledge generating), to process this knowledge according to the target group and so to contribute to the development of the organization they work with.

Innovative forms of IT-based learning (e-learning) like open distance learning (ODL) by using the Internet - telelearning - are required in order to achieve such key competencies for the staff of the companies.

In part 2 of this paper we present some aspects of knowledge management (KM), particularly knowledge distribution (KD) which are important for manufacturing companies and some possibilities to support it by using IT -[1].

In part 3, telelearning methods and virtual communities which are based on the World Wide Web (WWW) are briefly described.

WWW is a global information system with a hypermedia structure; it is independent on the operating system and facilitates the searching, retrieval and browsing of information via the Internet.

Virtual Competence centres which can be developed as suitable forums for KM, communication and experience

exchange between experts and for learning in a determined field are presented in part 4.

A Web-based virtual environment which supports a competence centre and a network of experts in telelearning from Germany and Romania and is being developed within the EU-Tempus project INCOT is presented in part 5.

2. KNOWLEDGE MANAGEMENT

Effective KM is considered as an important step to intelligent company building. The tasks knowledge acquisition (KA), knowledge identification (KI), knowledge codification (KC), knowledge distribution (KD) -[2] and knowledge use (KU) are assigned to KM in order to deal successfully with knowledge in the companies.

Many companies lose big potentials because knowledge which already exists (e.g. tacit knowledge in the brain of experts or explicitly one in archives) is not further processed, intensified, distributed and used.

Some experts consider that many companies know only a fraction of their knowledge which is potential accessible. Some causes can be the following -[3]:

- Missing methods for the identification of internal and external sources (e.g. individual knowledge of experts) and for the processing of containing knowledge,
- Missing structure for the transformation of it in knowledge accessible for all company staff,
- Troubles at the separation of "important knowledge" from "garbage".

Figure 1 shows three important trends which are connected between them and which produce "turbulence" instead of "transparency" within the knowledge society: extremely fast propagation of knowledge, far-reaching fragmentation and increasing globalisation of it. These determine also the structure of knowledge environment in the companies, which is more complex than some decades ago.

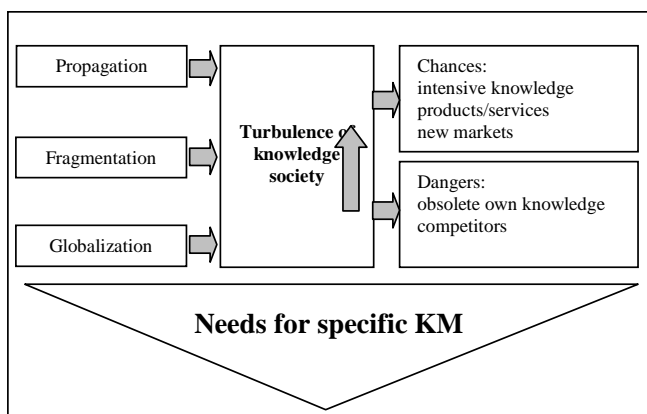


Figure 1 - Trends -[3]

KD is one of the main modules of KM. Some of main functions of it are the followings:

- Multiplication of knowledge through fast distribution on many persons,

- Saving and sharing of previous experience,
- Simultaneous knowledge exchange, which led to the creation of new knowledge.

Knowledge is distributed within the organization through carriers, named vehicles of knowledge -[4]. A vehicle of knowledge is a means by which knowledge can be transferred in spite of existing time and space barriers in the organization.

Such vehicles are the following:

- Organizational, like methods and organizational structures; they can help e.g. in case of projects to provide a common language,
- People as a source and carrier of knowledge; it is important to combine individual learning with organizational learning while people work co-operatively,
- Knowledge management systems (KMS), e.g. IT that support knowledge management.

Knowledge can be distributed:

- Through face to face, personal communication between individuals within work processes, meetings, training, seminars, social activities,
- By using IT like the Internet,
- By creating efficient virtual knowledge networks -[5] supported by virtual competence centres (see part 4).

The Web, which was developed in 1990 at CERN in Geneva to help researchers in sharing information and documents, is an important client-server based knowledge retrieval tool.

The management often controls knowledge processes in companies. Some barriers for KD in many companies are individual or cultural ones. They could have functional sources or hierarchical ones and refer to questions of power or confidence (Figure 2).

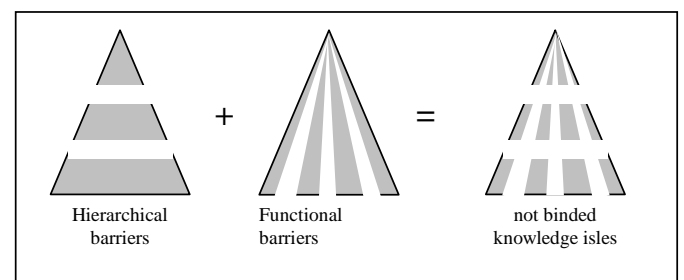


Figure 2 - Barriers -[3]

3. TELELEARNING AND VIRTUAL LEARNING COMMUNITIES

Telelearning e.g. ODL via the Internet is one of the "most promising" forms of learning which can be applied within continuous vocational training programs but it means the beginning of a new paradigm of education: tutor and participants are "separated" in space and time, there is no necessity for courses to be place- or time-based, the student can control the learning process more than in traditional class rooms.

Today, the Web is a basic element in educational telelearning environments -[6].

It is able to support images as well as many other file types, including hypermedia documents containing sound and videos, which can be viewed with external viewer programs. In order to be able to send and receive hypermedia files, servers and clients communicate using the HyperText Transmission Protocol (HTTP). Files are written by using the Hypertext Markup Language (HTML), which is a language for logical document layout. Web-based learning technologies are drastically changing how distance learners communicate with their professors and with each other. Nowadays, distance students have not to wait days and weeks for responses; e-mails are answered faster than phone calls or faxes. They can access not only local libraries but also other vast amounts of information available on the Internet.

A main advantage of the Web is that information is dynamic. So trainers can update their material whenever there is a need for it, students are kept up to date with the most recent changes and news about the learning material. The hyperlinks allow students to browse, jump between links in order to find relevant information and to perform searches easily. Flexibility is another main advantage of web-based learning. This refers to flexibility of time or/and of place of study, to the choice ("modularity") of learning content, to possible frequent updates and to the ability to respond to specific demands. So the Internet and Web-based learning facilitates the development of continuous and life-long learning processes by serving ever-changing needs of learners throughout their life span and by supporting connections between learning and working. In order to achieve this objective, the content of web-based learning modules can be closely integrated with the work of the trainees, and the knowledge and skills they are learning should find immediate application within their work and life.

Features offered by telelearning Web-based software are useful to students, instructors and administrators.

In the following we present some of these features. At the development of our telelearning environment we are going to implement some of them (see part 5).

Student-related features typically include communication tools like asynchronous conferencing among course participants, e-mails to individual course participants or to groups and discussion forum, course content search tools, tests to be performed at specific points in the course, self-evaluation multiple choice quizzes that can be performed at any time as well as presentation tools including students home pages.

Instructor-related features typically are content-related tools referring authoring and/or the ability to import content from other software, management tools to set up students groups for discussion, testing tools including automatic marking of multiple choice quizzes.

Features for administration could be creation and deletion of individual courses, performance monitoring for response time of Web servers, statistics related to number of students accessing the system at different times of day, access control and on-line registration.

The use of the Web has also limitations like the following:

- powerful computers are needed for users in order to fully benefit from the potential of the WWW,

- much time is consumed for downloading multimedia files or accessing popular URLs because of limited band width on Internet,

- the evaluation of thousands of links which are found by students related to their subject is often difficult,

- sometimes-active addresses given by instructors as references disappear before the students access them due to the dynamic process of updating and modifying materials.

Another downside of Web-based virtual classrooms is that distance students are not able to meet trainers/tutors and classmates face-to-face. Social communication, which is an essential component of every educational activity, is missing. The Co-operation, which in traditional classroom settings happens sometimes instantaneously, has to be facilitated by a conscious effort to build learning communities -[7].

An online learning community cannot be created by one person. Although the trainer (professor) is responsible for facilitating the creation process, the community or learning network is entirely dependent on participants and their willingness to contribute time, attention, ideas and commitment. In order to be considered "present" in a virtual classroom, the students have to access the course site online regularly, and to put in their answers and comments.

Stephen Covey -[8] affirms that the collaboration involved in learning in communities creates a sense of synergy and an atmosphere of excitement and passion for learning and working together, supporting the process of knowledge acquisition. In the traditional model of training the tutor imparts knowledge to participants, who are expected to absorb it. In the telelearning processes, the relationships and interactions among people are main objectives for generating knowledge. It is important that the tutor acts as a facilitator of the knowledge acquisition process. This process is supported by the choice of exercises to be solved by the participants. Therefore, the exercises phrasing has to be clear, to promote the distribution of existing individual knowledge and the creation of new individual and collective knowledge to be used by the participants. Another important factor - particularly in the field of staff qualification - is that learning environments foster personal meaning through interactions with communities of learners.

In the following we present the Web-based virtual competence centres, which can promote telelearning processes.

4. VIRTUAL COMPETENCE CENTRES

Virtual competence centres (VCCs) can be understood as a combination of competence centres (CC) and virtual communities (e.g. telelearning ones). VCC structure the relations between different participating groups (persons) in a new way.

In many cases VCCs are developed as professional platforms for KD, learning or communication and facilitate staff of different organizations to get information referring to different topics, products, services as well as to exchange experiences within discussion forums, per e-mail or in chat-rooms.

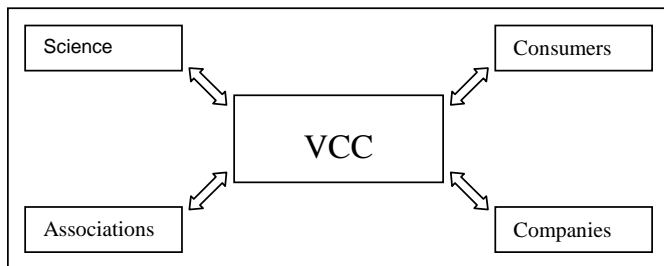


Figure 3 - Potential participants within a VCC -[9]

Within a Web-based VCC, the user can have advantages from the hypermedia-based preparing of document contents. In this case, the corresponding computers have to fulfil some requirements (see part 3) in order to avoid long waiting periods for the downloading of multimedia documents.

One of the main objectives of a VCC is to develop a durable partnership between the participating persons (groups) not only to support a one or two time visit of the centre.

VCCs facilitate the learning from experience of other people through the advantages of new IT.

Through the unification of previous CC and virtual communities advantages, the VCCs support the combination of proved static elements with dynamical, forward-looking ones.

There exist many VCCs, also in German-speaking area. They differ in the type and structure of the co-ordination and of the users, in topics, in regional focus as well as in the content.

Some characteristics of the existing VCCs are the followings -[9]:

- Many of them are an early stage of development,
- In most cases they are not really VCCs but service centres with additional entrance for the Internet,
- Only few VCCs can realize a close connection with the clients through registration and limitation of the users,
- Often they only supply information and do not offer interaction and communication services,
- Many of the VCCs could be improved by using multimedia and other new IT,
- They are rarely regional oriented.

5. A WEB-BASED ENVIRONMENT TO SUPPORT A COMPETENCE CENTRE

For the development of the environment supporting the virtual centre we chose WEB because is the IT-distributed medium the most easy to access. Nowadays, you can access the web from your office, from your home, from your car, and even from your mobile phone. Every firewall, no matter how tight the security is, allows outgoing web connections. In short, World-Wide-Web, with it's support for images as well as many other file types, including sound and videos, is the most popular environment on the Internet. That's why, most of the times, it is associated with the Internet.

In the following we present some requirements on the software environment:

- A web server (Apache, IIS, etc), PHP 3.0+, and a MySQL server (either on the same machine, or on a remote one),

- For the client, a cookie-enabled browser (IE 3.0+, Netscape 2.0+).

The Security Module provides the support for several security levels clearance, as well as a way to verify if a certain person who requires accessing the VCC is a member of the community. This module can add new members, granting them the user level, after they complete an on-line form containing miscellaneous information (name, age, e-mail, fields of interest, etc.).

The Virtual Community Database is the database where all the information regarding a member is kept: name, e-mail, fields of interest, age, etc. This database can be accessed for updates by the security module, the testing module and the members that are administrators.

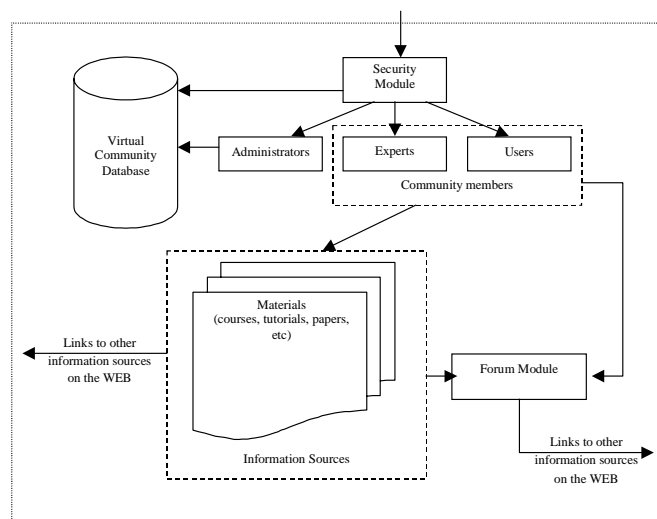


Figure 4 - The environment, which supports the VCC

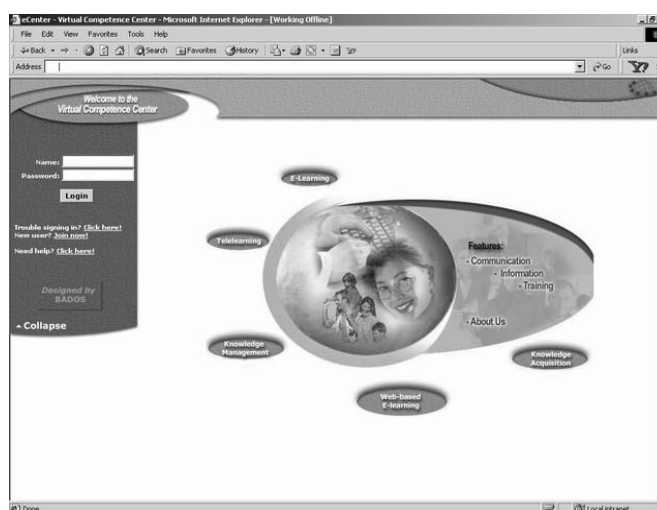


Figure 5 - The homepage of the competence centre

Information Sources module is a WEB-based document collection regarding various fields of interest. An expert, who will be the owner of the source, can add an information source.

The Forum Module provides further discussion regarding an information source where members can share information, experience and opinions.

Non-members of the centre will be able to get information about this Competence Centre, browse the categories available, but not access them, or the discussion forum. Of course, they will be able to register on-line and get access to all the information.

The community members are divided into three categories:

- Administrators: They add other administrators and experts, remove other users from the community, to change the owners of an information source, delete information sources and entries in any discussion forum).

- Experts: They can add new information sources, they are considered to be the "owners" of the information source they have added (they can modify or delete it), and they can share the ownership of an information source with other experts. They are also allowed to create test question (and modify and delete them) for the training materials that they own. If they post an entry in a discussion forum regarding an information source that they own, their entry is shown in a different, high contrast font, being considered as an "authorized entry".

- Users: This category of community members represents the "knowledge seekers". They can access here information sources, access the discussion forums, and contact the "owners" of an information source, or the poster of an entry in a discussion forum, for additional information.

Because the environment uses a modular approach, it is

easy to maintain, update, develop or reuse it. The object-oriented facilities offered by the language PHP allows that also the modules have this orientation. They are easier and faster even a module-level upg.

CONCLUSION

Virtual forms of co-operation and communication like VCCs which support processes of KM and Internet-based ODL have the future before them.

Within our EU-co-operation we would like to work out didactical principles of Internet-based ODL and to expand our VCC for other European countries.

REFERENCES

- [1] *Brodner P., Hamburg I.* (Eds.) *Strategische Wissensnetze: Wie Unternehmen die Resource Wissen nutzen*, Schneider, Gelsenkirchen, 1999.
- [2] *Brodner P., Helmstadter E., Widmaier B.* (Hg.) *Wissensteilung - zur Dynamik von Innovation und kollektivem Lernen*, Munchen/Mering, 1999.
- [3] Probst, G. J. B. *Wissen managen: wie Unternehmen ihre wertvollste Ressource optimal nutzen*, Frankfurt am Main. New York, 1998.
- [4] HART H., 1999: Knowledge transfer and knowledge management in a global organisation // P. Brodner, I. Hamburg (Hrsg.), *Strategische Wissensnetze: Wie Unternehmen die Ressource Wissen nutzen*, Schneider, Gelsenkirchen, 27-46.
- [5] *Engert S., Hamburg I., Terstriep J.* *Kompetenznetzwerke zur Kontextsteuerung von Wissensteilung: Ein Beispiel*. 1999.
- [6] *Engert S., Hamburg I., Terstriep J.* *Web based Training: Experience within a German Project*. // Supporting Learning Communities in Education, Paper presented at the i3 Spring Days 2000, Athens.
- [7] *Pallof R., Pratt K.* *Building Learning Communities in Cyberspace*, San Francisco, 1999.
- [8] *Covey S.*, *The Seven Habits of Highly Effective People: Powerful Lessons in Personal Change*, Fireside, New York 1989.
- [9] *Schmidtman A., Felser W.* *Virtuelle Kompetenzzentren*. // WIRTSCHAFTSINFORMATIK, 1999, 41 6, S. 554-560.

УДК 681.32:007.58

ОБЗОР СОВРЕМЕННЫХ ТИПОВ НЕЙРОННЫХ СЕТЕЙ

М.В.Аникеев, Л.К.Бабенко, О.Б.Макаревич

В работе рассматривается современное состояние теории нейронных сетей. Проведен обзор некоторых наиболее часто применяемых в настоящее время нейросетевых архитектур. Рассмотрены различные подходы к обучению, проектированию и настройке нейронных сетей. Особое внимание уделено различным вариантам метода обратного распространения ошибки.

The current theory of neural networks is considered in this paper. Some of the most widely used architectures of neural networks are reviewed. Various methods of training and design are analyzed. Particular attention is paid to the backpropagation method of training.

ВВЕДЕНИЕ

Возможности современных компьютеров позволяют производить различные вычисления со скоростью на десятки порядков превышающей возможности челове-

ческого мозга. Однако ряд даже тривиальных для человека задач, не связанных с вычислениями, остается весьма сложным для вычислительной техники. Способность человека к ассоциативному хранению информации, обучению, обобщению и обработки информации с учетом контекста остается непревзойденной даже для современных суперкомпьютеров. Целью проектирования искусственных нейронных сетей (ИНС) является построение вычислительной структуры или алгоритма, работающего по принципам естественного интеллекта. К основным можно отнести следующие свойства нейронных сетей.

1. Нейронные сети, по аналогии с мозгом человека и животных, строятся из множества простых элементов, выполняющих элементарные действия и соединенных между собой различными связями.

2. Нейронные сети способны совершенствовать свою работу (обучаться или адаптироваться), используя примеры.

3. Нейросетевое решение задачи не требует от разработчика формулировки алгоритма решения поставленной задачи и его программирования. Нейронные сети, как правило, используют примеры "правильной" работы для построения своего метода решения задачи. При этом существует возможность выявления сетью скрытых закономерностей в задаче, неизвестных разработчику.

1. ОСНОВНЫЕ ПОНЯТИЯ АРХИТЕКТУРЫ ИСКУССТВЕННЫХ НЕЙРОННЫХ СЕТЕЙ

Основным элементом искусственных нейронных сетей является аппаратный или программный узел, имитирующий работу нейрона. Обобщенная схема базовой модели нейрона представлена на рис. 1.1.

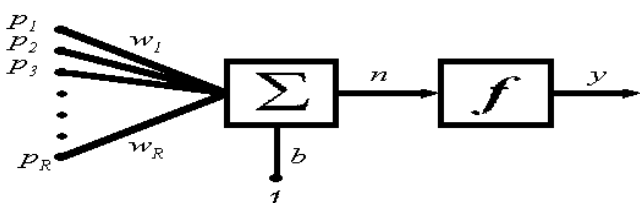


Рисунок 1.1

Здесь $\{p_1, p_2, \dots, p_R\} \equiv \mathbf{p}$ - сигнал, подаваемый на вход нейрона, $\{w_1, w_2, \dots, w_R\} \equiv \mathbf{w}$ - вектор весовых коэффициентов нейрона, b - коэффициент смещения (порог) нейрона, f - функция активации, y - выходной сигнал. Таким образом, нейрон реализует функцию $y = f(n)$, где $n = \sum_{j=1}^R p_j w_j + b$. Обобщенная функция нейрона в векторной форме записывается следующим образом

$$y = f(\mathbf{w} \cdot \mathbf{p} + b). \tag{1.1}$$

Перечислим основные типы активационных функций, используемые в ИНС.

1. *Пороговые функции.* Используются при проектировании сетей персептронного типа, предназначенных для решения задач линейной классификации. Пороговая функция существует в двух модификациях: асимметричная, описываемая формулой

$$f(n) = \begin{cases} 0, & \text{если } n < 0 \\ 1, & \text{если } n \geq 0 \end{cases}, \tag{1.2}$$

и симметричная

$$f(n) = \begin{cases} -1, & \text{если } n < 0 \\ 1, & \text{если } n \geq 0 \end{cases}. \tag{1.3}$$

2. *Линейная функция.* Используется в различных классах сетей для линейной аппроксимации линейных и нелинейных функций. Выражается формулой

$$f(n) = n. \tag{1.4}$$

3. *Сжимающие функции.* Существуют в двух основных модификациях: сигмоидальной (логистической)

$$f(n) = \frac{1}{1 + e^{-n}} \tag{1.5}$$

и тангенциальной

$$f(n) = \text{th}(n) = \frac{e^{2n} - 1}{e^{2n} + 1}. \tag{1.6}$$

Смысл использования сжимающих функций состоит в преобразовании входного сигнала из бесконечного диапазона в конечный: $[0, 1]$ в первом, и $[-1, 1]$ во втором случае. Кроме того, сжимающие функции обладают полезным свойством - они непрерывны и дифференцируемы на всей области действительных чисел, что позволяет обучать сети нейронов со сжимающими активационными функциями методом обратного распространения ошибки. В [1] предложена функция (1.7), обладающая всеми достоинствами логистической функции (дифференцируемость, конечный интервал значений, возрастание, простая производная), но просчитываемая быстрее на вычислительной технике.

$$f(n) = \frac{n}{1 + |n|}. \tag{1.7}$$

4. *Радиально-базисная функция (Гауссиана).* Используется в радиально-базисных сетях и описывается формулой

$$f(n) = e^{-n^2}.$$

Эта функция достигает своего максимума при $n=0$ и поэтому применяется для определения степени близости двух векторов - входного \mathbf{p} и вектора весовых коэффициентов \mathbf{w} .

Применяемые в нейронных сетях активационные функции не ограничиваются приведенным выше списком. Однако представленные четыре класса функций охватывают подавляющее число нейронных сетей.

Слоем нейронов называется совокупность параллельно подключенных нейронов, обрабатывающих в один и тот же момент времени один входной вектор. Выход нейронного слоя выражается формулой

$$\bar{y}_{S \times 1}(\bar{p}_{R \times 1}) = f(\mathbf{W}_{S \times R} \bar{p}_{R \times 1} + \bar{b}_{S \times 1}). \tag{1.8}$$

Матрица \mathbf{W} имеет размерность $S \times R$, где S - число нейронов в слое, а R - размерность входного вектора.

Все нейронные сети можно разделить на два больших

класса - сети прямого распространения и сети с обратными связями. Сети прямого распространения представляют собой упорядоченные последовательности нейронных слоев, в которых сигналы передаются последовательно от входного слоя к выходному. Слои, расположенные между входным и выходным слоями называются скрытыми (иногда входной слой называют первым скрытым). В сетях с обратными связями, как видно из их названия, сигналы могут передаваться от последующих слоев к предыдущим, а также между нейронами одного и того же слоя.

Для успешного функционирования нейронной сети необходимо соответствующим образом настроить ее весовые коэффициенты. Для некоторых типов сетей (например, для адаптивных линейных сетей, описанных в гл. 3) существуют методы построения обученных конфигураций из постановки задачи. Однако подавляющее большинство нейронных сетей необходимо обучать после построения. Существуют две основные парадигмы обучения: обучение "с учителем" (supervised manner) и обучение "без учителя" (unsupervised manner) [2-4].

Обучение "с учителем" характерно для сетей прямого распространения. В процессе обучения "с учителем" сети последовательно предъявляются пары $\langle \mathbf{p}_i, \mathbf{d}_i \rangle$, где \mathbf{p}_i - входной вектор, \mathbf{d}_i - желаемый выходной вектор. При этом сеть пытается настроить свои веса таким образом, чтобы при предъявлении на вход вектора \mathbf{p}_i , получать на выходе вектор максимально близкий к \mathbf{d}_i . После окончания процедуры обучения можно надеяться, что нейронная сеть правильно отреагирует на входные вектора, не входящие в обучающую выборку. Способ обучения "без учителя" характерен для сетей с обратными связями. Получая входное воздействие, сеть подстраивает свои веса согласно своим внутренним целям. Например, соревновательные сети (гл. 6) в процессе работы учатся относить близко расположенные входные вектора к одному классу.

2. ОДНОСЛОЙНЫЕ ПЕРСЕПТРОНЫ

Однослойными персептронами называются однослойные сети прямого распространения, содержащие нейроны, которые активизируются только пороговыми функциями. Правило обучения однослойного персептрона формулируется следующим образом. Пусть имеется множество пар $\{\langle \mathbf{p}_1, \mathbf{d}_1 \rangle, \langle \mathbf{p}_2, \mathbf{d}_2 \rangle, \dots, \langle \mathbf{p}_Q, \mathbf{d}_Q \rangle\}$, называемое обучающей выборкой. Здесь \mathbf{p}_Q - входной сигнал, \mathbf{d}_Q - желаемый отклик персептрона на соответствующее входное воздействие, называемый целью ($q = 1 \dots Q$). Поскольку все нейроны активизируются пороговой функцией, координаты векторов \mathbf{d}_q должны принимать два устойчивых значения (обычно 0 и 1). Первоначально, веса и смещения принимают любые случайные значения, соизмеримые с величинами координат входных векторов из обучающей выборки, или нулевые значения. Процесс обучения состоит в последовательном предъявлении сети векторов \mathbf{p}_q . Вве-

дем для каждого нейрона понятие ошибки $e = d - y$, где d - цель, y - фактический выход нейрона. Правило обучения однослойного персептрона можно записать следующим образом:

$$\Delta W_{S \times R} = (\bar{d}_{S \times 1} - \bar{y}_{S \times 1})(\bar{p})_{1 \times R}^T = \bar{e}_{S \times 1}(\bar{p})_{1 \times R}^T. \quad (2.1)$$

В процесс обучения можно включить также смещения нейронов, принимая их за веса, на вход которых всегда подается 1. Для смещений формулируется следующее правило обучения:

$$\Delta \bar{b}_{S \times 1} = \bar{d}_{S \times 1} - \bar{y}_{S \times 1} = \bar{e}_{S \times 1}. \quad (2.2)$$

Итерацию алгоритма обучения, включающую в себя последовательное предъявление сети всех пар обучающей выборки, принято называть эпохой (epoch). Процесс модификации весов и смещений останавливается на той эпохе, которая не вызвала никаких изменений в настройке сети. Если образцы, предъявляемые однослойному персептрону, линейно разделяемы, то такая процедура обучения гарантирует, что персептрон построит классификатор за конечное число итераций. Для персептронов весьма важной является возможность настройки смещений в нейронах, поскольку при фиксированных смещениях (в частности при $b=0$ для всех нейронов) гиперплоскости классификатора привязываются к определенной точке (к началу координат). Такое ограничение не гарантирует сходимости процедуры обучения, даже если задача классификации линейно разрешима. Это происходит из-за того, что при исключении смещений из рассмотрения, сеть ограничивается классификаторами вида $a_1 p_1 + a_2 p_2 + \dots + a_n p_n = 0$ в базисе $\{p_1, p_2, \dots, p_n\}$, а все множество гиперплоскостей n -мерного пространства описывается уравнениями вида

$$a_1 p_1 + a_2 p_2 + \dots + a_n p_n + a_{n+1} = 0.$$

Каждый нейрон в однослойном персептроне описывает одну гиперплоскость в пространстве входных векторов и своим выходным значением, определяет по какую сторону от этой гиперплоскости находится поданный на вход вектор. Таким образом, при проектировании однослойных персептронов необходимо выбирать количество нейронов, равное количеству гиперплоскостей, достаточному для классификации всего пространства входного воздействия. Кроме того, персептроны не могут решать задачи классификации, если классы не могут быть разделены между собой при помощи линейных функций. Например, из множества всех возможных булевых функций двух переменных, однослойный персептрон не может реализовать "эквивалентность" и "исключающее ИЛИ". Поскольку однослойные персептроны могут распознавать только линейно разделяемые образы, они используются только в простейших случаях классификации. Для более сложных случаев обычно применяются другие типы нейронных сетей - адаптивные линейные сети и сети, обучаемые методом обратного распространения ошибки.

3. СЕТИ АДАПТИВНЫХ ЛИНЕЙНЫХ ЭЛЕМЕНТОВ

Адаптивными линейными элементами (ADALINE = ADAPtive LINear Element) называются нейроны с линейной функцией активации [2]. Однослойные сети прямого распространения, состоящие из нескольких адаптивных линейных элементов, иногда объединяют под термином MADALINE (Multi ADALINE).

Сети ADALINE, как и перцептроны, способны решать линейно разделяемые задачи классификации образов. Однако диапазон выходных значений нейронов в сетях ADALINE не ограничен только двумя выходными значениями. Это свойство позволяет обучать такие сети методом наименьших среднеквадратических отклонений, известным в литературе как алгоритм LMS (Least Mean Square) [2, 5], называемым также алгоритмом Уидроу-Хоффа или дельта-правилом [6]. Этот метод строит из сети более точный линейный классификатор, чем правило обучения однослойного перцептрона. Метод LMS заключается в минимизации среднеквадратической ошибки сети. Функция ошибки ($mse = \text{Mean Square Error}$) выражается формулой

$$mse = \frac{1}{Q} \sum_{k=1}^Q e(k)^2 = \frac{1}{Q} \sum_{k=1}^Q (d(k) - y(k))^2. \quad (3.1)$$

Доказано, что функция ошибки в сетях ADALINE с любой обучающей выборкой представляет собой R -мерную параболу, где R - число входных линий сети. Поэтому поверхность ошибки имеет не более одного минимума. Изменение весов и смещений сети согласно методу LMS выражается следующими формулами:

$$W(k+1)_{S \times R} = W(k)_{S \times R} + 2\eta \bar{e}(k)_{S \times 1} \bar{p}^T(k)_{1 \times R}; \quad (3.2)$$

$$b(k+1)_{S \times 1} = W(k)_{S \times 1} + 2\eta \bar{e}(k)_{S \times 1}.$$

Здесь η - скорость обучения. Чем больше η , тем быстрее происходит обучение, однако слишком большое значение ведет к нестабильности алгоритма. Несмотря на то, что алгоритм LMS работает только для однослойных линейных сетей, это его свойство не снижает универсальности, поскольку добавление других линейных слоев к сетям ADALINE не улучшает характеристик сетей и, следовательно, бесполезно.

Адаптивные линейные сети дают хорошие результаты, только в случае, если зависимость между входными и выходными значениями сети линейна или линейно аппроксимируема. Если линейные сети не способны решать задачу классификации, аппроксимации или предсказания, поставленную перед ними, то для решения этой задачи могут быть использованы нелинейные сети прямого распространения, обучаемые методом обратного распространения ошибки.

4. ОБРАТНОЕ РАСПРОСТРАНЕНИЕ ОШИБКИ

Метод обратного распространения ошибки (ОРО, Backpropagation) был разработан как обобщение метода Уидроу-Хоффа для многослойных сетей с нелинейными дифференцируемыми функциями активации. В современном виде метод был впервые формализован в [7].

4.1. Базовый алгоритм ОРО [3, 8, 9]

Принцип ОРО основан на настройке сети таким образом, чтобы минимизировать среднеквадратическую ошибку на обучающей выборке. Для этого используется понятие градиента функции ошибки. Сети последовательно представляются входные вектора из обучающей выборки, и, начиная с последнего слоя, вычисляются градиенты функции ошибки для каждого нейрона. Веса и смещения изменяются в направлении, противоположном градиенту. Поэтому базовый алгоритм ОРО еще часто называют методом градиентного спуска (gradient descent).

Приведем базовый алгоритм ОРО в виде последовательности шагов:

1. Все веса и смещения инициализируются маленькими ненулевыми величинами обычно из интервала от -1 до +1.
2. Представление входного вектора \mathbf{p} и желаемого выходного вектора \mathbf{t} .
3. Расчет фактического выхода сети \mathbf{a} при подаче на вход вектора \mathbf{p} .
4. Настройка весов и смещений по формуле:

$$w_{ij}^{(k)}(t+1) = w_{ij}^{(k)}(t) + \eta \delta_i^{(k)} y_j^{(k-1)}. \quad (4.1)$$

Здесь $\delta_i^{(k)}$ - вычисляется для выходного слоя по формуле:

$$\delta_i^{(K+1)} = f'(y_i^{(K+1)})(d_i^{(K+1)} - y_i^{(K+1)}), \quad (4.2)$$

а для скрытых слоев по рекуррентной формуле:

$$\delta_i^{(k)} = f'(y_i^{(k)}) \sum_j \delta_j^{(k+1)} w_{ji}^{(k+1)} \quad (k = 1, \dots, K). \quad (4.3)$$

Здесь η - скорость обучения (небольшая положительная постоянная величина, обычно берущаяся из интервала от 0 до 1), $f'(y_i^{(k)})$ - производная активационной функции i -го нейрона, находящегося на k -м слое. Если используются нейроны с сигмоидной функцией активации, то для них

$$f'(y_i^{(k)}) = y_i^{(k)}(1 - y_i^{(k)}). \quad (4.4)$$

Процедура настройки весов происходит в направлении от выходного слоя к первому, отсюда и название метода.

5. Если не удовлетворено условие окончания обучения, то переход к шагу 2.

Существует множество различных модификаций базово-

го алгоритма ОРО (градиентного спуска). В [10] они условно разделены на пять групп:

- 1) алгоритмические изменения (ОРО с моментом [2, 11], Quickprop [10, 11], методы второго порядка [2, 4, 10], адаптация скорости обучения [10, 12, 13], RPROP [10, 14]);
- 2) эвристики, применяемые к алгоритму (эвристические методы инициализации весов [2], оценка оптимальной скорости обучения);
- 3) регулировки (падение весов [11, 15]);
- 4) модификации сетей (нестандартные нейроны и активационные функции, рекуррентные узлы в сетях прямого распространения);
- 5) сети с динамически изменяющейся структурой (каскадная корреляция [6]).

Рассмотрим некоторые наиболее популярные модификации алгоритма ОРО.

4.2. Пакетная модификация [2, 11]

В стандартном (пошаговом) алгоритме ОРО веса и смещения модифицируются после каждого предъявления обучающего вектора. При использовании пакетной модификации, веса и смещения обновляются только после окончания эпох. То есть после предъявления очередного обучающего вектора изменения весов вычисляются, но они не влияют на конфигурацию сети, а суммируются. Только после предъявления последнего вектора, полученные суммы прибавляются к текущим весам. Пакетная модификация не гарантирует ускорения сходимости метода, но избавляет процесс обучения от его зависимости от порядка предъявления обучающих пар. Если алгоритм ОРО реализуется в пошаговом режиме, то желательно на каждой эпохе случайным образом менять порядок предъявления обучающих векторов.

4.3. Градиентный спуск с моментом [2, 10, 11]

Стандартный метод ОРО часто сходится очень медленно, двигаясь вдоль плоских участков поверхности ошибки. Для ускорения сходимости алгоритма предлагается ввести понятие взвешенного момента. Выражение для изменения весов в этой модификации будет следующим:

$$w_{ij}(t+1) = w_{ij}(t) + \Delta w_{ij}(t) + \alpha \Delta w_{ij}(t-1). \quad (4.5)$$

Здесь t - номер очередной итерации обучения, w_{ij} - рассматриваемый вес (или смещение), Δw_{ij} - изменение веса (смещения), предлагаемое стандартным алгоритмом ОРО, α - константа момента. Значение константы момента должно лежать в диапазоне от 0 до 1 (обычно ? выбирается равным 0,9).

4.4. Разрушение весов [16]

При обучении часто возникает проблема недопустимо низкого уровня обобщающей способности нейронных сетей. То есть сеть может давать на выходе низкий уровень ошибок для данных обучающей выборки, но для других

данных значения ошибок будут недопустимо высоки. Такая ситуация обычно вызывается двумя причинами: переобученностью или избыточным количеством нейронов скрытых слоев. Переобученность сети (overfitting) возникает из-за стремления обучающего алгоритма максимально приблизить значение среднеквадратической ошибки к нулю на обучающей выборке. При этом сеть может потерять способность получать правильные выходные данные для данных, не входящих в обучающую выборку. Если низкая обобщающая способность возникает из-за избыточного количества нейронов, то достаточно сложно точно оценить от какого числа нейронов можно отказаться, чтобы сеть сохранила в себе способность решать требуемую задачу.

Универсальный подход к решению проблемы низкой обобщающей способности, называемый разрушением весов (weight decay) был предложен в [16]. Метод заключается в вычитании из каждого весового коэффициента незначительной части перед модификацией веса по алгоритму ОРО, согласно формуле (4.6).

$$w_{ij}(t) = (1-h)w_{ij}(t-1) - \eta \delta_i y_i. \quad (4.6)$$

Здесь h - небольшая константа. В [10] утверждается, что даже значение $h = 10^{-5}$, как правило, позволяет существенно улучшить обобщающую способность сети. В [15] используются значения h , равные $1 \cdot 10^{-3}$ и $5 \cdot 10^{-3}$. В [11] метод разрушения весов используется для устранения переполнения разрядной сетки, иногда возникающего при обучении алгоритмом Quickprop.

4.5. Эластичная модификация (RPROP) [14]

Эластичный алгоритм ОРО (Resilient backpropagation = RPROP) был предложен Ридмиллером и Брауном в 1993 году. В отличие от других алгоритмов, основанных на градиентном спуске, RPROP не использует величину градиента, а только его знак. Каждому весу сопоставляется так называемая величина обновления (update value), которая инициализируется небольшим числом. После каждой итерации алгоритма значения этих параметров увеличиваются на коэффициент увеличения, если знак текущего градиента совпадает со знаком предыдущего, или уменьшаются на коэффициент уменьшения, если эти знаки не совпадают. Коэффициент увеличения принято брать равным 1,2, а коэффициент уменьшения - 0,5. Если знак (направление) градиента отрицателен, то величина обновления прибавляется к весу, а если знак положителен, то величина обновления вычитается из веса.

Алгоритм RPROP часто дает более быструю сходимость, чем большинство других алгоритмов, и поэтому он является одной из наиболее популярных модификаций ОРО из известных в настоящее время. Однако в [11] отмечается, что для некоторых особых случаев RPROP не может найти оптимальную настройку сети, попадая в ложные локальные минимумы на поверхности ошибки.

4.6. Методы второго порядка [2, 10]

Базовый алгоритм ОРО двигается к минимуму функции ошибки в направлении наискорейшего спуска. Такой принцип не всегда означает максимально быструю сходимость поиска. Методы второго порядка пытаются найти более короткий путь к минимуму, более тщательно исследуя кривизну поверхности ошибки. Все методы второго порядка основаны на методе Ньютона, использующем первые три члена разложения функции ошибки в ряд Тейлора. Правило для изменения весовых коэффициентов в методе Ньютона следующее:

$$w(t+1) = w(t) - A^{-1}(t)\eta \frac{\partial E}{\partial w}. \quad (4.7)$$

Здесь $A(t)$ - гессиан (матрица вторых производных) функции ошибки по текущим значениям весов и смещений. В теории метод Ньютона должен ускорять сходимость поиска в несколько раз, но вычисление гессиана настолько трудоемкая операция, что практические реализации метода Ньютона находят минимум ошибки намного медленнее стандартного метода ОРО. В силу своей трудоемкости, метод Ньютона на практике не применяется, но существует множество эффективных алгоритмов, использующих основную идею метода Ньютона.

Одним из них является алгоритм Левенберга-Маркардта (Levenberg-Marquardt), который использует нетрудоемкую аппроксимацию гессиана для метода Ньютона. В этом алгоритме, гессиан оценивается по формуле $A^{-1} \approx J^T \cdot J$, а градиент рассчитывается как $\mathbf{g} = J^T \cdot \mathbf{e}$, где J - якобиан (матрица первых производных) функции ошибки, \mathbf{e} - вектор ошибки сети. Так как активационные функции в нейронных сетях прямого распространения имеют простые производные, якобиан рассчитывается намного быстрее матрицы Гесса, тем более он и так вычисляется в стандартном методе ОРО.

На методе Ньютона основаны также так называемые алгоритмы сопряженных градиентов (conjugate gradient algorithms). В алгоритмах сопряженных градиентов направления изменения весов определяются по следующей рекуррентной формуле:

$$V(t) = -G(t) + z(t, t-1)V(t-1), \quad (4.8)$$

где $G = \frac{\partial E}{\partial w}$ - градиент функции ошибки, параметр $z(t, t-1)$ определяется, в зависимости от выбранной конкретной модификации алгоритма, по формуле (4.9) (алгоритм Флетчера-Ривса) или (4.10) (алгоритм Полака-Рибьера).

$$z(t, t-1) = \frac{|G(t)|^2}{|G(t-1)|^2}, \quad (4.9)$$

$$z(t, t-1) = \frac{|G(t)|^2 - |G(t) \cdot G(t-1)|}{|G(t-1)|^2}. \quad (4.10)$$

После выбора направления поиска, находится величина изменения веса по формуле $\Delta w = \lambda V(t)$, где коэффициент λ вычисляется, исходя из условия минимума функции $E(w(t-1) + \lambda V)$.

4.7. Алгоритм Quickprop [10, 11]

Алгоритм Quickprop, предложенный Фальманом [11] в 1988 году, является одной из наиболее популярных модификаций обратного распространения ошибки. Quickprop сочетает в себе идеи алгоритмов второго порядка с простотой реализации и низкой трудоемкостью. Правило изменения весов в методе Quickprop сводится к следующему уравнению.

$$\Delta w(t) = -\eta \frac{\partial E}{\partial w} + \frac{\frac{\partial E}{\partial w}(t)}{\frac{\partial E}{\partial w}(t-1) - \frac{\partial E}{\partial w}(t)} \Delta w(t-1). \quad (4.11)$$

Фальман также вводит в алгоритм коэффициент максимального роста (μ), который используется как верхний предел изменения весов для предотвращения колебаний сети в окрестности минимума ошибки. Без использования этого коэффициента весовые коэффициенты могут постоянно расти, не увеличивая общую производительность сети. Для большинства практических задач значение коэффициента максимального роста рекомендуется брать из интервала от 1,75 до 2,25. Полностью алгоритм Quickprop представлен в [10] в виде псевдокода.

4.8. Модификации ОРО с переменной скоростью обучения [2, 10]

В классическом методе ОРО скорость обучения (η) инициализируется и в течение всего процесса обучения остается постоянной. Существует несколько модификаций ОРО, которые подразумевают изменение параметра η во время обучения.

В 1981 году Саттоном [12] был предложен метод адаптивного изменения скорости обучения, называемый методом Delta-Delta. Согласно этому методу каждому весу сопоставляется своя скорость обучения, которая изменяется по следующей формуле

$$\Delta \eta_i(t) = \gamma \frac{\partial E(t)}{\partial w_i(t)} \frac{\partial E(t-1)}{\partial w_i(t-1)}. \quad (4.12)$$

Здесь γ - небольшой положительный коэффициент, являющийся аналогом параметра η в алгоритмах ОРО с постоянной скоростью обучения.

В 1988 Джейкобс [13] указал на недостатки алгоритма

Саттона и модифицировал его, предложив следующую формулу для изменения скорости обучения:

$$\Delta\eta_i(t) = \begin{cases} k, & \text{если } \bar{\delta}(t-1)\delta(t) > 0 \\ -\Phi\eta(t), & \text{если } \bar{\delta}(t-1)\delta(t) < 0, \\ 0, & \text{если } \bar{\delta}(t-1)\delta(t) = 0 \end{cases} \quad (4.13)$$

где $\delta(t) = \frac{\partial E(t)}{\partial w(t)}$, $\bar{\delta}(t) = (1 - \theta)\delta(t) + \theta\bar{\delta}(t-1)$. Здесь

$\delta(t)$ - частная производная функции ошибки по весу, $\bar{\delta}(t)$ - экспоненциальное среднее текущей и ранее вычисленных производных. Джейкобс использовал для коэффициента линейного роста k значения от 0,35 (для сложных задач) до 5,0 (для простых задач); коэффициент экспоненциального падения Φ от 0,1 до 0,35 и θ в районе 0,7. Алгоритм Джейкобса получил название Delta-Bar-Delta. Его достоинством является то, что в отдалении от минимума функции ошибки скорость обучения растет линейно, ускоряя сходимость, а при попадании в окрестность минимума, она падает экспоненциально, повышая точность поиска.

5. РАДИАЛЬНО-БАЗИСНЫЕ СЕТИ

Радиально-базисные сети относятся к сетям прямого распространения, но они не обучаются методом ОРО. Свое название эти сети получили из-за использования нейронов с радиально-базисными функциями. На входе нейрона вычисляется расстояние между векторами \mathbf{w} и \mathbf{p} , затем полученное значение умножается на смещение нейрона и от результата берется радиально-базисная функция (гауссиана). Радиально-базисная функция выражается формулой $H(x) = \exp(-x^2)$. Таким образом, нейроном реализовывается функция $a = H(\|\mathbf{w} - \mathbf{p}\|b)$. Поскольку гауссиана достигает максимума при значении аргумента, равном нулю, рассмотренный нейрон выдает значение, показывающее насколько близко расположен входной вектор к весовому вектору.

Основная разновидность радиально-базисных сетей представляет собой двухслойную сеть прямого распространения, первый слой которой состоит из радиально-базисных нейронов, а второй слой представляет собой линейную сеть (гл. 4). В отличие от сетей, обучаемых методом ОРО, радиально-базисные сети не обучаются, а разрабатываются под конкретную задачу. Например, известен алгоритм разработки радиально-базисной сети, дающей нулевую ошибку на любой обучающей выборке (функция newtbe в [2]). Но такой алгоритм может построить сеть с избыточным числом нейронов. Поэтому, как правило, при разработке радиально-базисных сетей берется максимально допустимое значение ошибки, большее нуля, и находится минимальная конфигурация системы, которая на обучающей выборке не превышает эту ошибку. Кроме основной разновидности радиально-базисных сетей существуют и другие, например сети обоб-

щенной регрессии и вероятностные сети [2, 4].

У радиально-базисных сетей имеется масса достоинств, но и есть существенный недостаток. Радиально-базисные нейроны используют больше вычислений для своего функционирования, поэтому радиально-базисные сети работают медленнее, чем другие сети прямого распространения.

6. САМООРГАНИЗУЮЩИЕСЯ СЕТИ

Самоорганизующиеся сети относятся к сетям с обратными связями. В отличие от всех рассмотренных ранее классов нейронных сетей, у самоорганизующихся сетей нет ярко выраженного разделения на фазу обучения и фазу функционирования. Самоорганизующиеся сети получают вектора входного воздействия, выдают на выходе свою реакцию и, возможно, корректируют свои настройки в зависимости от поступившего входного вектора. Самоорганизующиеся сети включают в себя соревновательные сети и самоорганизующиеся карты признаков (Self-Organizing Feature Map = SOFM) Кохонена.

6.1. Соревновательные сети

Соревновательные сети - это разновидность однослойных самоорганизующихся сетей. На входе сеть вычисляет вектор, содержащий отрицательные расстояния между вектором \mathbf{p} и элементами матрицы \mathbf{W} , затем прибавляет к нему вектор смещений \mathbf{b} . На выход сети, соответствующий максимальному полученному значению подается 1, а на остальные выходы - 0. Если все смещения равны 0, то слой выбирает тот весовой вектор, который находится ближе всех к входному вектору. Это происходит потому, что расстояния между входным вектором и весовыми векторами берутся с противоположными знаками, и поэтому наименьшее расстояние между этими векторами дает наибольшее значение на выходе.

Кохонен [17] сформулировал следующее правило обучения соревновательных нейросетей. Пусть поданный на вход вектор активизировал i -й выход соревновательной сети. Тогда i -ая строка весовой матрицы меняется по формуле:

$${}_iW_{11}(q) = {}_iW_{11}(q-1) + \eta(p(q) - {}_iW_{11}(q-1)). \quad (6.1)$$

В результате увеличивается вероятность того, что рассматриваемый нейрон при предъявлении сети похожего вектора также активизируется. В то же время, снижается вероятность активизации этого же нейрона, если сети будет предъявлен непохожий вектор. Часто правило Кохонена дополняется специальной процедурой, заключающейся в периодическом увеличении смещений нейронов, которые редко выигрывают (или никогда не выигрывают) соревнование с другими нейронами. Необходимость в такой процедуре возникает из-за того, что при инициализации системы некоторые нейроны оказываются заведомо далеко от центров кластеров и, как бы долго ни шел процесс обучения, такие нейроны не активизируются и, следовательно, не участвуют в процессе кластеризации.

Если соревновательная сеть содержит достаточное количество нейронов, то в конечном итоге эта сеть должна выделить кластеры в пространстве входных сигналов, к которым она будет относить поступающие входные вектора.

6.2. Самоорганизующиеся карты признаков (SOFM)

Самоорганизующиеся карты признаков (Self-Organizing Feature Maps = SOFM) Кохонена, как и соревновательные сети обучаются кластеризации векторов по принципу близости их расположения во входном пространстве. Сети SOFM отличаются тем, что в них близкие по расположению нейроны учатся распознавать близкие области. Таким образом, карты признаков обучаются не только распределению входных векторов, но и их топологии. Самоорганизующиеся карты можно представить в виде взвешенного графа, в вершинах которого расположены нейроны, а взвешенные ребра характеризуют расстояния между нейронами. Правило Кохонена в самоорганизующихся картах применяется не только для победившего нейрона, но и для соседних с ним. В некоторых случаях, для соседних нейронов используется меньший коэффициент скорости обучения, чем для победившего нейрона (например, в два раза). Архитектура карт Кохонена похожа на архитектуру соревновательных сетей, за исключением того, что в картах не используются нейронные смещения.

6.3. Соревновательные сети с квантованием обучающего вектора (LVQ)

Квантование обучающего вектора (Learning Vector Quantization = LVQ) - это метод обучения соревновательных сетей "с учителем". Стандартные соревновательные сети обучаются классификации входных векторов "без учителя" в процессе работы, используя в качестве единственного признака функцию расстояния. Если два входных вектора очень похожи, соревновательный слой, вероятно, поместит их в один класс. При этом не существует механизма, позволяющего разработать соревновательный слой, относительно которого можно было бы с уверенностью сказать: попадут ли два произвольно заданных входных вектора в один и тот же класс или нет. При использовании метода квантования обучающего вектора, к соревновательному слою достраивается линейный слой. В этом случае, первый (соревновательный) слой обучается классифицировать входные образы по расстоянию между векторами, а второй (линейный) приводит получающиеся подклассы к классам, определяемым пользователем.

7. РЕКУРРЕНТНЫЕ СЕТИ

Рекуррентные сети являются подклассом нейронных сетей с обратными связями. В отличие от самоорганизующихся сетей, в рекуррентных сетях обратные связи располагаются не между нейронами одного и того же слоя, а идут от выхода слоев к входам. Рассмотрим две разновидности рекуррентных сетей: сети Эльмана и сети Хопфилда.

7.1. Сети Эльмана

Сеть Эльмана представляет собой двухслойную сеть с первым тангенциальным и вторым линейным слоем. Обратная связь проходит от выхода первого слоя на его вход через элемент задержки и матрицу весов. Обратная связь позволяет сети Эльмана распознавать не только пространственные, но и временные образы. Не смотря на то, что сети Эльмана относятся к рекуррентным сетям, их можно обучать слегка модифицированным методом обратного распространения ошибки. Если с помощью сети Эльмана требуется распознать временные образы, то обучающие входные вектора необходимо предъявлять последовательно, как они будут появляться при работе сети. При обучении сети Эльмана, весовые коэффициенты можно корректировать как после предъявления каждого входного вектора, так и после предъявления последовательности входных векторов (в пакетном режиме). В [2] рассмотрен пример практического применения сети Эльмана для адаптивного определения амплитуды произвольного сигнала.

7.2. Сети Хопфилда

Сети Хопфилда представляют собой однослойные рекуррентные сети с линейно-пороговой функцией активации и элементом задержки на обратной связи. Под линейно-пороговой активационной функцией понимается функция вида:

$$f(n) = \begin{cases} n, & \text{если } |n| < 1 \\ \text{sign}(n), & \text{если } |n| \geq 1 \end{cases} \quad (7.1)$$

Принцип функционирования сети состоит в хранении некоторого множества точек устойчивого равновесия (аттракторов) в пространстве выходного сигнала. При предъявлении входного вектора \mathbf{p} , сеть Хопфилда, через некоторое число итераций переходит в устойчивое состояние, соответствующее ближайшему аттрактору для вектора \mathbf{p} . Таким образом, основное применение сети Хопфилда - реализация ассоциативной памяти, способной восстановить весь вектор по вектору с утраченными или зашумленными фрагментами. Сети Хопфилда не обучаются, а проектируются для заданного множества аттракторов. В [2] реализован алгоритм проектирования сети Хопфилда с минимизацией количества ложных аттракторов. К сожалению, не известен алгоритм разработки сетей Хопфилда, гарантирующий полное отсутствие ложных и неустойчивых точек равновесия для любой задачи.

ЗАКЛЮЧЕНИЕ

В результате работы был проведен обзор основных типов современных архитектур нейронных сетей и нейросетевых алгоритмов. Работа проводилась в рамках договора 11711 Таганрогского радиотехнического университета для разработки нейросетевой программы биометрической идентификации пользователя по клавиатурному почерку.

ПЕРЕЧЕНЬ ССЫЛОК

1. Elliot D. L., "A Better Activation Function for Artificial Neural Networks", Institute for System Research: Maryland, 1993 (ftp://ftp.isr.umd.edu/pub/TechReports/1993/TR_93-8.pdf).
2. Demuth H., and M. Beale "Neural network toolbox for use with Matlab. User's guide", MathWorks, Inc., 1997 (<http://www.mathworks.com>).
3. Учебно-методическое пособие для самостоятельной работы по курсам "Системы искусственного интеллекта", "Методы распознавания образов" / Сост. Вишняков Ю. М., Кодачигов В. И., Родзин С. И.- Таганрог: ТРТУ, 1999.-132 с.
4. Hagan M. T., H. B. Demuth, M. H. Beale Neural Networks Design. Boston, MA: PWS Publishing, 1996.
5. Widrow B., and M. E. Hoff, Adaptive switching circuits, 1960 IRE WESCON Convention Record, New York IRE, 1960, pp. 96-104.
6. Fahlman S. E. and C. Lebiere The Cascade-Correlation Learning Architecture, Pittsburgh, PA: Carnegie Mellon University, 1991 (<ftp://ftp.cs.cmu.edu/afs/cs/project/connect/tr/cascor-tr.ps.Z>).
7. Rumelhart D. E., Hinton, G. E., and Williams, R. J. (1986) Learning Internal Representations by Error Propagation// Rumelhart D. E. and McClelland, J. L., Parallel Distributed Processing: Explorations in the Microstructure of Cognition, MIT Press, 1986.
8. Вороновский Г. К., и др. Генетические алгоритмы, искусственные нейронные сети и проблемы виртуальной реальности. - Харьков: ОСНОВА, 1997. - 112 с.
9. Wasserman P. Neurocomputing. Theory and practice, Nostram Reinhold, 1990. (Рус. пер. Ф. Уоссермен. Нейрокомпьютерная техника. М. Мир, 1992).
10. Gibb, J. Back Propagation Family Album, NSW, Australia: Macquarie University, 1996, 72 p. (<ftp://ftp.mpce.mq.edu.au/pub/comp/techreports/96C005.gibb.ps>).
11. Fahlman S. E. An Empirical Study of Learning Speed in Back-Propagation Networks, Pittsburgh, PA: Carnegie Mellon University, 1988 (<ftp://ftp.cs.cmu.edu/afs/cs/project/connect/tr/qp-tr.ps.Z>).
12. Barto A. G. and R. S. Sutton Goal seeking components for adaptive intelligence. Technical Report AFWAL-TR-81-1070, Wright-Patterson Air Force Base, Ohio, 1981.
13. Jacobs R. A. Increased rates of convergence through learning rate adaptation// Neural Networks , N1, 1988, pp. 295-307.
14. Riedmiller M., and H. Braun, A direct adaptive method for faster backpropagation learning: The RPROP algorithm// Proceedings of the IEEE International Conference on Neural Networks, 1993 (<ftp://i11s16.ira.uka.de/pub/neuro/papers/riedml.icnn93.ps.Z>).
15. Tveter D. R., Backpropagator's Review, 2000 (<http://www.dontveter.com/bpr/bpr.html>).
16. Krogh A., and J. A. Hertz. A simple weight decay can improve generalization. Technical report. Niels Bohr Institute / Nordita, Blegdamsvej 17, DK-2100 Copenhagen, Denmark, 1990.
17. Kohonen T., Self-Organization and Associative Memory, 2nd Edition, Berlin: Springer-Verlag, 1987.

УДК 519.873 + 519.7.24/25

ЛОГИКО-АЛГЕБРАИЧЕСКИЕ ОСНОВЫ МАТЕМАТИЧЕСКОЙ ТЕОРИИ НАДЕЖНОСТИ

А.И.Волгин

На базе логико-алгебраического аппарата аддитивно-мультипликативной (АМ) алгебры разработаны логические основы теории надежности изделий. На основе изоморфизма логической АМ-модели $P = P_{AM}(p_1, p_2, \dots, p_n)$ и топологической модели (логической схемы надежности - ЛСН) $ЛСН = P_{AM}(E_1, E_2, \dots, E_n)$ построена единая совмещенная структурно-параметрическая модель надежности, где $p_i \in [0, 1]$ - вероятность безотказной работы элементов E_i ЛСН со временем безотказной работы t_i .

The logical bases of production reliability were developed on the basis of the additive-multiplicative algebra.

ВВЕДЕНИЕ

Любой технический объект может быть представлен триадой (S, P, Q) , где $S = S(E_1, E_2, \dots, E_n)$ - структура объекта, содержащая множество $\{E_1, E_2, \dots, E_n\}$ взаимодействующих элементов E_i , каждому из которых ставится в соответствие параметр $p_i = p(E_i)$, Q - критерий оптимизации [1]. Задача синтеза заключается в построении структуры объекта, задача анализа (при заданной структуре) заключается в определении результирующего параметра $P = p[S(E_1, E_2, \dots, E_n)]$ структуры объекта. Оптимизация сводится к обеспечению экстремального значения

функционала $Q = Q(S, P, D) \rightarrow \text{ext}$ (максимум или минимум), где D - условия оптимизации.

Если ввести символы топологических операций параллельного \oplus и последовательного \cdot соединения элементов топологической модели (схемы) объекта, то для классов параллельно-последовательных схем (ПП-схемы) топологическая модель может быть представлена в аналитической форме записи $S = S(E_1, E_2, \dots, E_n)$. В частности, на рис.1 представлены логические (структурные) схемы надежности (ЛСН) при последовательном и параллельном (горячее резервирование) соединении элементов в ЛСН, аналитическое представление которых имеет вид:

$$S = E_1 \cdot E_2 \cdot \dots \cdot E_n, S = E_1 \oplus E_2 \oplus \dots \oplus E_n. \quad (1)$$

Математическая модель объекта записывается в базе операций арифметической алгебры (А-алгебра).

В настоящее время для описания объекта проектирования повсеместно используется раздельное представление топологической и математической моделей. Топологическая модель дает наглядно-образное представление структуры объекта и служит в качестве исходных данных для построения математической модели, которая необходима для расчета результирующего (выходного) параметра объекта $P = P_A(p_1, p_2, \dots, p_n)$.

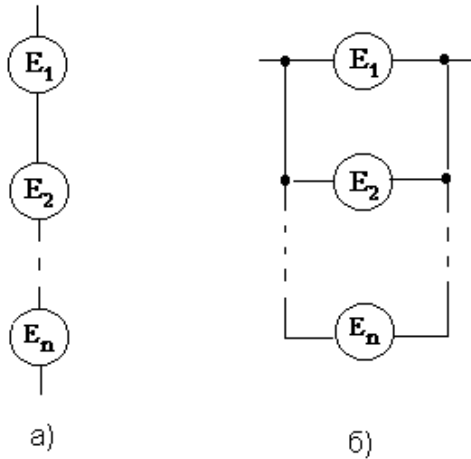


Рисунок 1

В силу межполушарной асимметрии мозга человека [2] наличие двух типов моделей с различными системами базовых операций для описания объекта кажется нам естественным и очевидным.

В [1] сформулирована общая постановка задачи построения совмещенной структурно-параметрической модели технических объектов. А именно. Необходимым и достаточным условием существования совмещенной структурной (топологической) и параметрической (математической) модели объекта является выполнение свойства изоморфизма [3]:

$$p[S(E_1, E_2, \dots, E_n)] = S[p(E_1), p(E_2), \dots, p(E_n)], \quad (2)$$

при представлении (2) в инфиксной форме записи P -операций, где $p_i = p(E_i)$.

Поставленная задача в каждом конкретном случае связана с переходом от базиса A -операций к подходящему базису P -операций математических проблемно-ориентированных (предметных) логик или специальных алгебр. При этом A -операции арифметической алгебры являются для предметных логик метаоперациями.

В общем виде поставленная задача не поддается формализации, но для отдельных классов задач (объектов) имеются решения общетеоретической значимости.

Настоящая работа посвящена дальнейшему развитию математической теории надежности [4,5] на основе логико-алгебраического аппарата мультипликативно-аддитивной алгебры (AM -алгебра) [6-9], которая на интервале $[0,1]$ является логикой исчисления надежности [10-12].

ОСНОВНЫЕ ОПРЕДЕЛЕНИЯ, ПОНЯТИЯ И ИЗОМОРФИЗМЫ

Под надежностью понимается способность изделия (радиоэлектронных средств, устройств, систем и др.) сохранять декларируемое качество на заданном интервале времени в нормируемых условиях эксплуатации. Качество изделия - это совокупность его свойств, определяющих

пригодность использования изделия по назначению.

Расчет надежности заключается в определении заданного показателя (характеристики) надежности изделия по аналогичным показателям надежности элементов, входящих в состав изделия. Показателями надежности являются вероятность безотказной работы p , вероятность отказа $\bar{p} = 1 - p$, время безотказной работы t_i , интенсивность отказов λ и др., связанные между собой известными зависимостями [4,5].

При традиционном подходе путем надежностного анализа исходной электрической схемы (функциональной, структурной или блок-схемы) изделия строится ее логическая схема надежности (ЛСН) [13]. Далее в базе операций A -алгебры (A -сложение, A -умножение) по ЛСН строится математическая модель надежности $P = P_A(p_1, p_2, \dots, p_n)$, по которой вычисляются числовые значения показателей надежности. В частности, при независимых отказах вероятность безотказной работы (ВБР) изделия без резервирования (рис.1а) и с горячим резервированием (рис.1б) определяются соответственно выражениями [4,5]:

$$P = P_A(p_1, p_2, \dots, p_n) = p_1 \cdot p_2 \cdot \dots \cdot p_n, \quad (3.1)$$

$$P = P_A(p_1, p_2, \dots, p_n) = 1 - (1 - p_1)(1 - p_2) \dots (1 - p_n). \quad (3.2)$$

При таком подходе, согласно (1) и (3), топологические (рис.1) и математические (3) модели не совпадают. При этом по представленной формуле для ВБР сложных ПП-схем невозможно восстановить ЛСН изделия, так как условие изоморфизма (2) не выполняется.

В [14] показано, что для получения совмещенной (структурно-параметрической) математической модели по ВБР ее необходимо строить в базе операций AM -логики, в которой элементарными операциями являются аддитивная инверсия $\bar{p}_i = 1 - p_i$, AM -конъюнкция (арифметическое умножение) и AM -дизъюнкция (вероятностное сложение):

$$M(p_1, p_2) = p_1 \cdot p_2 = 0,5[p_1 + p_2 - (p_1 + p_2 - 2p_1p_2)] \quad (4.1)$$

$$W(p_1, p_2) = p_1 \oplus p_2 = 0,5[p_1 + p_2 + (p_1 + p_2 - 2p_1p_2)], \quad (4.2)$$

где \oplus есть символ операции вероятностного сложения

$$p_1 \oplus p_2 = p_1 + p_2 - p_1p_2.$$

Здесь и в дальнейшем изложения для наглядности используется инфиксная $p_1 \oplus p_2$ и функциональная $W(p_1, p_2)$ форма записи операций и функций AM -алгебры. В классе ПП-схем надежности имеет место тождество:

$$P = P_A(p_1, p_2, \dots, p_n) = P_{AM}(p_1, p_2, \dots, p_n), \quad (5)$$

в которой AM -модель $P_{AM}(p_1, p_2, \dots, p_n)$ является функциональной сверткой A -модели $P_A(p_1, p_2, \dots, p_n)$, представленной в базе арифметической алгебры.

Расширение класса AM -функций, представленных в базе A -операций и AM -операций, осуществляется соответственно через операции суперпозиции и композиции. Тогда выражения (3), записанные в базе AM -операций, приводятся к виду:

$$M(p_1, p_2, \dots, p_n) = p_1 \cdot p_2 \cdot \dots \cdot p_n, \\ W(p_1, p_2, \dots, p_n) = p_1 \oplus p_2 \oplus \dots \oplus p_n. \quad (6)$$

Например, при $n=3$ и $n=4$ из (3) соответственно получим:

$$p_1 \oplus p_2 \oplus p_3 = p_1 + p_2 + p_3 - p_1 p_2 - p_1 p_3 - p_2 p_3 + p_1 p_2 p_3, \\ p_1 \oplus p_2 \oplus p_3 \oplus p_4 = p_1 + p_2 + p_3 + p_4 - p_1 p_2 - p_1 p_3 - p_1 p_4 - p_2 p_3 - p_2 p_4 - p_3 p_4 + p_1 p_2 p_3 + p_1 p_2 p_4 + p_1 p_3 p_4 + p_2 p_3 p_4 - p_1 p_2 p_3 p_4.$$

При произвольном n n -местная AM -дизъюнкция определяется полилинейной функцией вида

$$P = p_1 \oplus \dots \oplus p_n = \sum_{i=1}^n p_i - \sum_{i_1, i_2}^{N_{n-1}} p_{i_1} p_{i_2} + \sum_{i_1, i_2, i_3}^{N_{n-2}} p_{i_1} p_{i_2} p_{i_3} - \dots \quad (7) \\ \dots + (-1)^{n-2} \sum p_{i_1} \dots p_{i_{n-1}} + (-1)^{n-1} p_1 p_2 \dots p_n,$$

где N_j есть количество сочетаний из n по j .

Из (1) и (6) следует, что символы "." и " \oplus " являются обозначениями операций как последовательного и параллельного соединений элементов ПП-схем надежности, так и операций AM -конъюнкций и AM -дизъюнкций (4). Указанный изоморфизм позволяет из AM -моделей $P = P_{AM}(p_1, \dots, p_n)$ заменой в них переменных p_i на обозначения элементов E_i ЛСН получить аналитическую запись топологической модели и обратно:

$$P_{AM}(p_1, \dots, p_n) \leftrightarrow P_{AM}(E_1, E_2, \dots, E_n) = S(E_1, E_2, \dots, E_n), \quad (8)$$

Покажем изоморфизм математических моделей по вероятностным p_i и временным t_i параметрам относительно отображения $(p_i, \cdot, \oplus) \leftrightarrow (t_i, \wedge, \vee)$, где t_i - время безотказной работы, $\wedge = \min$ и $\vee = \max$ - базовые операции непрерывной логики выделения (селекции) соответственно минимальной и максимальной континуальных переменных [15,16].

При замене в выражениях (6) переменных p_i и символов ".", " \oplus " на t_i и символы операций непрерывной логики (НЛ) " \wedge ", " \vee " получим:

$$T = T(t_1, t_2, \dots, t_n) = t_1 \wedge t_2 \wedge \dots \wedge t_n = \min(t_1, t_2, \dots, t_n), \quad (9.1)$$

$$T = T(t_1, t_2, \dots, t_n) = t_1 \vee t_2 \vee \dots \vee t_n = \max(t_1, t_2, \dots, t_n), \quad (9.2)$$

где T - результирующее время безотказной работы изделий без резервирования (рис.1а) и с горячим резервированием (рис.1б).

В (9) первое и второе равенства представлены соответственно в инфиксной и функциональной формах записи.

Если в соответствии с отображением $(p_i, -, +, \cdot) \leftrightarrow (w_i, /, \cdot, +)$ в базовых операциях AM -алгебры $\bar{p}_i = 1 - p_i$, $p_1 \cdot p_2$, $p_1 \oplus p_2 = p_1 + p_2 - p_1 p_2$, представленных в базе A -операций, операции вычитания "-", сложения "+" и умножения "." заменить соответственно на дуальные A -операции деления "/", умножения "." и сложения "+", то приходим к базовым операциям логики исчисления иммитансов (W -логика) [18-20]: $\bar{w}_i = 1/w_i$ (мультипликативная инверсия),

$$\Lambda(w_1, w_2) = w_1 \parallel w_2 = (w_1^{-1} + w_2^{-1})^{-1} = w_1 w_2 / (w_1 + w_2), \\ V(w_1, w_2) = w_1 + w_2 \quad (10)$$

(обозначения p_i , M и W заменены соответственно на w_i , Λ и V).

Здесь переменные w_i являются иммитансами (сопротивлениями или проводимостями) элементов ПП-схем много-элементных электрических двухполюсных цепей, символом конкатенации "||" (от англ. сочленение) и сложения "+" обозначены соответственно параллельное и последовательное соединения электрической схемы. Одновременно символы "||" и "+" являются обозначениями базовых операций W -логики (10). Следовательно, ПП-схемы много-элементных электрических двухполюсников также описываются совмещенными структурно-параметрическими моделями.

Установленный здесь изоморфизм позволяет при совпадающих конфигурациях ЛСН и исходной для ее ПП-схемы электрической цепи от логической схемы надежности вернуться к порождающей ее схеме.

СВОЙСТВА И ЗАКОНЫ ЛОГИКИ ИСЧИСЛЕНИЯ НАДЕЖНОСТИ

Базовыми операциями в AM -алгебре являются аддитивная инверсия, A -умножение и вероятностное сложение (4), в которых в общем случае переменные p_i являются действительными или комплексными числами. В логике исчисления надежности (P -логика) переменные p_i определены на единичном интервале [0,1].

Все законы и свойства AM -алгебры выполняются и для P -логики: ассоциативность

$$p_1 \oplus (p_2 \oplus p_3) = (p_1 \oplus p_2) \oplus p_3,$$

$$p_1 \cdot (p_2 \cdot p_3) = (p_1 \cdot p_2) \cdot p_3; \quad (11)$$

- коммутативность

$$p_1 \oplus p_2 = p_2 \oplus p_1, p_1 \cdot p_2 = p_2 \cdot p_1; \quad (12)$$

- инверсно-идемпотентная обратимость

$$M(p, p) = W(p, -p); \quad (13)$$

- преобразования (законы) де Моргана

$$\overline{p_1 \oplus p_2 \oplus \dots \oplus p_n} = \bar{p}_1 \cdot \bar{p}_2 \cdot \dots \cdot \bar{p}_n, \\ \overline{p_1 \cdot p_2 \cdot \dots \cdot p_n} = \bar{p}_1 \oplus \bar{p}_2 \oplus \dots \oplus \bar{p}_n; \quad (14)$$

- ноль-доминантность для AM -конъюнкций (свойство конъюнктивности)

$$M(p_1, \dots, p_i, \dots, p_n) = 0 \text{ при } p_i = 0, \quad (15.1)$$

т.е. если хотя бы одна переменная равна нулю, то $M(p_1, \dots, p_n) = 0$;

- единичная доминантность для AM -дизъюнкций (свойство дизъюнктивности)

$$W(p_1, \dots, p_i, \dots, p_n) = 1 \text{ при } p_i = 1, \quad (15.2)$$

т. е. если хотя бы одна переменная равна единице, то $W(p_1, \dots, p_n) = 1$;

- свойство согласованности AM -операций с операцией A -сложения

$$M(p_1, p_2) + W(p_1, p_2) = p_1 + p_2 \text{ и др.} \quad (16)$$

AM -функции совместимы с обычными операциями A -алгебры. В частности, имеют место тождества:

$$M(p_1, p_2) + W(\bar{p}_1, \bar{p}_2) = 1,$$

$$W(p_1, p_2) = p_1 \oplus p_2 = p_1 + p_2 \bar{p}_1 = p_2 + p_1 \bar{p}_2,$$

$$W(p_1, \dots, p_n) = W(p_1, \dots, p_{i-1}, p_{i+1}, \dots, p_n) + \\ + p_i \overline{W(p_1, \dots, p_{i-1}, p_{i+1}, \dots, p_n)},$$

$$W(p_1, \dots, p_n, p_{n+1}) = W(p_1, \dots, p_n) + p_{n+1} \overline{W(p_1, \dots, p_n)} \text{ и др.}$$

Любая AM -дизъюнкция может быть разложена по заданному аргументу

$$W(p_1, \dots, p_n) = p_i W(p_1, \dots, p_{i-1}, 1, p_{i+1}, \dots, p_n) + \\ + (1-p_i) W(p_1, \dots, p_{i-1}, 0, p_{i+1}, \dots, p_n). \quad (17)$$

С учетом свойства единичной доминантности (15.2) выражение (17) приводится к виду

$$W(p_1, \dots, p_n) = p_i + \bar{p}_i (p_1 \oplus p_2 \oplus \dots \oplus p_{i-1} \oplus p_{i+1} \oplus \dots \oplus p_n). \quad (18)$$

Начиная процедуру разложения с $i=1$, получим

$$p_1 \oplus \dots \oplus p_n = p_1 + \bar{p}_1 p_2 + \bar{p}_1 \bar{p}_2 p_3 + \dots + \\ + \bar{p}_1 + \dots + \bar{p}_{n-1} p_n + \bar{p}_1 \bar{p}_2 \dots \bar{p}_n. \quad (19)$$

В AM -алгебре имеет место тождество

$$W(1 + \delta_1, \dots, 1 + \delta_n) = 1 + (-1)^{n-1} M(\delta_1, \dots, \delta_n). \quad (20)$$

Если в (6) часть переменных p_1, \dots, p_{i_m} ($m < n$) равны соответственно единице и нулю, то размерность AM -конъюнкций и AM -дизъюнкций уменьшается на m единиц при сохранении их структуры (свойства замкнутости).

При $p_i \in [0, 1]$ дополнительно имеют место следующие свойства:

вложенность

$$p_1 \oplus p_2 \oplus \dots \oplus p_n \in [0, 1], p_1 \cdot p_2 \cdot \dots \cdot p_n \in [0, 1]; \quad (21)$$

субдистрибутивность и супрадистрибутивность [17]

$$p_1 \cdot (p_2 \oplus p_3) \leq p_1 \cdot p_2 \oplus p_1 \cdot p_3, \\ p_1 \oplus p_2 \cdot p_3 \geq (p_1 \oplus p_2) \cdot (p_1 \oplus p_3) \quad (22)$$

концентрирования и реконцентрирования

$$M(p, p) = p \cdot p < p, W(p, p) = p(1 + \bar{p}) > p \text{ и др.} \quad (23)$$

При $p_i \in \{0, 1\}$ AM -алгебра как и непрерывная логика вырождается в алгебру логики Дж.Буля.

Все P -функции ($p_i \in [0, 1]$) имеют свою надежностную интерпретацию. В частности, (21) отражает свойство недо-стижимости абсолютной надежности ($P=1$) через резервирование при $p_i \neq 1$ и конечном n . Свойство конъюнктивности (16) указывает на тот факт, что при выходе из строя хотя бы одного элемента $p_i = 0$ в изделиях с последовательной ЛСН оно становится не работоспособным и пр.

Множество функций, порождаемых операциями композиции, совместно с операциями аддитивной инверсии, арифметического умножения и вероятностного сложения образуют AM -алгебру, которая при $p_i \in [0, 1]$ является логикой исчисления надежности.

Законы и свойства AM -алгебры через отображение $(p_i, -, +, \cdot) \leftrightarrow (w_i, /, \cdot, +)$ имеют аналоги в логике исчисления иммитансов: ассоциативность, коммутативность, преобразования де Моргана

$$\overline{w_1 \parallel w_2 \parallel \dots \parallel w_n} = \bar{w}_1 + \bar{w}_2 + \dots + \bar{w}_n, \\ \overline{w_1 + w_2 + \dots + w_n} = \bar{w}_1 \parallel \bar{w}_2 \parallel \dots \parallel \bar{w}_n, \quad (24)$$

субдистрибутивный и супрадистрибутивный законы

$$\begin{aligned} w \parallel (w_1 + w_2) &\leq (w \parallel w_1) + (w \parallel w_2), \\ w + (w_1 \parallel w_2) &\geq (w + w_1) \parallel (w + w_2), \end{aligned} \quad (25)$$

свойство недостижимости нуля (при конечных w_i и n).

$$w_1 > w_1 \parallel w_2 > w_1 \parallel w_2 \parallel w_3 > \dots > w_1 \parallel \dots \parallel w_n > 0 \text{ и др.} \quad (26)$$

Согласно (26), $w_1 \parallel \dots \parallel w_n \rightarrow 0$ при $n \rightarrow \infty$, $w_i \neq 0$.

При $w_i \in \{0, \infty\}$ логика исчисления иммитансов вырождается в алгебру логики Дж.Буля.

P-ЛОГИКА В ЗАДАЧАХ СИНТЕЗА И АНАЛИЗА НАДЕЖНОСТИ

В [6-10] рассмотрены применения АМ-алгебры в различных областях науки и техники. Здесь рассмотрим некоторые применения логики исчисления надежности.

В общем случае t_i в выражениях (9) являются случайными величинами. В [15] показано, что при независимых отказах функции распределения вероятностей непрерывно-логических конъюнкции (9.1) и дизъюнкции (9.2) определяются соответственно выражениями:

$$\begin{aligned} F(T) &= 1 - \prod_{i=1}^n [1 - F_i(t_i)] = 1 - \prod_{i=1}^n \bar{F}_i(t_i), \\ F(T) &= \prod_{i=1}^n F_i(t_i), \end{aligned} \quad (27)$$

где $F_i(t_i)$ есть функция распределения вероятностей t_i .

Нетрудно показать, что выражения (27) являются функциями P-алгебры:

$$F(T) = \prod_{i=1}^n [1 - F_i(t_i)] = \prod_{i=1}^n \bar{F}_i(t_i) = F_1(t_1) \oplus F_2(t_2) \oplus \dots \oplus F_n(t_n), \quad (28.1)$$

$$F(T) = \bar{F}_1(t_1) \oplus \dots \oplus \bar{F}_n(t_n) = \overline{F_1(t_1) \cdot F_2(t_2) \cdot \dots \cdot F_n(t_n)}, \quad (28.2)$$

где чертой сверху обозначены аддитивные инверсии.

Рассмотрим некоторые примеры типовых структур с последовательно-параллельной ЛСН. На рис.2,3 приведены ЛСН с поэлементным (рис.2) и поканальным (рис.3) резервированием. ЛСН, изображенная на рис.2, содержит n последовательно соединенных фрагментов E_1, E_2, \dots, E_n с параллельной структурой (рис.1б), а ЛСН, представленная на рис.3, содержит m параллельно соединенных фрагментов с последовательной структурой (рис.1а). Согласно (3), для рис.2,3 соответственно получим:

$$P = (1 - \bar{p}_{11} \bar{p}_{21} \dots \bar{p}_{m1}) (1 - \bar{p}_{12} \bar{p}_{22} \dots \bar{p}_{m2}) \dots (1 - \bar{p}_{1n} \bar{p}_{2n} \dots \bar{p}_{mn}), \quad (29.1)$$

$$P = 1 - (1 - p_{11} p_{12} \dots p_{1n}) (1 - p_{21} p_{22} \dots p_{2n}) \dots (1 - p_{m1} p_{m2} \dots p_{mn}). \quad (29.2)$$

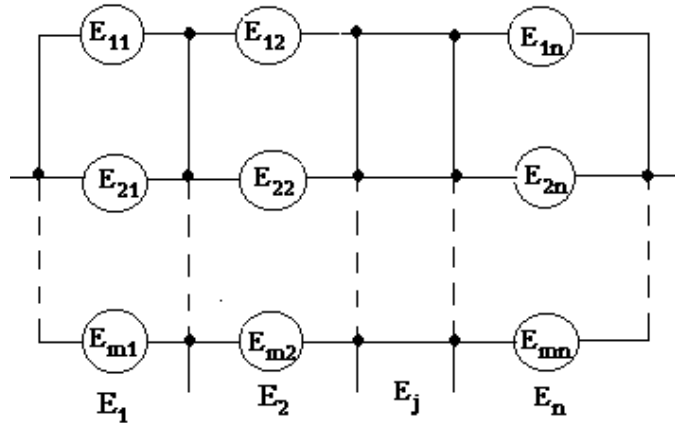


Рисунок 2

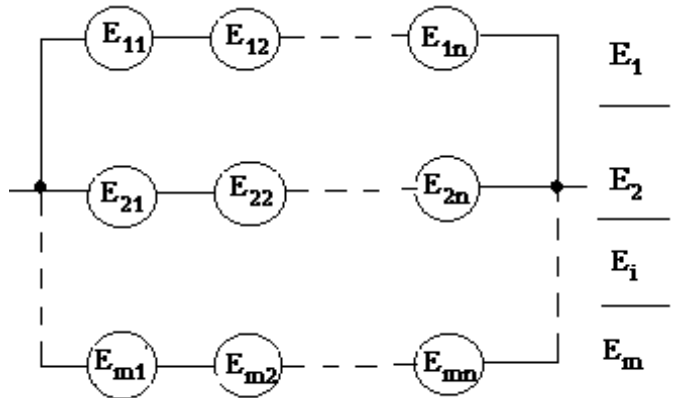


Рисунок 3

Выражения (29), записанные в базе операций P-логики, имеют вид:

$$P = (p_{11} \oplus p_{21} \oplus \dots \oplus p_{m1}) \cdot (p_{12} \oplus p_{22} \oplus \dots \oplus p_{m2}) \times \dots \times (p_{1n} \oplus p_{2n} \oplus \dots \oplus p_{mn}), \quad (30.1)$$

$$P = (p_{11} \cdot p_{12} \cdot \dots \cdot p_{1n}) \oplus (p_{21} \cdot p_{22} \cdot \dots \cdot p_{2n}) \oplus \dots \oplus (p_{m1} \cdot p_{m2} \cdot \dots \cdot p_{mn}). \quad (30.2)$$

Выражения (30) связаны между собой дуальным преобразованием P-операций: $(\cdot, \oplus) \rightarrow (\oplus, \cdot)$. Более высокую надежность обеспечивают структуры с поэлементным резервированием. Например, при равнонадежных элементах при $n=m=2$ и $p_i = 0,900$ ВБР схем (рис.2,3) соответственно равны 0,980 и 0,964. Заменяв в (30) вероятности p_{ij} на обозначения соответствующих элементов E_{ij} приходим к аналитическим записям логических схем надежности:

$$\begin{aligned} \text{ЛСН} &= (E_{11} \oplus E_{21} \oplus \dots \oplus E_{m1}) \times \\ &\times (E_{12} \oplus E_{22} \oplus \dots \oplus E_{m2}) \times \dots \times (E_{1n} \oplus E_{2n} \oplus \dots \oplus E_{mn}), \\ \text{ЛСН} &= (E_{11} \cdot E_{12} \cdot \dots \cdot E_{1n}) \oplus \\ &\oplus (E_{21} \cdot E_{22} \cdot \dots \cdot E_{2n}) \oplus \dots \oplus (E_{m1} \cdot E_{m2} \cdot \dots \cdot E_{mn}). \end{aligned}$$

В соответствии с отображением $(P, p_i, \cdot, \oplus) \rightarrow (T, t_i, \wedge, \vee)$ от формул (30) для ВБР приходим к формулам для времени безотказной работы:

$$\begin{aligned} T &= (t_{11} \vee \dots \vee t_{m1}) \wedge (t_{12} \vee \dots \vee t_{m2}) \wedge \dots \wedge (t_{1n} \vee \dots \vee t_{mn}) = \\ &= \min[\max(t_{11}, \dots, t_{m1}), \max(t_{12}, \dots, t_{m2}), \dots, \max(t_{1n}, \dots, t_{mn})], \end{aligned} \quad (31.1)$$

$$\begin{aligned} T &= (t_{11} \wedge t_{12} \wedge \dots \wedge t_{1n}) \vee (t_{21} \wedge t_{22} \wedge \dots \wedge t_{2n}) \vee \\ &\vee \dots \vee (t_{m1} \wedge t_{m2} \wedge \dots \wedge t_{mn}) = \max[\min(t_{11}, \dots, t_{m1}), \\ &\min(t_{12}, \dots, t_{m2}), \dots, \min(t_{1n}, \dots, t_{mn})]. \end{aligned} \quad (31.2)$$

Согласно свойству идемпотентности непрерывнологических функций [16], при равных $t_{ij} = \tau$ всех элементов E_{ij} из (31) следует, что $T = \tau$. Это означает, что НЛ-модель (31) надежности справедлива при работе резервирующих элементов в нагруженном режиме (горячее резервирование) при постоянном резервировании. При работе всех резервирующих элементов в ненагруженном режиме при резервировании замещением символы $\vee = \max$ в (31) необходимо заменить на символы \wedge сложения:

$$\begin{aligned} T &= (t_{11} + \dots + t_{m1}) \wedge (t_{12} + \dots + t_{m2}) \wedge \dots \wedge (t_{1n} + \dots + t_{mn}) = \\ &= \min(t_{11} + \dots + t_{m1}, t_{12} + \dots + t_{m2}, \dots, t_{1n} + \dots + t_{mn}) = \\ &= \min(T_1, T_2, \dots, T_n), \end{aligned}$$

$$\begin{aligned} T &= (t_{11} \wedge t_{12} \wedge \dots \wedge t_{1n}) + (t_{21} \wedge t_{22} \wedge \dots \wedge t_{2n}) + \\ &+ \dots + (t_{m1} \wedge t_{m2} \wedge \dots \wedge t_{mn}) = \min(t_{11}, t_{12}, \dots, t_{1n}) + \\ &+ \min(t_{21}, t_{22}, \dots, t_{2n}) + \dots + \min(t_{m1}, t_{m2}, \dots, t_{mn}), \end{aligned}$$

где $T_i = t_{1i} + t_{2i} + \dots + t_{mi}$ - время безотказной работы i -го последовательно соединенного вертикального фрагмента ЛСН (рис.2).

При расчете надежности изделий с не параллельно-последовательной ЛСН ($\overline{\text{ПП}}$ -схемы) ее необходимо преобразовать в эквивалентную по надежности с параллельно-последовательной структурой. Вопросы эквивалентных и квазиэквивалентных топологических преобразований рассмотрены в [14]. При строгом подходе в Р-логике дистрибутивные законы не выполняются. Но при p_i близким к нулю или к единице (например, в интервалах $[0; 0,05]$ и $[0,95; 1]$) можно считать, что дистрибутивные законы, соответствующие знаку равенства в (22), практически имеют место (тем более, что в практике используются элементы с $p_i > 0,95$). Это дает возможность использовать методы топологических преобразований, применяемых в релейно-контактных схемах [21].

Строгим способом таких эквивалентных по ВБР топологических преобразователей является метод расчле-

нения исходной $\overline{\text{ПП}}$ -схемы надежности по выбранному базовому элементу E_0 [13], короткое замыкание ($p_0 = 1$) и удаление ($p_0 = 0$) которого превращает исходную не параллельно-последовательную ЛСН в две ПП-схемы с ВБР соответственно $P_{\text{кз}}$ и $P_{\text{ух}}$. Тогда, согласно (17), ВБР исходной схемы будет определяться выражением:

$$P = p_0 P_{\text{кз}} + \bar{p}_0 P_{\text{ух}}, \text{ где } p_0 + \bar{p}_0 = 1. \quad (32.1)$$

Рассмотрим применения метода расчленения на примере мостовой структуры (рис.4а), которая наиболее часто встречается в качестве фрагментов $\overline{\text{ПП}}$ -схем надежности.

Выбрав в качестве базового элемент E_0 , при его коротком замыкании ($p_0 = 1$) и удалении ($p_0 = 0$) приходим соответственно к ПП-схемам короткого замыкания (рис.4б) и холостого хода (рис.4в), для которых:

$$P_{\text{кз}} = (1 - p_1 p_2)(1 - p_3 p_4) = (p_1 \oplus p_2) \cdot (p_3 \oplus p_4),$$

$$P_{\text{ух}} = 1 - (1 - p_1 p_4)(1 - p_2 p_3) = (p_1 \cdot p_4) \oplus (p_2 \cdot p_3).$$

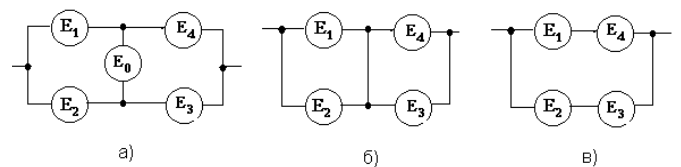


Рисунок 4

Согласно (17), результирующая ВБР исходной мостовой ПП-схемы определяется выражением:

$$\begin{aligned} P &= p_0 P_{\text{кз}} + \bar{p}_0 P_{\text{ух}} = p_0 [(p_1 \oplus p_2) \cdot (p_3 \oplus p_4)] + \\ &+ \bar{p}_0 [(p_1 \cdot p_4) + (p_2 \cdot p_3)]. \end{aligned} \quad (32.2)$$

Выражения (32) являются одной из базовых операций коммутарной алгебры [22,23] с предметными переменными $P_{\text{кз}}$ и $P_{\text{ух}}$, которые являются функциями Р-логики.

Обобщением метода расчленения является способ топологического преобразования в параллельно-последовательные ЛСН разложением исходной ПП-схемы по начальным или конечным элементам [24].

Другой областью применения логики исчисления иммитансов является проверка эмпирических (эвристических) показателей надежности на корректность (при корректном подходе они должны также являться функциями Р-логики). В частности, в [25] надежность изделий, обеспечиваемая при их изготовлении, предлагается оценивать по показателю

$$H = [H_{\text{пр}} + (1 - H_{\text{пр}})a]Д,$$

где $H_{пр}$ - надежность, достигнутая при проектировании; a - коэффициент, выражающий эффективность мероприятий, направленных на обеспечение качества изделия в процессе производства; D - коэффициент качества (дефектности) изготовления изделия. Предлагаемый показатель представим в базисе операций Р-логики: $H = (H_{пр} \oplus a) \cdot D$, т.е. показатель H введен корректно.

Еще одна обширная область применения Р-логики - это синтез высоконадежных изделий (электронных средств) на грубой элементной базе. В качестве примера рассмотрим процедуру синтеза, базирующегося на тождестве (20). Процедура базируется на свойстве операции вероятностного сложения Р-логики:

$$S = y/x = k_1 \oplus k_2 = k_1 + k_2 - k_1 k_2 = k_0(1 + \gamma), \quad (33)$$

где $\gamma = -\delta_1 \delta_2$, подавлять влияние вариации δ_1 и δ_2 коэффициентов передачи $k_i = k_0(1 + \delta_i)$ грубых (по постепенным и внезапным отказам) преобразователей Π_1 и Π_2 при их отклонении от заданных (номинальных) значений $S_0 = k_0 = 1$.

Размерность $[k_1 k_2]$ слагаемого $k_1 k_2$ в (33) отличается от размерности $[k_1] = [k_2]$ слагаемых k_1 и k_2 . Поэтому при топологическом (схемном) отображении (33) (при построении структуры) необходимо ввести обратный преобразователь (ОП) с передачей β , как это показано на рис.5 (схема Блэка), патент США № 1686792, 1928 г.), для которого

$$S = k_1 + k_2 - k_1 k_2 \beta = k_0(1 + \delta_i), \quad (34)$$

где

$$\gamma = (1 - k_0 \beta)(1 + \delta_1 + \delta_2) - k_0 \beta \delta_1 \delta_2. \quad (35)$$

При $k_0 \beta = 1$ (условие настройки) мультипликативная относительная погрешность $\gamma = -\delta_1 \delta_2$, т.е. точность и надежность по постепенным отказам при $|\delta_2| < 1$ повышается по сравнению с вариантом использования одного преобразователя Π_1 в $1/\delta_2$ раз. При этом если один из преобразователей Π_1 или Π_2 выходит из строя (дефект типа "отсутствие прохождения сигнала"), то устройство загрубляется, но продолжает оставаться работоспособным.

Усиление указанных свойств осуществляется через операцию композиции:

$$S = k_1 \oplus k_2 \oplus \dots \oplus k_n = k_0(1 + \gamma), \quad (36)$$

где $\gamma = (-1)^{n-1} \delta_1 \delta_2 \dots \delta_n$, т.е. точность (надежность по

постепенным отказам) при $|\delta_2| < 1$ увеличивается в $1/|\delta_2 \dots \delta_n|$ раз по сравнению с вариантом использования одного преобразователя Π_1 . При равенстве нулю части коэффициентов k_i (при наличии грубого отказа типа "отсутствие прохождения сигнала" при выходе из строя $m \leq n - 1$ преобразователей Π_i) многоканальная система загрубляется $\gamma = \delta_{i_1} \dots \delta_{i_{n-m}}$, но остается работоспособной, что является следствием О-замкнутости АМ-дизъюнкций (6) и (7).

При топологическом отображении алгоритма (36) операциям композиции соответствует процедура итерирования, когда преобразователь Π_1 и Π_2 в исходной схеме (рис.5) многократно замещается самой исходной схемой [26].

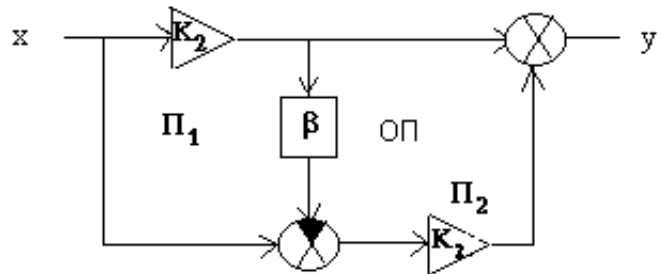


Рисунок 5

В основе логических методов нахождения показателей надежности изделий положены следующие логико-алгебраические структуры:

- АМ-алгебра, которая при $p_i \in [0, 1]$ является логикой исчисления надежности (применительно к задачам построения функций ВБР и ЛСН параллельно-последовательного типа);
- непрерывная (бесконечнозначная, нечеткая) логика (применительно к задачам построения функций времени безотказной работы)
- комлементарная алгебра (применительно к задачам построения функций ВБР для не параллельно-последовательных ЛСН).

При $p_i \in \{0, 1\}$ АМ-модели и непрерывнологические модели вырождаются в булевы модели надежности, теряя при этом свойство континуальности, а функции комлементарной алгебры вырождаются в функции предикатной алгебры выбора (применительно к задачам построения логических функций работоспособности и ЛСН по грубым отказам [27]).

Базовые операции АМ-алгебры являются частными реализациями комлементарной алгебры (КА) с базовыми операциями КА-конъюнкции и КА-дизъюнкции:

$$\Lambda_{(\alpha_1, \alpha_2)}(y_1, y_2) = y_1 \alpha_2 + y_2 \alpha_1, \quad (37.1)$$

$$V_{(\alpha_1, \alpha_2)}(y_1, y_2) = y_1 \alpha_1 + y_2 \alpha_2, \quad (37.2)$$

где $(\alpha_1, \alpha_2) = A$ есть вектор весовых коэффициентов, удовлетворяющих условию комплементарности $\alpha_1 + \alpha_2 = 1$, $(y_1, y_2) = Y$ - вектор предметных переменных.

Изменив обозначения переменных $\alpha_1 = p_1$, $\alpha_2 = \bar{p}_1$, $y_2 = p_2$ и, положив в (37.1) $y_1 = 0$, в (37.2) $y_1 = 1$, приходим к базовым операциям АМ-алгебры:

$$\Lambda_{(p_1, \bar{p}_1)}(0, p_2) = p_1 \cdot p_2, \quad V_{(p_1, \bar{p}_1)}(1, p_2) = p_1 \oplus p_2.$$

В свою очередь комплементарная алгебра [22,28] включает в себя предикатную алгебру выбора (ПАВ) [22,29,30], непрерывную логику (НЛ) [15,16,31] и двоичную булеву алгебру логики (БА) [32], образуя последовательность вложенных логико-алгебраических структур $КА \supset ПАВ \supset НЛ \supset БА$ [33], являющихся математическим базисом для построения логической теории надежности, которая требует своего дальнейшего развития.

В [15] развита теория надежности сложных систем, в которой первичными показателями надежности сложных систем считаются последовательные моменты времени t_i отказов (восстановлений) элементов и аналогичные моменты T_j для всей системы, а поставленная задача состоит в нахождении зависимости T_j от t_i ($j=1,2,\dots,m$) многоканальной системы.

"При этом оказалось, что функции $T_j = f_j(t_i)$ всегда выражаются суперпозицией операций непрерывной логики, а это позволяет говорить о логической теории надежности" [15]. Здесь мы пришли к этому выводу на основе выявленных изоморфизмов при выборе в качестве первичных показателей надежности вероятностей безотказной работы.

Может возникнуть вопрос: "А где же триадный подход [34,35], декларируемый автором в начале статьи?" Он присутствует на начальном этапе проектирования изделия при разработке его структуры (для электронных средств - это электрическая схема объекта проектирования). При этом ЛСН является одной из сопутствующих (производной от основной структуры изделия) вторичных структур заданной диадой (S_p, P) , которая совпадает с триадой (S, P, Q) при $P=Q$ (однокритеральная оптимизация по доминирующему показателю качества). Указанное подтверждает примат формы (структуры) над функцией [36].

ЗАКЛЮЧЕНИЕ

Изложенные в настоящей статье логические основы математической теории надежности являются "сверткой" традиционной математической теории надежности [4,5] на

основе раскрытых изоморфизмов и совмещения в единой структурно-параметрической модели топологической (структурной) и математической (параметрической) моделей надежности.

Достоинством логических представлений являются: компактность и наглядность формул, возможность непосредственного представления топологической модели (ЛСН) путем мысленного "сканирования" структурно-параметрических моделей по вероятностным p_i (Р-логика) и переменным t_i (непрерывная логика) показателям одна в другую, возможность прямого восстановления исходной электрической схемы (W-логика) по заданной ЛСН (Р-логика) и обратно (при одинаковых конфигурациях электрической схемы и ЛСН) и пр.

ПЕРЕЧЕНЬ ССЫЛОК

1. Волгин Л.И. Совмещенные структурно-параметрические модели технических объектов // Модели физических и технических объектов и процессов: Труды Международной научно-техн. конф. "Нейронные, реляторные, непрерывнологические сети и модели".-Ульяновск: УЛГТУ, 1998.-Том 3.- С.3-6.
2. Аршавский В. В. Различные модели мира в свете полиморфизма типов полшарного реагирования // Модели мира / Под ред. Д.А. Пospelова. -М.: Российская Ассоциация искусственного интеллекта, 1997. -С.125-136.
3. Волгин Л. И. Изоморфизмы логической теории надежности //Актуальные проблемы анализа и обеспечения надежности и качества приборов, устройств и систем: Сб.докл. Международной научно-техн.конф. Пенза: Изд-во ПГГУ, 1997. - С.9-10.
4. Гнеденко Б.В, Беляев Ю.К., Соловьев А.Д. Математические методы в теории надежности. - М.: Наука, 1965. - 524 с.
5. Нечипоренко В.И. Структурный анализ ж методы построения надежных систем. - М.: Советское радио, 1968. - 256 с.
6. Волгин Л.И. Свойства и законы двоичных булевых функций на множестве действительных и комплексных чисел //Автоматика и вычислительная техника. -1994. -№5. -С.5-21.
7. Волгин Л.И. Представление функций двоичной булевой алгебры на плоскости комплексных чисел // Электронное моделирование. - 1995. - 17. - № 4. -С.3-10.
8. Волгин Л.И. Свойства и законы двоичной булевой алгебры на плоскости комплексных чисел // Электронное моделирование. - 1994. - № 5. -С.3-10.
9. Волгин Л.И. АМ-алгебра и ее применения // Автоматика и вычислительная техника. - 1996. - № 1. - С.19-33.
10. Волгин Л.И. АМ-алгебра и ее применения // Ульяновск: УЛГТУ, 1997. -52с.
11. Волгин Л.И. АМ-алгебра и надежность // Методы и средства оценки и повышения надежности приборов, устройств и систем: Тезисы докл. Международной научно-техн. конф.-Пенза: ПГТУ, 1995. - С.15-17.
12. Волгин Л.И. Логические основы математической теории надежности // Проблемы и решения современной технологии: Сб. научных трудов ПТИС. - Тольятти: Изд-во Поволжского технологического института сервиса, 1996. - Вып.2. - Часть 2. - С.18-23.
13. Рябинин И.А., Черкесов Г.Н. Логико-вероятностные методы исследования надежности структурно-сложных систем. - М.: Радио и связь, 1981. -264с.
14. Волгин Л.И. Логические основы математической теории надежности. - Ульяновск: УЛГТУ, 1997. - 44 с.
15. Левин В.И. Логическая теория надежности сложных систем. -М.: Энергоатомиздат, 1985. - 128 с.
16. Волгин Л.И., Левин В.И. Непрерывная логика. Теория и применения. - Таллинн: АН Эстонии, 1990. - 210 с.
17. Волгин Л.И. Субдистрибутивный и супрадистрибутивный законы логики исчисления надежности //Надежность и качество: Книга докл. международного симпозиума. - Пенза: Изд-во ПГУ, 1999. С.34-36.
18. Шестаков В.И. Алгебра двухполюсных схем, построенных исключительно из двухполюсников (алгебра А-схем) // Автоматика и телемеханика. - 1941.- № 2. - С.15-24.
19. Волгин Л.И. Логика исчисления иммитансов многоэлементных электрических двухполюсников. - Ульяновск: УЛГТУ, 1998. - 24с.
20. Волгин Л.И. Определение сопротивлений и проводимостей алгебро-логическим методом //Электричество. -1998. -№7.

- С.64-69.
21. Шеннон К. Символический анализ релейных и переключательных схем // Работы по теории информации и кибернетике. - М.: Изд-во иностранной литературы, 1963. - С.9-43.
 22. Волгин Л.И. Комплементарная алгебра и предикатная алгебра выбора. - Ульяновск: УЛГТУ, 1996. - 68с.
 23. Волгин Л.И. Векторная комплементарная алгебра и ее применения. - Ульяновск: УЛГТУ, 1996. - 52с.
 24. Рогинский В.Н. Построение релейных схем управления. - М.-Л.: Энергия, 1964. - 423с.
 25. Чекмарев А.Н. Разработка научно-технических основ управления качеством изготовления радиоэлектронных средств: Автореферат дисс. д-ра техн.наук. - Самара: СГГУ, 1997. - 32 с.
 26. Волгин Л.И. Аналоговые операционные преобразователи для измерительных приборов и систем. - М.: Энергоатомиздат, 1983. - 32с.
 27. Голинкевич Т.А. Прикладная теория надежности. - М.: Высшая школа, 1977. - 160с.
 28. Волгин Л.И. Свойства и законы комплементарной алгебры // Известия АН ЭССР. Физика, математика. - 1988. - №4. - С.417-427.
 29. Волгин Л.И. Предикатная алгебра выбора и синтез функционально-логических преобразователей в элементном базисе релейторов // Электронное моделирование. -1988. -10. -№2. -С.3-9.
 30. Волгин Л.И. Представления функций порядковой логики в предикатной алгебре выбора // Электронное моделирование. - 1990. -12. - №2. -С.3-9.
 31. Волгин Л.И. Непрерывная логика и ее схемотехнические применения. - Ульяновск: УЛГТУ, 1996. -108с.
 32. Поспелов Д.А. Логические методы анализа и синтеза схем. - М.: Энергия, 1968. -228с.
 33. Волгин Л.И. Континуальные логики и предметные алгебры, порождаемые функцией взвешенных степенных средних // Информационные технологии. - 1999, № 9.
 34. Баранцев Р.Г. О тринитарной методологии // Философский вестник. -Вып.7: Между физикой и метафизикой: наука и философия. -СПб, 1998. -С.51-61.
 35. Волгин Л.И. Вельмисов П.А. Триада парадигма познания, //Любимцевские чтения: Сб. докладов. -Ульяновск: УГПУ им. И.Н.Ульянова, 1998. - С.66-68.
 36. Шрейдер Ю.А. Примат формы над функцией // Теория и общие вопросы обработки аналоговой информации: Труды международной конференции "Методы и средства преобразования и обработки аналоговой информации" / Под ред. Л.И. Волгина, 8-10 июня. - Ульяновск: УЛГТУ, 1999. -Том 1. - С.5-6.

УДК 621.396.664

ПОСТРОЕНИЕ АЛГОРИТМА ПОИСКА НЕИСПРАВНОСТЕЙ МНОГОКАНАЛЬНЫХ СИСТЕМ

Д.А.Габриэльян, В.В.Шацкий, Н.В.Шацкий

Описано построение алгоритма поиска неисправностей в многоканальной системе, основанного на учете вкладов ее элементов в определяющий параметр системы. Приведен пример построения алгоритма контроля по разработанному принципу для цилиндрической антенной решетки с дискретными фазовращателями. Показан выигрыш при использовании предложенного алгоритма по сравнению с ранее используемым алгоритмом контроля антенных решеток.

Construction of fault search algorithm for the multichannel system based on taking into account its element contributions into the defining system parameter has been described. An example for constructing the control algorithm in terms of the principle developed for the cylindrical antenna array with the discrete phase shifters has been given. The advantage in using the algorithm suggested has been shown as compared to that for antenna array control used previously.

Объективной тенденцией развития техники на современном этапе является постоянный рост сложности устройств, что объясняется расширением круга решаемых ими задач при одновременном повышении требований к эффективности функционирования. Принципиальная особенность многоканальных систем, относящихся к системам с квазиизбыточностью, заключается в возможности полного или частичного сохранения эффективности функционирования даже при наличии отказов в одном или нескольких каналах. Последнее достигается изменением алгоритма функционирования системы с учетом информации о характере и местоположении отказавшего элемента.

При таком подходе контроль многоканальной системы в отличие от контроля работоспособности требует разбиения области отказовых состояний, ранее рассматриваемой как

единое целое, на множество подобластей, соответствующих подлежащим различению неисправностям. В результате имеющая место при контроле работоспособности двухальтернативная задача превращается в многоальтернативную, конечной целью решения которой является отнесение контролируемой многоканальной системы - объекта контроля - к одной из заданных подобластей пространства отказовых состояний.

Важным вопросом при контроле многоканальной системы является минимизация снижения интегральной эффективности системы в целом за время поиска неисправности. Для этого алгоритм определения подобласти отказовых состояний должен учитывать не только вероятность нахождения системы в некотором отказовом состоянии, но и снижение эффективности функционирования системы при этом. Однако существующие в настоящее время подходы к построению алгоритмов поиска неисправностей опираются только на вероятности появления отказов [1].

Таким образом, построение алгоритма поиска неисправностей в многоканальной системе представляет собой не только теоретический интерес, но имеет и практическую значимость.

Целью данной статьи является построение алгоритма контроля многоканальной системы, при котором минимизируется интегральное снижение эффективности функционирования системы за время определения характера и местоположения отказавшего элемента.

Рассмотрим N -канальную систему, которая может находиться в одном из L состояний, определяемых наличием i -кратной ошибки

$$L = \sum_{i=0}^N C_N^i, \quad (1)$$

где C_N - биномиальный коэффициент [2,3]. Сразу отметим, что случай $i = 0$ соответствует полностью исправной системе, случай $i = N$ - полностью неисправной системе, а каждое состояние возникновения i -кратного отказа может быть реализовано C_N различными вариантами.

Вероятность возникновения i -кратного отказа в N -канальной системе при одинаковых вероятностях возникновения отказов q в каждом канале определяется выражением

$$P_i = p^{N-i} q^i, \quad (2)$$

в котором $p = 1 - q$.

Как следует из (1) и (2), все события образуют полную группу, поскольку $\sum_{i=0}^N C_N^i p^{N-i} q^i = (q + 1 - q)^N \equiv 1$. Это доказывает, что данные состояния являются случайными независимыми событиями.

В современных системах, построенных с использованием высоконадежных элементов, вероятность отказа каждого из каналов не превышает $10^{-6} \dots 10^{-7}$. В этом случае для системы с $N = 100$ вероятность того, что в системе будет не более одного отказа, составит $0,999 \dots 0,9999$. В связи с этим в дальнейшем ограничимся только рассмотрением случаев возникновения однократных отказов, вероятности которых при указанных параметрах составят $10^{-4} \dots 10^{-5}$.

Определим эффективность функционирования системы функционалом вида

$$K = K(\vec{A}, \vec{Q}), \quad (3)$$

в котором $\vec{A} = (a_1, a_2, \dots, a_N)$ - вектор, описывающий определяющие параметры функционирования каждого из N каналов в исправном состоянии; $\vec{Q} = (q_1, q_2, \dots, q_N)$ - вектор, определяющий вероятность появления отказов в каналах системы. Отметим, что однократному отказу в n -м канале системы соответствует вектор $\vec{Q} = (0, \dots, 0, q_n, 0, \dots, 0)$.

Выбор вида функционала K определяется рядом факторов, к которым можно отнести назначение системы, алгоритм функционирования, условия функционирования и т.д.

Введем понятие снижения эффективности функциони-

рования системы, связанного с возможностью возникновения отказа в n -м канале, которое определим соотношением

$$\delta_n = K(\vec{A}, \vec{Q}_0) - K(\vec{A}, \vec{Q}_n), \quad (4)$$

где $\vec{Q}_0 = (0, \dots, 0)$.

Интегральное снижение эффективности функционирования системы за время существования отказа может быть оценено с помощью выражения вида

$$\delta K = \sum_{n=1}^N t_n \delta_n,$$

или связанного с ним среднего снижения показателя эффективности, определяемого соотношением

$$\hat{\delta K} = \delta K / T. \quad (5)$$

Выше принято, что t_n - время поиска неисправности, определяемое порядком контроля (вектором предпочтения) каналов системы, а T - период контроля системы. Вектор предпочтения должен учитывать не только требуемое время поиска неисправности в n -м канале t_n , но и вклад данного канала в снижение эффективности функционирования системы, вызванное возникновением неисправности. Для системы, при контроле каналов которой выполняется условие $t_n = nt$ ($n = 1, \dots, N$), вектор предпочтения может быть записан в виде

$$\vec{W}(r) = \max_{r \in N} (\delta_r). \quad (6)$$

Значениями вектора предпочтения являются N чисел от 1 до N , определяющие порядок, в котором осуществляется контроль каналов системы. При этом первым контролируется канал с номером r_1 , для которого величина δ_{r_1} является максимальной, следующим проверяется канал с номером r_2 , в котором δ_{r_2} - вторая по величине и т.д. Последним проверяется канал с номером r_N , для которого δ_{r_N} минимально.

Таким образом, использование критерия (5) или его частного случая (6) позволяет при известных соотношениях для эффективности функционирования многоканальной системы определить порядок контроля каналов, минимизирующий интегральный показатель.

В качестве примера многоканальной системы рассмотрим цилиндрическую антенную решетку (ЦАР), эффективность функционирования которой определяется точностью установки луча в заданном направлении. Геометрия задачи и отсчет угла θ показаны на рис.1.

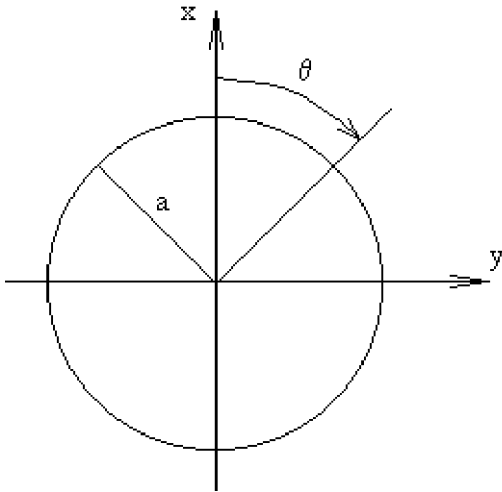


Рисунок 1 - Геометрия задачи для ЦАР и отсчет угла θ

Исследования выполним на примере ЦАР с излучателями $n = \overline{(1, N)}$ в виде ориентированных параллельно образующей идеально проводящего кругового цилиндра магнитных диполей (продольных щелей). Нормированная диаграмма направленности $\mu_n(\theta)$ указанных излучателей определяется в виде суммы по бесконечным азимутальным гармоникам выражением [4,5]

$$\mu_n = \sum_{m=0}^{\infty} \varepsilon_m i^m \frac{\cos[m(\theta - \theta^{(n)})]}{H_m^{(2)'}(\beta a)}, \quad (7)$$

где ε_m - число Неймана ($\varepsilon_m = 1$ при $m = 0$ и $\varepsilon_m = 2$ при $m \neq 0$); a - радиус цилиндра; $\theta^{(n)}$ - угловое положение n -го излучателя; $\beta = 2\pi/\lambda$ - волновое число свободного пространства; λ - длина волны; i - мнимая единица; $H_m^{(2)'}(\bullet)$ - производная функции Ганкеля 2-го рода m -го порядка.

Вид функционала (3) для рассматриваемой антенной решетки определяется смещением главного максимума функции $|F_0(\theta_0)|^2$ из точки θ_0 в точку θ_1 , причем эта функция изменяется на $|F_0(\theta_1)|^2 + |\Delta f(\theta_1)|$ и характеризуется дисперсией положения луча ЦАР при отказах (σ_θ^2). Функционал при этом, в случае использования в ее каналах устройств управления фазой (фазовращателей), принимает вид [2,3,5,6]

$$K(\vec{A}, \vec{Q}) = \sigma_\theta^2 = (|\Delta f(\theta_1)|^2)' \cdot \sigma^2 / (|F_0(\theta_0)|^2)', \quad (8)$$

где $|\Delta f(\theta_1)|^2$ - приращение ДН по мощности ЦАР, обусловленное неисправностями в ее каналах управления;

σ^2 - дисперсия отказов устройств управления фазой; $|F_0(\theta_0)|^2$ - ДН по мощности антенны без отказов. Дисперсия σ^2 характеризуется видом отказа (короткое замыкание или разрыв), типом устройства (непрерывное или дискретное; биномиального типа или лестничного) и выражение для нее приведено в [2,3,5,6]. Вектор $\vec{A} = (a_1, a_2, \dots, a_N)$ в функционале (8) определяется амплитудным и фазовым распределениями в каналах, входящими в ДН по мощности ЦАР; наименьшим дискретом фазовращателя в дисперсии σ^2 ; а вектор $\vec{Q} = (q_1, q_2, \dots, q_N)$ - определяется вероятностью отказа устройств управления фазой, также входящей в дисперсию σ^2 .

С учетом вышесказанного выражение (4) принимает вид

$$\delta_n = 2A_n^2 \left| \operatorname{Re} \{ \exp(i\varphi_n) [\mu_n(\theta) F_0^*(\theta_0)]'_{\theta=\theta_0} \} \right| \times \sigma_n^2 / (|F_0(\theta_0)|^2)'_{\theta=\theta_0}, \quad (9)$$

где $A_n(\varphi_n)$ - номинальное значение амплитуды (фазы) возбуждения в n -м канале ЦАР; σ_n^2 - дисперсия отклонения фазы возбуждения в n -м канале от номинального вследствие возникновения отказов; θ_0 - угловое положение максимума главного лепестка ДН в отсутствие отказов; * - знак комплексного сопряжения.

Следует подчеркнуть, что соотношение (9) предполагает лишь независимость ошибок в излучателях и их однородность. Величины (дисперсии) ошибок и закон распределения их могут быть произвольными. Допускается также корреляция различных видов ошибок в отдельных излучателях, характер которых одинаков для всех излучателей.

Соотношение (9) определяет степень влияния отказов отдельных устройств управления на флуктуации направления главного лепестка ДН в зависимости от положения излучающего элемента в составе раскрыва и вида амплитудно-фазового распределения. При проведении вычислительного эксперимента полагалось, что круговой цилиндр имеет радиус $a = 2,5\lambda$; элементы дуговой антенной решетки из 16 излучателей располагались равномерно в пределах полуокружности и возбуждались по закону $A_n = [\cos(\theta^{(n)})]^2$ и $\varphi_n = 2\pi a/\lambda [1 - \cos(\theta^{(n)})]$, а реализация фазы в каждом канале ЦАР производится с помощью четырехсекционных бинарных фазовращателей, причем $\theta_0 = 0^0$.

На рис.2 приведены результаты исследований величины дисперсии положения максимума главного лепестка ДН ЦАР в зависимости от вероятности q выхода из строя ее канала управления в виде кривых 1-8. Индекс 8 соответствует центральному, относительно направления фазирования излучателю, а остальные излучатели располага-

ются слева от центрального на излучающем раскрыве дуговой антенной решетки в порядке убывания. Как показывают расчеты, поведение кривых определяется вкладом комплексной амплитуды канала в суммарную ДН, положением излучателя на раскрыве, типом устройства управления и его дискретностью.

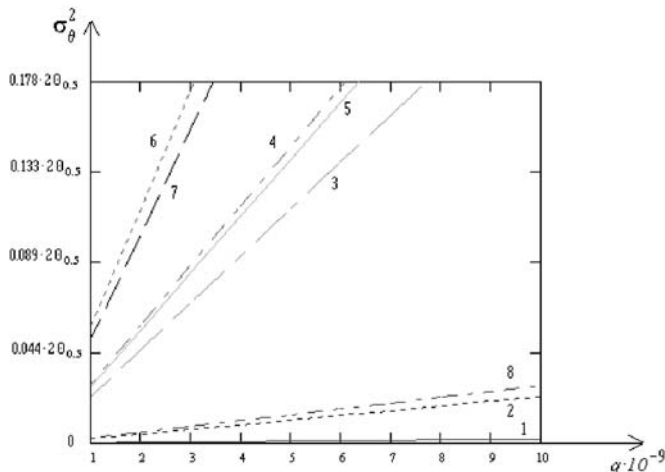


Рисунок 2 - Зависимость дисперсии положения максимума главного лепестка диаграммы направленности ЦАР от вероятности выхода из строя ее каналов управления: 1 и 16 каналы - кривая 1; 2 и 15 каналы - кривая 2; 3 и 14 каналы - кривая 3; 4 и 13 каналы - кривая 4; 5 и 12 каналы - кривая 5; 6 и 11 каналы - кривая 6; 7 и 10 каналы - кривая 7; 8 и 9 каналы - кривая 8

На рис.3 приведены кривые зависимостей, характеризующие снижение дисперсии σ_{θ}^2 за время контроля всех элементов ЦАР по алгоритму, разработанному авторами (кривая 1), и по строчно-столбцевому принципу (кривая 2) [7]. При сравнении двух алгоритмов на основе выражения (5) принято, что время контроля одного канала одинаково, т.е. $t = \text{const}$.

Необходимо отметить, что при оценке эффективности алгоритмов контроля необходимо говорить о среднем времени проведения проверки элементов, которая определяется из выражения (5) в виде равенства

$$\hat{\delta K}^{(cc)} = \sum_{n=1}^N t_n n \delta_n / T = \sum_{n=1}^N n \delta_n / N \quad (10)$$

для строчно-столбцевого алгоритма и уравнения вида

$$\hat{\delta K}^{(vec)} = \sum_{r=1}^N n W_r / N \quad (11)$$

для алгоритма, разработанного авторами.

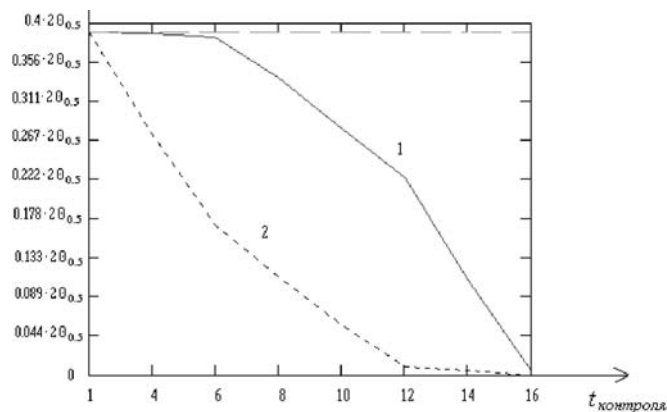


Рисунок 3 - Снижение дисперсии отклонения положения главного максимума диаграммы направленности ЦАР с отказами при контроле каналов антенны по строчно-столбцевому алгоритму (кривая 1) и по разработанному алгоритму (кривая 2)

Из выражений (10) и (11) видно, что выигрыш алгоритма контроля определяется весовыми добавками n (в выражении (10)) и r (в выражении (11)), стоящими перед рассчитанными вкладами элементов ОК в определяющую характеристику системы. С учетом сказанного эффективным считается тот алгоритм, у которого сумма произведений весовых добавок и вкладов наибольшая. Например, для рассмотренного в статье случая при строчно-столбцевом алгоритме контроля $\hat{\delta K}^{(cc)} = 2,9825$, а при алгоритме, разработанном авторами, среднее время равно $\hat{\delta K}^{(vec)} = 1,4349$; т.е. эффективность контроля при применении нового алгоритма возросла в 2,078 раза. Таким образом, установлено, что проведение контроля по разработанному алгоритму позволяет существенно сократить время восстановления направленных свойств антенной решетки, по сравнению со строчно-столбцевым контролем.

ВЫВОДЫ

1. Предложенный алгоритм контроля элементов многоканальных систем позволяет определить порядок проверки ее элементов по величине вносимого ими вклада в определяющий параметр, что должно обеспечить ускоренное восстановление характеристик системы при осуществлении дифференциального контроля за счет уменьшения времени существования в системе отказов с наибольшим вкладом в определяющую характеристику.

2. Полученные при численном исследовании зависимости дисперсии положения максимума диаграммы направленности ЦАР от вероятности выхода из строя ее каналов позволяют утверждать о большем влиянии на σ_{θ}^2 каналов антенной решетки, расположенных в угловом секторе $35^\circ \dots 150^\circ$, а также определяют последовательность проверки элементов антенны.

3. Построенный алгоритм поэлементного контроля элементов управления 16-элементной ЦАР при обеспечении точной установки луча антенны в заданном направлении показал, что эффективность контроля антенной решетки возросла в 2,078 раза по сравнению с контролем по строчно-столбцовому алгоритму.

ПЕРЕЧЕНЬ ССЫЛОК

1. Кудрицкий В.Д., Сеница М.А., Чинаев П.И. Автоматизация контроля радиоэлектронной аппаратуры/Под ред. Чинаева П.И. - М.: Сов. Радио, 1977. - 256с.
2. Габриэлян Д.Д., Шацкий В.В., Шацкий Н.В. Алгоритм поэлементного контроля фазированной антенной решетки по одному определяющему параметру // Вопросы радиоэлектроники. Сер. Общие вопросы радиоэлектроники. - 1999,

- вып.№19.
3. Шацкий Н.В. Принцип построения алгоритма поиска неисправностей в антенной решетке с оценкой его эффективности // Исследовано в России, 38, стр.499-507, 2000. <http://zhurnal.mipt.rssi.ru/articles/2000/038.pdf>.
4. Марков Г.Т., Чаплин А.Ф. Возбуждение электромагнитных волн. - М.: Радио и связь, 1983.
5. Габриэлян Д.Д., Звездина М.Ю., Шацкий Н.В. Точность установки луча в цилиндрической антенной решетке при наличии отказов в каналах управления амплитудой // Изв. вузов: Радиоэлектроника.- 1999, т.42, №5-6, с.19-23.
6. Габриэлян Д.Д., Шацкий Н.В. Оценка характеристик плоской фазированной антенной решетки при наличии отказов в каналах управления фазой // Вопросы радиоэлектроники. Сер. Общие вопросы радиоэлектроники. - 1998, вып.№18, с. 64-69.
7. Самойленко В.И., Шишов Ю.А. Управление фазированными антенными решетками. - М.: Радио и связь, 1983. - 240с.

УДК 517+ 681.3

МОДЕЛИРОВАНИЕ ФУНКЦИИ ЛАНЖЕВЕНА ЦЕПНОЙ ДРОБЬЮ

В.И.Грядун, В.П.Пинчук, В.П.Шаповалов

Для получения контролируемого с заданной точностью расчета функции Ланжевена предложено разложение в цепную дробь. Разработанный алгоритм расчета реализован в системе программирования MATLAB. Модель позволяет достигать минимальную машинную погрешность $\epsilon = 0$, которая эквивалентна точности $\sim 10^{-16}$.

Для проведения контролируемого с заданной точностью расчета функции Ланжевена предложено разложение в цепную дробь. Разработанный алгоритм расчета реализован в системе программирования MATLAB. Модель позволяет достигать минимальную машинную погрешность $\epsilon = 0$, которая эквивалентна точности $\sim 10^{-16}$.

For receipt of the Langeven's function computation controlled with set exactness is offered decomposition in continued fraction. Worked up computation algorithm realized in the MATLAB programming system. Model allows to achieve minimum machine error $\epsilon = 0$, which equivalent to exactness $\sim 10^{-16}$.

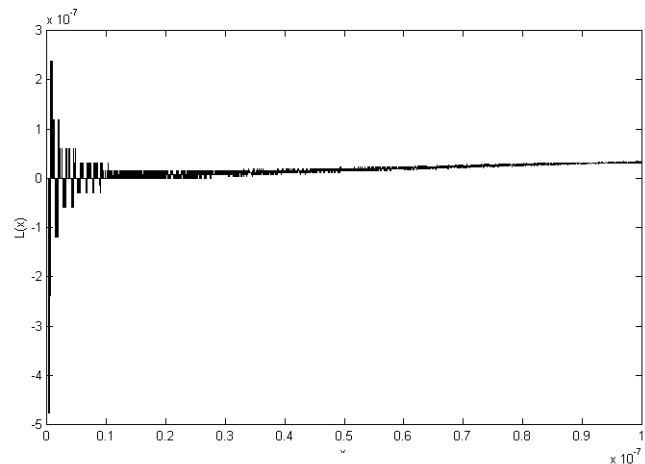


Рисунок 1 - Расчет функции Ланжевена по определению (1)

ВВЕДЕНИЕ

Расчет функции Ланжевена

$$L(x) = \text{cth}(x) - \frac{1}{x} \tag{1}$$

на ЭВМ в области малых положительных величин x дает результаты с большой погрешностью, достигающей миллионов процентов. Действительно, на рис. 1 хорошо видны недостатки расчета, причем, наблюдаются даже отрицательные значения функции, которая является положительной в рассматриваемом диапазоне аргументов [1].

Такие недостатки расчетов функции Ланжевена проявляются на всех типах ЭВМ и характерны для различных систем программирования, например, MathCAD, Matlab, Pascal и др.

Рассматриваемая погрешность негативно проявляется в физических приложениях. Так, при вычислении средней проекции электрического диполя p_e частицы на направление электрического поля по формуле

$$\overline{p_e} = p_e L(x), \tag{2}$$

как раз реализуются малые значения аргумента

$$x = \frac{p_e E}{k_B T}, \quad (3)$$

где p_e - электрический дипольный момент молекулы, Кл·м; E - напряженность электрического поля, В/м; k_B - постоянная Больцмана, Дж/К; T - абсолютная температура, К.

Для молекул воды, например, электрический дипольный момент, составляет величину $6,1 \cdot 10^{-30}$ Кл·м [2] и при комнатной температуре в электрических полях $10^1 \dots 10^6$ В/м аргумент x составляет величины $\sim 10^{-8} \dots 10^{-3}$, которые попадают в критическую область расчета средней проекции диполя по формуле (2).

В табл. 1 приведены результаты расчета средней проекции электрического дипольного момента молекулы анилина ($p_e = 5 \cdot 10^{-30}$ Кл·м) в широком диапазоне электрических полей. Видим, что только в сильных электрических полях $\sim 10^7$ В/м абсолютная погрешность заметно снижается.

Таблица 1- Расчет зависимости средней проекции дипольного электрического момента молекулы анилина от величины электрического поля

| Электрическое поле E , В/м | Аргумент, x | Проекция дипольн. момента молекулы, Кл·м | |
|------------------------------|-----------------------|------------------------------------------|-------------------------|
| | | Расчет по формуле (2) | Корректное значение |
| 1 | $1,208 \cdot 10^{-9}$ | $-1,161 \cdot 10^{-29}$ | $2,0167 \cdot 10^{-39}$ |
| 10 | $1,21 \cdot 10^{-8}$ | $7,42 \cdot 10^{-31}$ | $2,0167 \cdot 10^{-38}$ |
| 100 | $1,21 \cdot 10^{-7}$ | $1,71 \cdot 10^{-32}$ | $2,0167 \cdot 10^{-37}$ |
| 1000 | $1,21 \cdot 10^{-6}$ | $1,91 \cdot 10^{-34}$ | $2,0167 \cdot 10^{-36}$ |
| 10^4 | $1,21 \cdot 10^{-5}$ | $1,8546 \cdot 10^{-35}$ | $2,0167 \cdot 10^{-35}$ |
| 10^5 | $1,21 \cdot 10^{-4}$ | $2,0128 \cdot 10^{-34}$ | $2,0167 \cdot 10^{-34}$ |
| 10^6 | $1,21 \cdot 10^{-3}$ | $2,0128 \cdot 10^{-33}$ | $2,0167 \cdot 10^{-33}$ |
| 10^7 | $1,21 \cdot 10^{-2}$ | $2,0128 \cdot 10^{-32}$ | $2,0167 \cdot 10^{-32}$ |
| 10^8 | $1,21 \cdot 10^{-1}$ | $2,0109 \cdot 10^{-31}$ | $2,0147 \cdot 10^{-31}$ |
| 10^9 | 1,21 | $1,841 \cdot 10^{-30}$ | $1,8438 \cdot 10^{-30}$ |

Аналогичная картина наблюдается при расчете средних проекций магнитных диполей на направление внешнего магнитного поля. В табл. 2 представлены результаты расчетов для титана, у которого магнитный момент атома

p_m составляет величину $\sim 1,64 \cdot 10^{-24}$ А·м² при удельной восприимчивости $3,19 \cdot 10^{-9}$ м³/кг [3]. Здесь также наблюдаются громадные абсолютные погрешности, которые снижаются только при достижении напряженности магнитного поля $\sim 10^5$ А/м.

Таким образом, вычисление функции Ланжевена по ее оригинальному определению (1) не представляется достаточно надежным и возникает вопрос о необходимости обеспечения контролируемой точности расчета. В связи с этим предлагается к рассмотрению метод цепных дробей в вычислительной математике [4].

Таблица 2 - Зависимость проекции магнитного момента атома титана на направление магнитного поля от его величины

| Напряженность магнитного поля, А/м | Аргумент, x | Проекция парамагнитн. момента атома, А·м ² | |
|------------------------------------|-----------------------|-------------------------------------------------------|-------------------------|
| | | Расчет по формуле (2) | Корректное значение |
| 1 | $4,98 \cdot 10^{-10}$ | $-3,14 \cdot 10^{-22}$ | $2,7224 \cdot 10^{-34}$ |
| 10 | $4,98 \cdot 10^{-9}$ | $1,64 \cdot 10^{-24}$ | $2,7224 \cdot 10^{-33}$ |
| 10^2 | $4,98 \cdot 10^{-8}$ | $1,75 \cdot 10^{-26}$ | $2,7224 \cdot 10^{-32}$ |
| 10^3 | $4,98 \cdot 10^{-7}$ | $-4,53 \cdot 10^{-28}$ | $2,7224 \cdot 10^{-31}$ |
| 10^4 | $4,98 \cdot 10^{-6}$ | $-8,75 \cdot 10^{-31}$ | $2,7224 \cdot 10^{-30}$ |
| 10^5 | $4,98 \cdot 10^{-5}$ | $2,72 \cdot 10^{-29}$ | $2,7224 \cdot 10^{-29}$ |
| 10^6 | $4,98 \cdot 10^{-4}$ | $2,72 \cdot 10^{-28}$ | $2,7224 \cdot 10^{-28}$ |
| 10^7 | $4,98 \cdot 10^{-3}$ | $2,72 \cdot 10^{-27}$ | $2,7224 \cdot 10^{-27}$ |
| 10^8 | $4,98 \cdot 10^{-2}$ | $2,72 \cdot 10^{-26}$ | $2,7219 \cdot 10^{-26}$ |
| 10^9 | 0,498 | $2,6781 \cdot 10^{-25}$ | $2,6784 \cdot 10^{-25}$ |
| 10^{10} | 4,98 | $1,311 \cdot 10^{-24}$ | $1,311 \cdot 10^{-24}$ |
| 10^{11} | 49,794 | $1,607 \cdot 10^{-24}$ | $1,607 \cdot 10^{-24}$ |

1. РАЗЛОЖЕНИЕ ФУНКЦИИ ЛАНЖЕВЕНА ЦЕПНОЙ ДРОБЬЮ

Для выражения функции Ланжевена через цепную дробь удобно воспользоваться представлением гиперболического тангенса в виде отношения [5]

$$\text{th}(x) = \frac{x}{C_1(x)}, \quad (4)$$

где $C_1(x)$ - бесконечная цепная дробь, определяемая выражением

$$C_1(x) = 1 + \frac{x^2}{3 + \frac{x^2}{5 + \frac{x^2}{7 + \dots}}} \quad (5)$$

Подставляя затем (4) в (1) с учетом (5) и проведя несложные упрощения, получим

$$L(x) = \frac{x}{C_2(x)}, \quad (6)$$

где $C_2(x)$ - часть цепной дроби (5), определяемая выражением

$$C_2(x) = 3 + \frac{x^2}{5 + \frac{x^2}{7 + \frac{x^2}{9 + \dots}}} \quad (7)$$

Теорема. Бесконечная цепная дробь (7) сходится в области значений аргумента от нуля до бесконечности.

Доказательство. Перепишем (7) в следующем виде

$$C_2(x) = 3 + \frac{1}{\frac{5}{x^2} + \frac{1}{7 + \frac{1}{\frac{9}{x^2} + \frac{1}{11 + \dots}}}} \quad (8)$$

Теперь, обозначая: $a_0 = 3$; $a_1 = 5/x^2$; $a_2 = 7$; $a_3 = 9/x^2$; $a_4 = 11$ и т.д., получим обыкновенную цепную дробь

$$C_2(x) = a_0 + \frac{1}{a_1 + \frac{1}{a_2 + \frac{1}{a_3 + \frac{1}{a_4 + \dots}}}} \quad (9)$$

которая сходится [6], если

$$\sum_{n=1}^{\infty} a_n = \infty. \quad (10)$$

Ряд (10) расходится, т.к. его нечетные члены составляют целочисленную арифметическую прогрессию.

Сходимость цепной дроби (9) позволяет моделировать функцию Ланжевена в области положительных аргументов. Но т. к. она бесконечна, то возникает вопрос о выборе количества звеньев дроби при расчетах. Естественным будет путь установления связи между погрешностью расчета и количеством звеньев в цепной дроби.

2. АЛГОРИТМ РАСЧЕТА И АНАЛИЗ ПОГРЕШНОСТЕЙ

Представим аргумент функции Ланжевена с учетом его погрешности в следующем виде

$$x = \tilde{x} + \delta x, \quad (11)$$

где \tilde{x} - приближенное значение аргумента; δx - погрешность. Тогда погрешность функции, рассчитанной по формуле (1), определится выражением

$$\delta L(x) = \left[1 - \operatorname{cth}(\tilde{x}) + \frac{1}{\tilde{x}^2} \right] \delta x. \quad (12)$$

Из (12) видно, что при малых значениях аргумента, когда $x \ll 1$, погрешность функции очень велика и составляет

$$\delta L(x) \approx \frac{1}{2\tilde{x}^2} \delta x, \quad (13)$$

а при больших значениях аргументов, когда $x \gg 1$ погрешность мала и составляет величины, определяемые отношением

$$\delta L(x) \approx \frac{1}{x^2} \delta x. \quad (14)$$

Таким образом, действительно, использование формулы (1) для расчета функции Ланжевена в области малых аргументов является крайне нежелательным из-за неконтролируемой точности вычислений.

Покажем, что расчет функции Ланжевена по формуле (6) с использованием цепной дроби (7) позволяет получать результат с наперед заданной точностью.

Известно [7], что если обыкновенная бесконечная цепная дробь сходится, то ее значение больше любой подходящей дроби четного порядка и меньше любой подходящей дроби нечетного порядка. Поскольку цепная дробь находится в знаменателе формулы (6), то функция Ланжевена для заданного положительного значения аргумента будет меньше ее значения рассчитанного с помощью любой подходящей дроби четного порядка и, соответственно, большей ее величины полученной с помощью любой подходящей дроби нечетного порядка.

Такое поведение функции Ланжевена, рассчитанной с помощью конечных цепных дробей, позволяет ввести максимальную погрешность ее расчета для двух ближайших подходящих дробей с помощью разности

$$\delta L(x) = L_{2k}(x) - L_{2k-1}(x). \quad (15)$$

Введя теперь желаемую точность расчета ϵ и наращивая количество звеньев подходящей конечной цепной дроби, можно достигнуть выполнения необходимого условия

$$\delta L(x) \leq \epsilon. \quad (16)$$

Так как разрядность представления чисел в ЭВМ ограничена, то можно потребовать достижения максимальной точности, то есть положить

$$\epsilon = 0, \quad (17)$$

что обеспечивает точность расчета $\sim 10^{-16}$.

Рассмотренный алгоритм расчета функции Ланжевена с помощью подходящих дробей по формуле (6) с максимально допустимой машинной точностью практически реализован в системе Matlab 5.3, рис. 2.

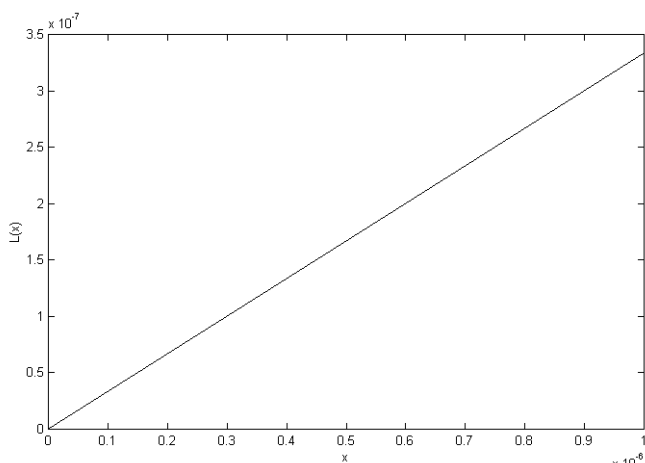


Рисунок 2 - Моделирование функции Ланжевена цепной дробью

ВЫВОДЫ И ОБСУЖДЕНИЕ РЕЗУЛЬТАТОВ

Построенный алгоритм расчета функции Ланжевена с помощью бесконечной цепной дроби $C_2(x)$ хорошо реализуется в критически неустойчивой области расчета

малых аргументов, причем, количество звеньев цепной дроби, необходимых для достижения максимальной точности расчета, небольшое и составляет величину ~ 7 .

При дальнейшем увеличении аргумента количество ветвей цепной дроби (6) для достижения предельной точности возрастает. Однако, следует учесть, что при достижении аргументом x величины ~ 48 расчет функции Ланжевена с помощью цепной дроби по точности не отличается от расчета по определяющей формуле (1). Более того, анализ показывает, что при $x \geq 48$ гиперболический котангенс с двойной точностью равен машинной единице, и поэтому функцию Ланжевена удобнее рассчитывать по формуле

$$L(x) = 1 - \frac{1}{x}, \quad (18)$$

которая дает такой же результат, как и формулы (1) и (6).

Таким образом, учтя в алгоритме расчета функции Ланжевена минимальную величину аргумента, при котором уже нет необходимости пользоваться цепной дробью, мы имеем во всей области определения функции Ланжевена предельную контролируемую точность расчета.

ПЕРЕЧЕНЬ ССЫЛОК

1. Е. Янке, Ф. Эмде, Ф. Леш. Специальные функции. - М.: Наука, 1968. - 344 с.
2. Краткий справочник физико-химических величин. / Под ред. К. Мищенко и А. Равделя. - Ленинград: Химия, 1974. - 200 с.
3. Свойства элементов. Часть 1. Физические свойства. / Под ред. Г. В. Самсонова. - М.: Металлургия, 1978. - 600 с.
4. В. Я. Скоробогатько. Теория ветвящихся цепных дробей и ее применение в вычислительной математике. - М.: Наука, 1983. - 312 с.
5. А. Н. Хованский. Приложение цепных дробей и их обобщений к вопросам приближенного анализа. - М.: Госизд. ТТЛ, 1956. - 203 с.
6. Stern. Ueber der Kennzeichen der Convergenz eines Kettenbruches // J. Reine und angew. Math. 37 (1848), 255-272.
7. А. Я. Хинчин. Цепные дроби. - М.: Гос. изд. физ.-мат. лит., 1961. - 112 с.

УДК 681.32:007.52

ПРЕДВАРИТЕЛЬНАЯ НАСТРОЙКА ЯДЕРНОЙ НЕЙРОННОЙ СЕТИ

А.Ю.Дорогов

В статье рассматриваются многослойные нейронные сети с ядерной организацией. Исследуются вопросы предварительной настройки синаптической карты нейронной сети перед выполнением процедуры обучения. Предложены принципы оптимальной настройки. Получены расчетные формулы вычисления синаптических весов. Представлена методика статического обучения.

The article multilayer neural networks with kernel organization are considered. The problems of initial setting of neural network synaptic card before learning procedure are investigated. The principles of optimum setting are offered. The formulas for

calculation of synapses weights are obtained. The technique of static tutoring is represented.

ВВЕДЕНИЕ

Для обучения многослойных нейронных сетей в подавляющем большинстве случаев используется алгоритм Error Backpropagation [1] или его модификации. Алгоритм обучения данного класса предполагает, что начальные значения синаптических весов предварительно установ-

лены, причем таким образом, чтобы сеть находилась в некотором общем положении, позволяющем поисковой процедуре найти глобальный минимум критерия обучения. Как правило, используется случайная установка начальных значений к уровням $\pm 1/2$. Такой вариант не всегда является удовлетворительным, особенно в тех случаях, когда истинные значения входных и выходных данных сильно отличаются от значений внутренних переменных нейронной сети. В результате процесс обучения сильно затягивается или даже останавливается (в этом случае говорят, что нейронная сеть попала в состояние паралича). В настоящей статье предлагается алгоритм, который позволяет улучшить процедуру начальной установки синаптических весов за счет использования дополнительной информации об обучающей выборке.

Можно выделить следующие цели статистической настройки:

- выполнить приведение истинных уровней входных и выходных данных к рабочему диапазону внутренних переменных сети;
- установить сеть в общее положение, позволяющее достигнуть глобального минимума;
- минимизировать вероятность паралича сети при обеспечении достаточной чувствительности к вариации параметров.

Для достижения указанных целей предлагается методика, основанная на двух эвристических принципах:

1) Принцип максимальной чувствительности. После статической настройки нейронная сеть должна обладать максимальной чувствительностью к вариации любого синаптического веса.

2) Принцип равного влияния. Для каждого нейрона сети изменение координат входных векторов в диапазоне истинных значений должно приводить к равному изменению выходного уровня этого нейрона.

Кроме того, при статической настройке необходимо обеспечить минимальную вероятность паралича сети. Состояние паралича наблюдается, когда аргументы функций активации настолько велики, что выход нейрона практически не чувствителен к изменению отдельного синаптического веса. В статье рассматривается алгоритм статической настройки для ядерных нейронных сетей [2], но в равной мере он может быть использован для обычных полносвязанных нейронных сетей, поскольку они также принадлежат к классу ядерных сетей. На рис.1 показан пример трехслойной ядерной нейронной сети с ограниченными связями, где каждой вершине графа соответствует одиночный нейрон. В ядерных сетях нейроны объединяются в группы - нейронные ядра. Постулируется, что рецепторные поля нейронных ядер не пересекаются. Для ядерных нейронных сетей можно построить укрупненное графическое представление, в котором каждой вершине соответствует нейронное ядро. Такое представление называется структурной моделью ядерной нейронной сети.

На рис.2 показана структурная модель соответствующая нейронной сети приведенной на рис.1. На структурном уровне каждое нейронное ядро характеризуется парой чисел (p, g) , где p - размерность рецепторного поля, g - размерность аксонового поля нейронного ядра. Каждая

связь на структурной модели соответствует оператору межядерного проектирования и характеризуется рангом оператора, который изображен на графе как вес дуги.

На рис.3 показана структурная модель классической полносвязанной нейронной сети.

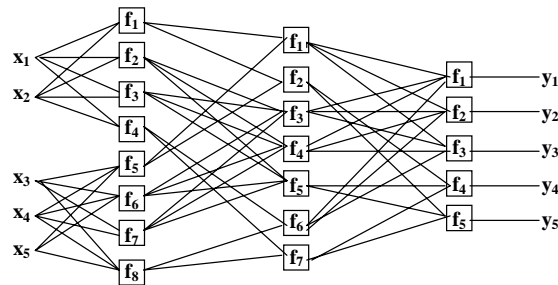


Рисунок 1

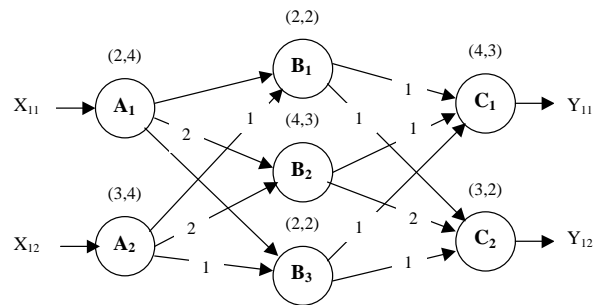


Рисунок 2

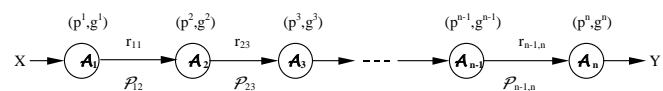


Рисунок 3

ФУНКЦИИ АКТИВАЦИИ

Формальный нейрон, который обычно используется в искусственных нейронных сетях, представляет собой сумматор взвешенных координат, выход которого поступает на нелинейный элемент с пороговой функцией активации. Математическая модель нейрона может быть представлена следующим образом

$$s = \sum_u x(u)w(u) + w(*), \quad y = f(s),$$

где u - номер входного рецептора, $f()$ - нелинейная функция активации, $w(*)$ - смещение функции активации. Используются различные функции активации, но наибольшее распространение получили сигмоидная и

гиперболический тангенс. Сигмоидная функция определяется аналитическим выражением

$$f(s) = \frac{1}{1 + e^{-\lambda s}}$$

и имеет граничные значения $f(-\infty) = 0$, $f(0) = 1/2$, $f(\infty) = 1$. Производная этой функции имеет вид

$$\frac{df}{ds} = \frac{\lambda \cdot e^{-\lambda s}}{(1 + e^{-\lambda s})^2} = \lambda \cdot (1 - f(s)) \cdot f(s), \quad f'(0) = \frac{\lambda}{4}.$$

Целесообразно выбрать значение $\lambda = 4$, в этом случае $f'(0) = 1$. На рис.4 показаны графики сигмоидной функции с различным значением параметра λ .

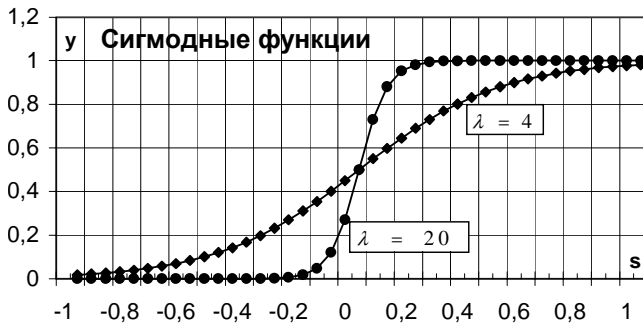


Рисунок 4

В нейронных сетях используемых для управления технологическими объектами в качестве функции активации часто используется гиперболический тангенс. В аналитическом виде эта функция имеет вид

$$f(s) = \text{th}(\lambda \cdot s) = \frac{e^{2\lambda s} - 1}{e^{2\lambda s} + 1}$$

с граничными значениями $f(-\infty) = -1$, $f(0) = 0$, $f(\infty) = 1$. Производная от этой функции будет равна

$$\frac{df}{ds} = \frac{4\lambda e^{2\lambda s}}{(e^{2\lambda s} + 1)^2} = \lambda(1 - f^2(s)), \quad f'(0) = \lambda.$$

Целесообразно выбрать значение $\lambda = 1$, в этом случае $f'(0) = 1$. На рис.5 показаны графики функции гиперболического тангенса с различным значением параметра λ . Там же показана линейная функция с угловым коэффициентом равным 1. Линейная функция, как правило, используется в выходном слое нейронной сети и служит для согласования с истинными уровнями выходных данных.

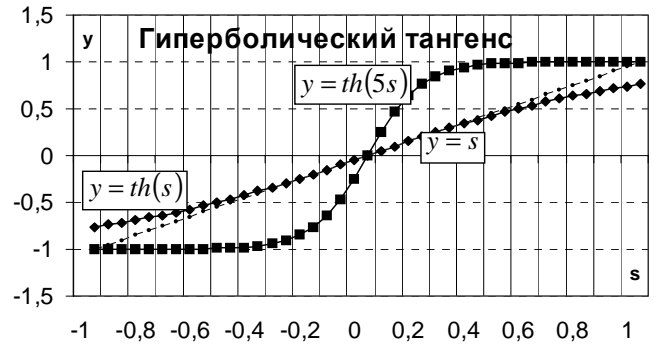


Рисунок 5

АЛГОРИТМ СТАТИЧЕСКОЙ НАСТРОЙКИ

Обозначим весовые коэффициенты нейронных ядер через $w_i^m(u, v)$, где i - номер ядра в слое m . Обработка данных для нейронного ядра определяется выражением

$$p_i^m = \sum_{u=1}^{p_i^m} x_i^m(u) w_i^m(u, v) + w_i^m(*, v), \quad (1)$$

$$y_i^m(v) = f(s_i^m(v)),$$

где $x_i^m(u)$ - входные переменные, p_i^m - число рецепторов в ядре, u - номер рецептора, v - номер нейрона.

СТАТИЧЕСКАЯ НАСТРОЙКА ВХОДНОГО СЛОЯ

Рассмотрим некоторое ядро входного слоя ($m = 1$). Для упрощения записи временно опустим идентифицирующие индексы слоя и номера ядра, тогда выражение (1) примет вид

$$s(v) = \sum_{u=1}^p x(u)w(u, v) + w(*, v), \quad y(v) = f(s(v)). \quad (2)$$

Предположим, что значения весовых коэффициентов $w(u, v)$ уже определены и пусть $\bar{x}(u)$ среднее значение по обучающей выборке для координаты u входного вектора. Для рассмотренных функций активации максимум производной достигается когда $s(u) = 0$, поэтому, следуя принципу максимальной чувствительности из (2) получим следующее правило установки уровня смещения:

$$w(*, v) = - \sum_{u=1}^p \bar{x}(u)w(u, v). \quad (3)$$

В этой формуле в качестве входных координат использованы средние значения, поэтому настройка выполняется по среднему вектору входного обучающего

множества.

Найдем теперь правило настройки синаптических весов ядра. Пусть числа $\Delta x(u)$ определяют ширину диапазона изменения входной переменной. Следуя принципу равного влияния необходимо принять

$$|\Delta x(u)w(u, v)| = \Delta s, \tag{4}$$

где Δs постоянное значение. Если функции активации в пределах ядра одинаковы, то Δs имеют совпадающие значения для всех нейронов ядра. Чем больше значение Δs , тем выше чувствительность сети к вариации параметров, но с возрастанием Δs увеличивается вероятность того, что аргумент s достигнет критической области, когда сеть может попасть в состояние паралича. Поэтому выбор значения Δs должен удовлетворять некоторому компромиссу между чувствительностью и склонностью сети к параличу. Исследуем это вопрос более подробно.

Будем рассматривать входные переменные нейронной сети как случайные величины с равномерным законом распределения в диапазонах $\Delta x(u)$. Подставляя (3) в (2) приходим к выражению

$$s(v) = \sum_{u=1}^p (x(u) - \bar{x}(u))w(u, v) + w(*, v) = \sum_{u=1}^p \xi(u, v).$$

Выполнение условия (4) приводит диапазоны изменения всех переменных $\xi(u, v)$ к значению Δs , поэтому аргумент u можно опустить, полагая $\xi(u, v) = \xi(v)$. Случайная величина $\xi(v)$ является центрированной с нулевым математическим ожиданием и равномерным законом распределения. Дисперсия случайной величины $s(v)$ будет удовлетворять неравенству

$$D[s(v)] \leq pD[\xi(v)] = p \frac{(\Delta s)^2}{12}. \tag{5}$$

Если случайные величины $\xi(v)$ (для различных значений v) независимы, тогда неравенство следует заменить равенством. Математическое ожидание случайной величины $s(v)$ очевидно равно нулю. По теореме Чебышева вероятность того, что случайная величина $s(v)$ отклонится от своего математического ожидания на величину больше s_0 , равна

$$P(s \geq s_0) \leq \frac{D[s(v)]}{(s_0)^2}. \tag{6}$$

Для практических целей значение вероятности можно выбирать в диапазоне $0,5 \pm 0,1$. Выбор величины s_0 зависит от допустимого значения производной функции активации. Это величина определяет область значений

переменной s , в пределах которой гарантируются отсутствие паралича сети. В таблице приведены значения s_0 для различных функций активации при допустимом значении $df(s_0)/ds = 0,5$. Из выражений (5) и (6) при $P = 0,083$ следует

$$1 \leq p(\Delta s/s_0)^2.$$

Таблица

| Тип функции $f(s)$ | s_0 | $f(s_0) - f(-s_0)$ |
|---------------------------------------|-------|--------------------|
| Сигмоидная $\lambda = 4$ | 0,45 | 0,72 |
| Сигмоидная $\lambda = 20$ | 0,17 | 0,94 |
| Гиперболический тангенс $\lambda = 1$ | 0,85 | 1,38 |
| Гиперболический тангенс $\lambda = 5$ | 0,37 | 1,92 |

Отсюда можно получить значение Δs . Далее используя (4) получим следующую расчетную формулу

$$|w(u, v)| = \frac{s_0}{\Delta x(u)} \frac{1}{\sqrt{p}}. \tag{7}$$

Знак синаптического веса следует выбирать случайным образом с равновероятным законом распределения. Это приводит сеть в общее положение и дополнительно рандомизирует входные переменные. После нахождения синаптических весов, смещение функций активации определяется по формуле (3).

СТАТИЧЕСКАЯ НАСТРОЙКА СКРЫТЫХ СЛОЕВ

Для определенности будем полагать, что в скрытых слоях используются тангенциальные функции активации с $\lambda = 1$. В этом случае $\Delta x(u) = 2$. Подставив в (7) получим следующую расчетную формулу

$$|w_i^m(u, v)| = \frac{s_0}{2} \frac{1}{\sqrt{p_i^m}}.$$

Эта формула справедлива в предположении, что входные величины $x_i^m(u)$ распределены равномерно в диапазоне $[-1 \div +1]$. Фактически о законе распределения можно сказать только, что он симметричен и что с большой вероятностью значения переменных лежат в диапазоне $[f(-s_0) \div f(s_0)]$. Первое утверждение следует из того, что знак синаптических весов в предшествующих слоях выбирается случайным образом с равной вероятностью, а второе следует из условия минимизации вероятности паралича сети которое также реализовано в предшес-

твующих слоях. Наиболее вероятная ширина диапазона переменных $x_i^m(u)$ для различных функций активации представлена в третьем столбце таблицы. Эти значения можно использовать в формуле (7) которая примет вид

$$|w_i^m(u, v)| = \frac{s_0}{f(-s_0) - f(s_0)} \frac{1}{\sqrt{p_i^m}}.$$

Уточненная формула позволяет увеличить начальную чувствительность сети к изменению параметров, сохранив при этом достаточно низкую вероятность паралича.

Знак значения синаптического веса, как и прежде, выбирается случайным образом. После того как синаптические веса определены, значения смещения функций активации вычисляются по формуле

$$w_i^m(*, v) = - \sum_{u=1}^{p_i^m} x_i^m(u) w_i^m(u, v).$$

Значения $x_i^m(u)$ определяются непосредственно из сети.

Для этого последовательно выполняется послойная обработка входных векторов в предшествующих слоях. Это всегда можно сделать, поскольку веса в предшествующих слоях уже определены оптимальным образом, а для самого первого слоя значения входного вектора определяются средним вектором входного обучающего множества.

СТАТИЧЕСКАЯ НАСТРОЙКА КОНЕЧНОГО СЛОЯ

Для того чтобы получить истинные значения выходных переменных, функции активации в последнем слое должны быть линейными. Для выходных переменных считается известными значения средних $\bar{y}(v)$ и их диапазон изменения. Эти величины могут быть получены на основе обучающей выборки. Полагая, что линейная функция активации имеет производную равную единице, выходные переменные будут определяться выражением

$$y(v) = \sum_{u=1}^p x(u)w(u, v) + w(*, v) \quad (8)$$

(В последнем выражении индексы идентификации слоя и нейронного ядра для упрощения записи опущены). Предположим, что синаптические веса уже установлены. Тогда на основании последнего выражения значения смещений могут быть найдены по формуле

$$w(*, v) = \bar{y}(v) - \sum_{u=1}^p x(u)w(u, v). \quad (9)$$

Значения $x(u)$ - предпоследнего слоя определяются также как и для скрытых слоев. Определим теперь значения синаптических весов. Из выражений (8) следует

$$D[y(v)] \leq \sum_{u=1}^p (w(u, v))^2 D[x(u)]. \quad (10)$$

Равенство выполняется в том случае, когда случайные величины $x(u)$ независимы. Если все функции активации в предпоследнем слое одинаковы, то переменные $x(u)$ имеют совпадающие диапазоны изменения. В предположении равномерного закона распределения эти случайные величины имеют и одинаковую дисперсию, которую обозначим D_x . Следуя принципу равного влияния в этом случае необходимо принять $w(u, v) = w(v)$. Тогда неравенство (10) можно переписать в виде

$$D[y(v)] \leq p(w(v))^2 D_x. \quad (11)$$

Поскольку при равномерном законе распределения $D[y(v)] = (\Delta y(v))^2/12$ и $D_x = (\Delta x(v))^2/12$ то из (11) следует следующая расчетная формула

$$|w(v)| = \frac{\Delta y(v)}{\Delta x \sqrt{p}}.$$

Ширина диапазона Δx для тангенциальных функций равна 2, но может быть несколько уменьшена по тем же основаниям, что и для скрытых слоев. Знаки синаптических весов, как и прежде, выбираются случайным образом. Смещения функций активации определяются, следуя формуле (9).

После процедуры начальной установки нейронная сеть настроена таким образом, что при векторе средних на входе, выходной вектор сети совпадает с вектором средних выходного множества обучающей выборки. Полученные формулы предварительной установки весов носят приближенный характер, поскольку точные законы распределения переменных состояния сети неизвестны. Аппроксимация этих распределений равномерным законом и использование неравенства Чебышева для оценки допустимого интервала обеспечивает, возможно, избыточный запас устойчивости сети к параличу, что ведет к снижению начальной параметрической чувствительности. Можно сказать, что полученные расчетные формулы устанавливают нижние границы значений синаптических весов гарантирующие обучаемость сети с минимальной вероятностью паралича. При сильно коррелированных данных значения синаптических весов можно увеличить, вводя в расчетные формулы поправочные множители.

ПЕРЕЧЕНЬ ССЫЛОК

1. Горбань А.Н., Россиев Д.А. Нейронные сети на персональном компьютере. Новосибирск: Наука, 1996. 276с.
2. Дорогов А.Ю., Алексеев А.А. Математические модели быстрых нейронных сетей. // Сб. научн. тр. СПбГЭТУ "Системы управления и обработки информации". Вып.490, 1996, С.79-84.

ВОЗМОЖНОСТИ ПРИМЕНЕНИЯ НЕЙРОННЫХ СЕТЕЙ В ЗАДАЧАХ ПРОГНОЗИРОВАНИЯ И СОБЛЮДЕНИЕ ПРИНЦИПА СОГЛАСОВАННОСТИ ПОЛУЧЕННЫХ РЕЗУЛЬТАТОВ

А.С.Дулесов

Представлены теоретические аспекты возможности использования нейронных сетей в задачах прогнозирования. Дано понятие принципа согласованности результатов и модель его соблюдения.

The theoretical aspects of the opportunity of the neuron network usage in the problems of prognostication are represented. The notion of the principal of the observation results coordination and model of its observance are given.

ВВЕДЕНИЕ

В практике прогнозирования одной из основных задач ставилась возможность на основе имеющейся информации построить (оценить) точно или приближенно зависимость влияния как внешних, так и внутренних факторов на характеристики объекта. Процесс такой оценки является моделированием и в большинстве случаев опирается на наличие статистических данных и теорию прогнозирования.

1. ОСНОВНЫЕ ПОНЯТИЯ СТАТИСТИЧЕСКОГО ОЦЕНИВАНИЯ

Под объектом исследования, в каком-то смысле однородном (например, предприятие, технологический процесс, отрасль и т.д.), будем понимать реально существующий объект, который в некоторый фиксированные моменты времени может занимать определенное состояние. Поведение объекта описывается через свойства (например, "мощность", "скорость", "стабильность" и др.), отражаемые через характеристики.

Конечной целью наблюдения за объектом можно считать получение информации относительно некоторого конечного набора X_1, X_2, \dots, X_n его характеристик. Зависимость между характеристиками Y и X объекта моделируется обычно в виде функции $Y = f(X)$. В данном равенстве функция f , также может рассматриваться как характеристика, но уже не объекта, а совокупности объектов. Тогда под f следует понимать некоторый класс функций, объединенных какими-либо свойствами (свойством линейности или свойством принадлежности к модели одного и того же объекта). По сути, мы имеем дело с задачей возможно более точной оценки функции f по данным наблюдений (обычно неполным или неточным) отдельных объектов. В конечном итоге речь идет о построении "истинной" зависимости между показателями (числовыми характеристиками) объекта и рассчитанной зависимостью,

найденной в ходе решения поставленной задачи.

В данной работе представлены возможности моделирования зависимости $Y = f(X)$ с использованием нейронной сети прямого распространения, а также дана оценка эффективности его применения.

О методах построения зависимости $Y = f(X)$ можно прочесть в любом учебнике по статистике. Основным инструментом здесь принято считать метод наименьших квадратов. Однако дела обстоят не так просто, как может показаться на первый взгляд. Так в [1] обобщен опыт в сфере разработки и совершенствования моделей построения статистических зависимостей и изложен принцип максимальной согласованности. Учитывая всю сложность задачи, обусловленную наличием различного рода ошибок, возникает вопрос: способна ли нейронная сеть дать согласованную оценку, как наилучшую из приемлемых.

Наибольшее распространение в практике построения зависимостей получили методы, позволяющие строить зависимости вида:

$$Y = f(X_1, \dots, X_n, \xi).$$

Здесь характеристика объекта Y зависит от характеристик (факторов) X_1, \dots, X_n с некоторой долей уверенности известные исследователю и "прочих факторов" $X_{n+1}, \dots, X_{n+2}, \dots$, агрегируемых в один неопределенный фактор ("оценку наблюдения"). Искомые зависимости относятся (чаще всего) к классу параметрических, т.е. имеют вид известной функции $Y = f(X, a)$ от объясняющей (векторной) переменной X и конечномерного вектора неизвестных параметров a .

Механизм построения зависимости $Y = f(X, a)$ с помощью аналитических методов достаточно хорошо проработан. Если вести речь об аналогичной задаче, но в качестве инструмента использовать нейронные сети, то, прежде всего, следует задаться вопросом: возможно ли представление функции многих переменных с помощью функции одного или нескольких переменных. Оказалось, что, рассматривая класс непрерывных функций, разрешение данного вопроса возможно. Обращая внимание на функцию $Y = f(X, a)$, следует говорить о точном ее воспроизведении (аппроксимации). Следовательно, решение задачи направлено на вычисление функций, имеющих наилучшее приближение к истинной. Теорема Вейерштрасса утверждает, что непрерывную функцию нескольких пере-

менных $f(X_1, \dots, X_n)$ на замкнутом ограниченном множестве Ω можно равномерно приблизить последовательностью полиномов: для любого положительно малого значения ε существует такой многочлен $F(X_1, \dots, X_n)$, что

$$\max_{\Omega} |f(X_1, \dots, X_n) - F(X_1, \dots, X_n)| < \varepsilon.$$

О возможностях представления функций нескольких переменных функцией одной переменной сказано в [2]. Будем следовать логике рассуждений, представленных в данной работе.

2. ПРЕДСТАВЛЕНИЕ ФУНКЦИЙ С ПОМОЩЬЮ НЕЙРОННОЙ СЕТИ

Кроме аппроксимации функций многочленами широкое применение в настоящее время находят способы приближения функций многих переменных с помощью линейных операций и суперпозиций функций одного или нескольких переменных. Опорой здесь являются нейронные сети. Из всего многообразия существующих ныне моделей нейронной сети возьмем за основу сеть прямого распространения (персептрон). Имеющая многослойную структуру, состоящую из нейронов, она способна решать задачи оценивания и аппроксимации функций по нескольким точкам (примерам) и по цепочке "вход-выход". В дальнейшем с их помощью задача прогноза сводится к получению неизвестных значений на выходе сети по известным входам.

Нейронные сети позволяют вычислять функции одного переменного и суперпозиции (функции от функций). Возможности нейронных сетей прямого распространения сколь угодно точно воспроизводить $Y = f(X)$ и их принцип работы представлен в [3].

Архитектура нейронной сети, в общем виде, состоит из входного, выходного (один выходной нейрон) и нескольких внутренних слоев. На вход подаются факторы X_1, \dots, X_n , на выходе - получают Y . С выходного нейрона снимается сигнал, представляющий собой вычисляемые этой сетью функции от входных сигналов. В общем виде, она представляется как

$$Y(X, a) = \Phi_l \sum_{i=1}^k \left(\Phi_{l-1, i} \left(\dots \left(\sum_{i=1}^k \left(\Phi_{1i} \left(\sum_{j=1}^n (a_{1ij} X_j) \right) \right) \dots \right) \right) \right),$$

$$(j = \overline{1, n}), (i = \overline{1, k}).$$

Здесь Σ - адаптивный линейный сумматор, вычисляющий скалярное произведение вектора входного сигнала (на первом слое - X) на вектор параметров (вес синапса) a ; Φ - нелинейный преобразователь, преобразующий выходной сигнал сумматора и являющийся функцией нескольких переменных $\Phi = (a, X_1, \dots, X_n)$; n - число переменных; k - число функций Φ ; l - число слоев в нейронной сети.

Возвращаясь к вопросу о возможностях нейронной сети как аппарата для построения зависимости $Y = f(X)$, следует обратиться к обобщенной аппроксимационной теореме Стоуна. Ее трактовка является утверждением об универсальных аппроксимационных свойствах любой нелинейности [3]: "с помощью линейных операций и каскадного соединения можно из произвольных нелинейных элементов получать любой требуемый результат с любой первоначально заданной точностью".

3. ОЦЕНКА И СОГЛАСОВАННОСТЬ ЗАВИСИМОСТЕЙ

Точность оценки неизвестных зависимостей нередко оценивается следующим образом. Пусть истинные значения показателей Y и X связаны зависимостью $Y = f(X)$, а по наблюдаемым значениям (X, Y) построена зависимость $Y = \Phi(X)$. Точность восстановления истинного значения может быть охарактеризована показателем отклонения U_t . Обычно пользуются абсолютным отклонением $U_t = |Y_t - \Phi(X_t)|$ и относительными отклонениями $U_t = |Y_t / \Phi(X_t) - 1|$ или $U_t = |\Phi(X_t) / Y_t - 1|$. Использование этих отклонений оправдано в случаях, когда объясняемая характеристика Y наблюдается с ошибками. Однако определение U_t теряет свое предназначение, если ошибки присущи наблюдениям объясняющих переменных. Надо отдать должное тому, что оценка точности истинного значения с помощью U_t не является единственной. Далее в работе будет представлен алгоритм оценки адекватности полученной зависимости с истинной. При любом способе оценки неизвестной зависимости приходится учитывать требования согласованности полученной зависимости с имеющейся информацией. Особая роль здесь отводится наличию различного рода ошибок. Их разнородность, трудность формализации, накладывают на процесс исследования дополнительные затраты. Выделяют систематические и случайные ошибки. Систематические ошибки постоянны при наблюдениях всех значений показателей, их можно предугадать и в последующем свести к минимуму. Что касается случайных ошибок, вызванных проявлением случайных факторов, то они колеблются и эти колебания трудно предсказуемы. Понятие "случайная ошибка" относительно, поскольку в зависимости от постановки задачи исследования она может иметь "временный" характер. Например, на некоторых интервалах времени влияние отдельных факторов имеет "вялотекущий" процесс и характеризуется как случайная ошибка. Однако под действием сторонних возмущений параметр претерпевает резкие (скачкообразные) изменения и тем самым переходит из разряда "случайных" в разряд значимых. Решение задачи о согласованности, чаще всего, опирается на "общеизвестный" метод наименьших квадратов. Но он нуждается в содержательном обосновании исходя из условий поставленной задачи.

Оценивая полученные зависимости с результатами наблюдений, должен соблюдаться общий принцип макси-

мальной согласованности. В соответствии с ним оценка зависимости должна быть наилучшей с точки зрения критерия, отражающего степень ее "согласованности" с имеющейся информацией. Центральным при этом становится вопрос о разумном выборе измерителя "степени согласованности". В числе одного из них можно предложить подход касающейся оценки статистических гипотез.

4. СТАТИСТИЧЕСКИЕ ГИПОТЕЗЫ О СОГЛАСОВАННОСТИ

В качестве примера согласованности сопоставим два метода, регрессионный и нейросетевой, при помощи которых строятся и воспроизводятся зависимости соответственного вида $Y = f(X, a)$ и $Y = f(X)$. Выбор этих методов не является случайным, поскольку первый широко применяется, и второй находит свое применение в задачах техники, экономики и т.д. Последующий подбор критерия оценки должен явиться дополнением к общей части процесса "согласованности".

Регрессионная зависимость может быть как линейной, так и нелинейной. Нейросетевая модель позволяет лишь получать конечные результаты после ее успешного обучения. Поэтому будем принимать во внимание значения параметров Y рассматриваемых моделей. При этом полагаем, что ошибки являются независимыми нормально распределенными случайными величинами с нулевыми средними значениями и одинаковой дисперсией.

В качестве гипотезы о "согласованности" примем при заданном уровне значимости α , гипотезу о сопоставлении отклонений с весьма малым значением U_0 , т.е.

$H_0: U_t \leq U_0$. Проверка данной гипотезы осуществляется исходя из известного закона распределения Стьюдента, поскольку среднее квадратическое отклонение случайных переменных неизвестно. Предположение о подчиненности случайной величины U_t нормальному закону распределения позволяет воспользоваться статистической характеристикой

$$t_{\text{расч}} = \frac{\bar{U}_t - U_0}{S} \sqrt{T-1},$$

где $\bar{U}_t = \frac{1}{T} \sum U_t$ - среднее отклонение; $S = \frac{1}{T-1} \sum (U_t - \bar{U})^2$ -

среднее выборочное квадратическое отклонение. Согласно распределению Стьюдента (определяется квантиль $t_{\alpha\nu}$, где $\nu = T-1$) при условии $|t_{\text{расч}}| \leq t_{\alpha\nu}$ гипотеза H_0 принимается, в противном случае - отвергается.

Выдвигая гипотезу H_0 и проверяя ее, можно судить о

степени "согласованности" модели с желаемыми результатами.

Выделение такого критерия в качестве меры согласованности обусловлено еще и тем, что он чувствителен к изменениям оцениваемых параметров, т.е. к "выбору общего вида строящейся (воспроизводящейся зависимости)". В этой связи полезно сослаться на указанные в [4] принципы выбора общего вида зависимости:

- максимальное использование априорной информации об анализируемой зависимости;
- предварительный анализ структуры исходных данных;
- возможность упростить искомую зависимость;
- построенная зависимость должна быть, возможно, более устойчива по отношению к той совокупности исходных данных, на основании которых она оценена.

Последний принцип тесно связан с возможностями использования нейронных сетей. Реализация этого принципа предусматривает координацию оценок искомой зависимости, полученных на основании различных комбинаций "подвыборок" факторов воздействия из общего числа наблюдений. Неудачная комбинация выделенных факторов и "подвыборок" приведет к необходимости в отклонении гипотезы H_0 и/или (в худшем случае) к невозможности обучаться. Дать четкие рекомендации по формированию требующихся "подвыборок" в общем случае трудно. Однако можно использовать подход, когда на вход нейронной сети первоначально вводятся более значимые факторы X . После того, как сеть обучится, выполняется проверка на выполнение H_0 . Если в ходе анализа последняя будет отвергнута, то следует добавить очередной (уже менее значимый, чем предыдущие) фактор. Выбор факторов, их уровня значимости - сложная задача и во многом зависит от опыта, "знаний и интуиции исследователя".

ЗАКЛЮЧЕНИЕ

Практика расчетов показала, что нейронные сети прямого распространения являются эффективным инструментом анализа и предсказания поведения объектов исследования. Можно ожидать, что в недалеком будущем реализованные на базе нейронных сетей алгоритмы анализа и прогноза займут достойное место в системе построения зависимостей.

ПЕРЕЧЕНЬ ССЫЛОК

1. Клейнер Г.Б., Смоляк С.А. Экономические зависимости: Принципы и методы построения. - М.: Наука 2000. - 104 с.
2. Горбань А.Н. Функции многих переменных и нейронные сети // Соросовский образовательный журнал, № 12, 1998. С. 105-112.
3. Горбань А.Н., Росснев А.А. Нейронные сети на персональном компьютере. Новосибирск: Наука, 1996. - 256 с.
4. Айвазян С.А., Енюков И.С., Мешалкин Л.Д. Прикладная статистика: Исследование зависимостей: Справ. Изд. М.: Финансы и статистика, 1985. - 185 с.

ДИАГНОСТИКА И КОНТРОЛЬ СОСТОЯНИЯ КОЛЕСА ПЕРВОЙ СТУПЕНИ РКНД АНСАМБЛЕМ НЕЙРОСЕТЕЙ

С.В.Жернаков

В статье рассматривается один из подходов к комплексному диагностированию и контролю состояния колеса первой ступени РКНД на базе гибридного ансамбля нейросетей. Проведен сравнительный анализ нейросетевых ансамблей и обычных нейросетей для решения подобных задач. Приводятся результаты моделирования.

One method of complex diagnostics and checking of the first stage of the low-pressure compressor-working wheel on the base of neurosets hybrid ensemble is shown. Comparison analysis of neurosets ensembles and usual neurosets is fulfilled. The article is illustrated with the results of simulating.

В процессе создания и доводки авиационных газотурбинных двигателей (ГТД) значительная часть экспертов в процессе принятия решений опирается на экспериментальные данные, полученные в ходе стендовых и летных испытаний.

Вибродиагностика ГТД позволяет на ранних этапах контроля выявлять опасные дефекты, способные приводить к отказам авиационных двигателей. Поэтому одной из актуальных задач подобных исследований является решение проблемы идентификации типов колебаний непосредственно в ходе эксперимента. Причем в силу того, что при подходе к области критических режимов работы ГТД градиенты уровней колебаний часто очень велики, поэтому актуальным является создание методов и средств ранней программно-аппаратурной диагностики. Это относится, в частности, к обнаружению изменений дисбалансов роторов и выявлению момента потери аэроупругой устойчивости рабочих колес компрессоров ГТД [1-11].

В данной работе, на базе экспериментальных данных, автором осуществлена диагностика состояния рабочего колеса первой ступени компрессора низкого давления (РКНД) на базе гибридного ансамбля нейросетей (НА), который распознавал и классифицировал отказы.

Известно [1-7,9-11], что в процессе экспериментальных исследований конструкции колес I ступени РКНД ГТД наблюдается устойчивая качественная картина вибронпряженности колеса I ступени РКНД. Частота собственных колебаний (ЧСК) лопаток после проточки больше, чем эти же частоты до проточки. ЧСК лопаток до проточки клеймятся на лопатках. Так бездефектные лопатки могут иметь допуск на ЧСК 2...3 Гц и как правило повышаться на эту величину. Поэтому вблизи максимальных оборотов и цикловых нагрузок на перемычке диска имеют место вибрационные напряжения в лопатках при их колебаниях [1-11]. Колебания лопаток происходят синхронно с некоторой единой час-

тотой для всех лопаток строго на одинаковых оборотах в соответствие с величиной собственной частоты всех лопаток (осредненно), следовательно, величину расположения резонансной точки по оборотам при колебаниях лопаток, можно скорректировать по их средней частоте колебаний. Настройка на расположение резонанса по оборотам зависит от средней величины ЧСК в комплекте. Этот резонанс можно "распространить" на длительные рабочие режимы эксплуатации, например, "Максимал". Статическая цикловая нагрузка на перемычках замков в диске будет максимальной и на нее наложится переменная составляющая от лопаток, которая может повлиять на несущую малоцикловую прочность перемычки диска. Максимальные вибронпряжения реализуются в лопатках I ступени РКНД, имеющих нижние частоты собственных колебаний, а из этой группы лопаток максимальные напряжения на лопатках с частотами на 1...3 Гц выше самой нижней ЧСК в комплекте. Отсюда следует, что если разрушение лопаток, например, связаны с их избыточными колебаниями, то разрушение следует ожидать в первую очередь у лопаток с ЧСК на 1...3 Гц выше самой минимальной частоты.

В ходе исследований экспериментальных данных в табл.1 (изделие В) имело место растрескивание большого количества пазов по замковой части диска с наибольшей трещиной, отвечающей пазу дефектной лопатки 9 с $f = 211$ Гц. При натурных испытаниях, проводимых с лопатками I ступени РКНД, была выявлена закономерность отказов: растрескивание нечетных пазов диска I ступени РКНД значительно чаще, относительно четных пазов [2,4-7,11].

Таблица 1 - Дефекты на РК с трещинами

| Комплект колеса | Fmax, Гц | Fmin, Гц | Fдеф., Гц | F ср., Гц |
|-----------------------|----------|----------|---------------------|-----------|
| Резонансные испытания | 227 | 112 | 214 (лопатка 24) | 221,42 |
| Изделие А | 232 | 216 | 217 (лопатка 9) | 223,4 |
| Изделие В | 231 | 211 | 211 (лопатка 9) | 225,6 |

Принципы, лежащие в основе вибрационных измерений и применяемые для этого процесса программно-аппаратные средства, подробно описаны в классических

работах [1-12]. Однако необходимо отметить следующее: в качестве аппаратных средств диагностирования колебаний лопаток могут применяться как стандартные средства (вибро и тензодатчики) [1-7], так и переносные, портативные средства [6,7,9-12]. Типы и возможности их различны. Некоторые обеспечивают лишь контроль общих уровней вибрации, другие - обработку, анализ и запись колебаний. В последнее время широкое распространение получили цифровые коллекторы данных, которые имеют встроенный микропроцессор, жидкокристаллический дисплей и память для хранения результатов измерения [7,9-12]. Среди распространенных приборов (специализированной аппаратуры) для виброисследований ГТД, его узлов и элементов в наземных условиях часто применяется портативный виброанализатор типа VM-3/Х фирмы "Виброметр".

Принимая во внимание изложенное выше, а также экспериментальный материал, основанный на частотном состоянии лопаток I ступени РКНД (табл.2) (символ d в которой, означает дефектную лопатку), автор применил аппарат нейросетей для их диагностики и контроля.

Таблица 2 - Фрагмент обучающей выборки по 750 РКНД

| N лопатки | Частота собственных колебаний лопаток (Гц) | | | | | | | |
|-----------|--------------------------------------------|-------|------|------|------|------|------|------|
| | 1 | 2 | 3 | 4 | 5 | 6 | 7 | 8 |
| 1 | 240 | 236 | 237 | 238 | 238 | 239 | 239d | 236 |
| 2 | 244 | 242 | 243 | 244 | 242 | 230 | 243 | 240 |
| 3 | 233 | 234 | 237 | 237 | 233 | 236 | 233 | 231 |
| 4 | 244 | 242 | 244 | 245 | 241 | 230 | 244 | 242 |
| 5 | 225 | 226 | 230 | 231 | 239 | 243 | 228 | 227d |
| 6 | 242 | 240 | 242 | 242 | 233 | 233d | 241 | 237 |
| 7 | 231 | 233 | 233 | 233 | 239 | 241 | 236 | 232 |
| 8 | 242 | 240 | 241 | 242 | 234 | 233d | 243 | 238 |
| 9 | 232d | 234d | 232 | 233d | 242 | 241 | 235 | 230 |
| 10 | 242 | 240 | 240 | 242 | 236 | 234 | 242 | 238 |
| 11 | 238 | 234 | 232 | 238 | 242 | 240 | 233d | 230 |
| 12 | 245 | 242 | 244 | 245 | 229 | 229d | 245 | 243 |
| 13 | 230 | 229 | 227d | 230 | 238 | 244 | 231 | 228 |
| 14 | 242 | 240 | 241 | 242 | 234 | 234 | 240 | 238 |
| 15 | 232 | 233 | 232 | 233 | 242 | 241 | 232 | 229 |
| 16 | 242 | 240 | 238 | 241 | 234 | 234 | 242 | 239 |
| 17 | 232 | 235 | 234 | 233 | 245 | 239 | 236 | 229 |
| 18 | 245 | 243 | 245 | 247 | 232 | 227 | 246 | 242 |
| 19 | 230 | 230 | 230 | 231 | 231 | 242 | 228 | 221 |
| 20 | 241 | 239 | 240 | 244 | 240 | 236 | 239 | 239 |
| 21 | 234 | 231 | 232 | 234 | 232d | 244 | 232 | 230 |
| fcp | 237,4 | 236,4 | 227 | 230 | 229 | 229 | 228 | 221 |
| fmin | 230 | 229 | 227 | 230 | 229 | 229 | 228 | 221 |
| fmax | 245 | 243 | 245 | 247 | 245 | 244 | 246 | 243 |

Решение такой задачи на базе нейронных сетей (НС) является перспективной и своевременной. Однако выбор архитектуры НС, полностью отвечающей требованиям качества диагностики и контроля параметров лопаток I ступени РКНД, является самостоятельной проблемой и зависит от специфики решаемой задачи [13]. Так в процессе решения данной задачи необходимо реализовать следующую последовательность действий: фильтрацию данных; прогнозирование развития ситуации; распознавание отказа; классификацию (кластеризацию). Эти задачи тесно взаимосвязаны и их можно разделить функ-

ционально, а они в свою очередь требуют разработки (доработки) соответствующих методов и методик (средства), реализуемых на различных моделях нейросетей. Схематично эта цепочка выглядит следующим образом: задача → функции → методы (средства) → реализация.

При решении этой задачи автором проведены исследования [14-18], показавшие перспективы выбора НС, базирующейся на ансамбле нейросетей (АН) (РБФ - Перцептрон - Кохонена) особенно в задачах большой размерности. Общая декомпозиция ЧСК лопаток I ступени РКНД ГТД, на отдельные составляющие с полной параметрической "привязкой" показывает, что выбор наиболее эффективной архитектуры НА будет иметь следующий вид: первый (входной слой) сеть НС РБФ (фильтрация данных), второй слой - Перцептрон (поле-концентратор, выполняющее роль распознавания ситуации), НС Кохонена - классификатор. При выборе архитектуры НА, автором проведены исследования по ее оптимальности. Обучающая выборка на входе НА (24*750) представляла собой собственные частоты лопаток РКНД (21) и частоты (среднюю, минимальную и максимальную) комплекта (3) из 750 исследованных РКНД. Архитектура НА: 4 НС РБФ по 6 параметров (частот) на входе и 2 - на выходе (состояния "1" - норма и "0" - отказ); НС Перцептрон: 8 параметров на входе и 2 - на выходе (состояние "1" или "0") (после операции контрастирования НС Перцептрон имела архитектуру (8-6-4-2) и НС Кохонена (2 параметра на входе - 1 на выходе)).

Эксперименты проведенные с полноразмерной НС Перцептрон, имеющей архитектуру (24-16-8-4-2) и НА (рис.1) показали преимущества последних. Результаты экспериментов приведены в табл.3 и табл.4.

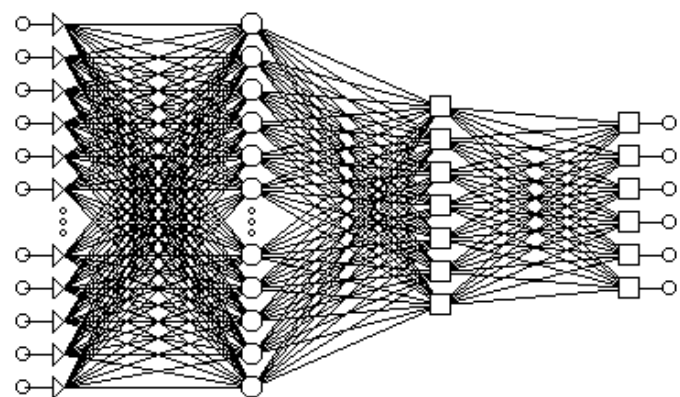


Рисунок 1 - Нейронный ансамбль (РБФ - Перцептрон)

Обучение НС Перцептрон проводилось следующими алгоритмами: back propogation, conjugate gradient, quick propogation, Quasi-Newton, Lewenberg-Marquardt, delta bar delta. Результаты обучения представлены в табл.3.

Таблица 3 - Результаты обучения НС Персептрон

| Алгоритм обучения | Ошибка распознавания персептроном | Количество шагов обучения |
|---------------------|-----------------------------------|---------------------------|
| Back propogation | 0.03267 | 1350 |
| Conjugate gradient | 0.01809 | 1250 |
| Quick propogation | 0.01759 | 1250 |
| Quasi-Newton | 0.01536 | 1300 |
| Lewenberg-Marquardt | 0.01706 | 1350 |
| Delta bar delta | 0.01507 | 1200 |

Таблица 4 - Результаты обучения НА

| Ошибка распознавания (Персептрон-Кохонен) | Ошибка распознавания (РБФ-Персептрон-Кохонен) | Количество шагов обучения |
|-------------------------------------------|-----------------------------------------------|---------------------------|
| 0.01465 | 0.01184 | 1550 |

Результаты моделирования показали, что процесс обучения НА уступает по времени процессу обучения обычной, полноразмерной НС, но эффективность при решении задач контроля, диагностики и прогнозирования состояния лопаток I ступени РКНД ГТД значительно выше (рис.2).

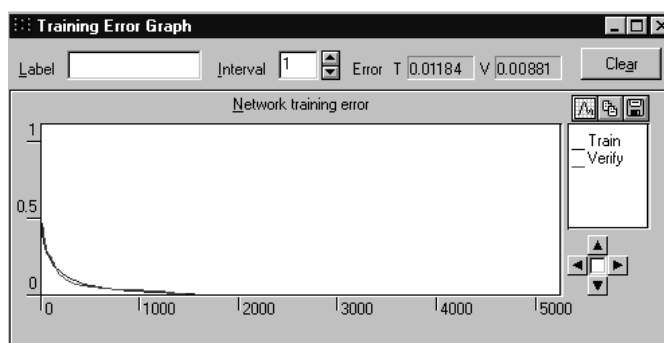


Рисунок 2 - Ошибка обучения и распознавания комбинированной НС

Для расширения диапазона диагностирования I ступени РКНД в условиях частичной или полной неопределенности контролируемых параметров на вход гибридного нейронечеткого ансамбля НС подаются функции принадлежности этих параметров, таким образом при решении задач контроля, диагностики и прогноза НС охватывает более широкий спектр протекающих в лопатках I ступени РКНД ГТД процессов и как следствие этого точность прогнозирования отказов увеличивается.

Программный комплекс (ПК) Трајан 4.0 является мощным нейроимитатором, позволяющим создавать НС любой конфигурации и сложности; содержит простые и комплексные алгоритмы обучения; позволяет проводить оптимизацию (контрастирование) архитектуры НС с

мощью встроенных процедур - генетических алгоритмов. ПК имеет удобный интерфейс. Созданная в нем НС может быть легко перенесена в приложения пользователя (OLE - технологии); на выходе из среды по желанию программиста формируется готовое приложение и синтезируется программа на С++. ПК совместим со всеми известными базами данных (Oracle, Informix, R-base, Visual Foxpro и т.д.) и работает в различных операционных средах (Windows 95/98/NT/2000, Unix и др.)

Подход, предложенный автором, апробирован в среде ПК Трајан 4.0 и показал высокую эффективность при решении задач контроля, диагностики и прогнозирования параметров состояния лопаток I ступени РКНД ГТД.

ПЕРЕЧЕНЬ ССЫЛОК

1. Акимов В.М. Основы надежности авиационных ГТД. - М.: Машиностроение, 1981.
2. Сачин В.М., Скибин В.А., Шатохин А.Г. Развитие методов аппаратурного анализа при исследовании динамических прочностных параметров ГТД. //Испытания авиационных ГТД. - Уфа: УАИ, 1980.
3. Саранцев В.В. Цели и методы контроля состояния двигателей при эксплуатации по состоянию. //Испытания авиационных ГТД. - Уфа: УАИ, 1986.
4. Карасев В.А., Булыгина М.М., Назаренко Г.М. и др. Развитие методов вибрационной диагностики в обеспечении эксплуатации по техническому состоянию. //Труды ЦИАМ. №10042, 1983.
5. Бауер В.О., Булыгина М.М., Исаев Р.И. и др. Расчетно-экспериментальные исследования с целью повышения эффективности вибрационного контроля технического состояния авиационных двигателей. // Труды ЦИАМ. №10058, 1983.
6. Иванов И.А., Карасев В.А., Кармалита В.А. и др. Разработка цифровых методов и средств бортового контроля вибрации ГТД. // Труды ЦИАМ. №10147, 1984.
7. Дорошко С.М. Контроль и диагностирование технического состояния газотурбинных двигателей по вибрационным параметрам. - М.:Транспорт,1984.
8. Дубравский Н.Г. Система диагностического контроля при стендовых испытаниях авиационных двигателей. //Испытания авиационных двигателей. - Уфа: УАИ, 1985.
9. Лебедев А.В., Зазулин В.А. Некоторые методы безразборной диагностики состояния изделий. //Испытания авиационных двигателей. - Уфа: УАИ, 1985.
10. Зимакова Т.Г., Коростелев Ю.А., Максимов В.П. Бесконтактное измерение вибрационных характеристик лопаток рабочих колес компрессоров ГТД с использованием магнитной регистрации. //Испытания авиационных двигателей. - Уфа: УАИ, 1983.
11. Сидоренко М.К. Исследование диагностических свойств вибросигналов ГТД на основе моделей импульсных процессов. //Испытания авиационных двигателей. - Уфа:УАИ,1988.
12. Васильев В.В., Гуров И.П. Компьютерная обработка сигналов. - Санкт-Петербург:БХВ,1998.
13. Горбань А.Н. Обучение нейронных сетей. - М.:СП Параграф, 1990.
14. Жернаков С.В. Об одной методике построения гибридных экспертных систем диагностики и контроля ГТД в условиях неопределенности. // Прогресс. Качество. Технология: Труды III Конгресса двигателестроителей Украины с иностранным участием. - Харьков: ХАИ,1998.
15. Жернаков С.В. Комбинированные модели нейросетей для диагностики и прогнозирования состояния ГТД. // Нейроинформатика -99. - М.: МИФИ, 1999.
16. Жернаков С.В. Нейросети для диагностики и контроля параметров ГТД в условиях неопределенности. //Тезисы докладов. III Конгресс двигателестроителей Украины. - Харьков: ХАИ, 1998.
17. Жернаков С.В. Прогнозирование комплексных отказов ГТД с использованием комбинированных моделей нейросетей в условиях неопределенности. // Нейроинформатика и ее приложения. VI Всероссийский семинар. - Красноярск: КГТУ, 1998.
18. Жернаков С.В. Диагностика и прогнозирование состояния газотурбинного двигателя гибридными нейронечеткими экспертными системами. //Теория и системы управления. Известия РАН, 1999, №5.

НЕДВІЙКОВЕ КОДУВАННЯ ПОВІДОМЛЕНЬ В КАНАЛАХ З ПОДИНОКИМИ ПОМИЛКАМИ

Ю.П.Жураковський

В работе рассматриваются двоичные коды: с повторением и проверкой по модулю q и итеративные. Описаны операции кодирования и декодирования предложенными кодами, а также представлены их основные характеристики. Приведены примеры кодирования и декодирования предложенными в работе кодами первичных кодовых комбинаций восьмеричного кода.

У роботі розглядаються недвійкові коди: з повторенням і перевіркою за модулем q та ітеративні. Описані операції кодування та декодування запропонованими кодами, а також наведені їх основні характеристики. Наведені приклади кодування та декодування запропонованими у роботі кодами первинних кодових комбінацій вісімкового коду.

There are recurrence, check on module q and iteration unbinary codes are considered in the work. The operations of coding and decoding by the offered codes are described, and also their basic characteristics are submitted. The examples of coding and decoding by the codes, offered in work, of primary code combinations of a code with $q=8$ are given.

1. ВСТУП

Для зменшення кількості помилок до деякого дозвольного рівня, що визначається вимогами до систем передачі даних (СПД) [1], в повідомленнях, які передаються по каналах з досить високим рівнем завад, використовується завадостійке кодування [1,2,3,4].

Останнім часом в СПД, поряд з двійковими кодами, почали широко застосовувати надмірні недвійкові коди, що виявляють та виправляють помилки [1,2,5].

Як відомо, потенційні можливості недвійкових (багато-позиційних, q -їчних) кодів значно вищі за двійкові. Це пояснюється збільшенням швидкості передачі інформації такими кодами, яка має пряму залежність від алфавіту коду q [1].

Тому розвиток теорії недвійкових кодів на даному етапі розвитку техніки передачі даних є досить актуальним з точки зору їх практичного застосування.

З недвійкових систематичних кодів, що виправляють однократні помилки широко відомі: код з багатократним повторенням та узагальнений код Хеммінга (УКХ) [1,5]. Однак, ці коди мають значні недоліки. Так коди з багатократним повторенням має велику надмірність $R_{\text{над}} = r/n$, де r - кількість перевірочних елементів у кодовій комбінації; n - довжина коду, $n = k + r$, тут k - кількість інформаційних елементів у кодовій комбінації, та досить низьку швидкість коду $R = k/n$. На відміну від вищезгаданого коду, УКХ має досить малу надмірність та більш високу швидкість коду, але технічна реалізація кодера/декодера цього коду дещо складніша, що

викликано значно більшою кількістю операцій множення, ділення та додавання, ніж для коду з багатократним повторенням. Така складність реалізації кодера/декодера позначається і на часі виконання операцій кодування та декодування, що зростає, а це у деяких випадках, особливо для систем, які працюють у реальному масштабі часу, просто неприпустимо.

У даній роботі розглядається запропоновані автором недвійкові систематичні коди: з повторенням та перевіркою за $\text{mod } q$ (НСКППМ) та ітеративний.

2. НЕДВІЙКОВИЙ СИСТЕМАТИЧНИЙ КОД З ПОВТОРЕННЯМ ТА ПЕРЕВІРКОЮ ЗА МОДУЛЕМ

У НСКППМ кількість перевірочних елементів визначається як $r = k + 1$. Тоді довжина коду $n = k + r = k + (k + 1) = 2k + 1$. Надмірність коду $R_{\text{над}} = r/n = (k + 1)/n = (k + 1)/(2k + 1)$, а швидкість коду $R = k/n = k/(2k + 1)$.

При побудові НСКППМ перевірочні елементи b_i визначаються як просте повторення інформаційних елементів a_i первинної кодової комбінації, тобто $b_i = a_i$, де $i = \{1, 2, \dots, k\}$. Крім одержаних перевірочних елементів b_i визначається ще додатковий перевірочний елемент b_0 як сума за $\text{mod } q$ всіх інформаційних елементів первинної кодової комбінації

$$b_0 = \sum_{i=1}^k a_i \text{mod } q \text{ або } b_0 = \sum_{i=1}^k b_i \text{mod } q.$$

Цей перевірочний елемент, як правило, розташовують в кінці кодової комбінації НСКППМ.

Кодове слово НСКППМ на виході кодера буде мати вигляд:

$$\bar{X} = \{a_1 a_2 \dots a_i \dots a_k b_1 b_2 \dots b_i \dots b_k b_0\}.$$

Утворена таким чином кодова комбінація НСКППМ поелементно передається у канал зв'язку, де вона може бути спотворена завадою. Не вдаючись у сутність явищ, що призводять до цього, сам процес спотворення зручно подати як

$$\vec{Y} = \bar{X} + \vec{E} = \{a'_1 a'_2 \dots a'_i \dots a'_k b'_1 b'_2 \dots b'_i \dots b'_k b'_0\},$$

де \bar{Y} - спотворена кодова комбінація на виході каналу зв'язку; \vec{E} - n -мірний вектор помилки з єдиною ненульовою компонентою e (передбачається однократна помилка), де e може приймати значення $1 \leq e \leq q-1$; a' та b' - відповідно інформаційні та перевірочні елементи спотвореної кодової комбінації.

Код може застосовуватися або для виявлення, або для виправлення помилок. НСКППМ виявляє всі одно-, дво- та трикратні помилки або виправляє всі однократні помилки.

При виявленні помилок НСКППМ може виявляти і деякі помилки кратності більше трьох. Однак, чотирикратні, так звані "дзеркальні" помилки код виявити не може. Як відомо, до "дзеркальних" відносять помилки в одному і тому ж розряді в інформаційній та перевірочній частинах прийнятої кодової комбінації.

При використанні НСКППМ для виявлення помилок у декодері виконуються такі операції:

- порозрядне порівняння інформаційної та перевірочної частин прийнятої кодової комбінації: $a_1' \leftrightarrow b_1'$, $a_2' \leftrightarrow b_2'$,
 \dots , $a_i' \leftrightarrow b_i'$, \dots , $a_k' \leftrightarrow b_k'$;

- визначення кодового синдрому:

$$s_1 = \left(\sum_{i=1}^k a_i' \oplus b_0' \right) \bmod q =$$

$$= (a_1' \oplus a_2' \oplus \dots \oplus a_i' \oplus \dots \oplus a_k' \oplus b_0') \bmod q;$$

$$s_2 = \left(\sum_{i=1}^k b_i' \oplus b_0' \right) \bmod q =$$

$$= (b_1' \oplus b_2' \oplus \dots \oplus b_i' \oplus \dots \oplus b_k' \oplus b_0') \bmod q.$$

Перша перевірка виявляє всі одно-, дво- і трикратні помилки та помилки більшої кратності, якщо при порівнянні є незбіг у деяких розрядах інформаційної та перевірочної частин прийнятої кодової комбінації. Двократна "дзеркальна" помилка при першій перевірці не виявляється, тому що при цьому, як і при відсутності помилок, буде повний збіг інформаційних і перевірочних елементів у всіх розрядах.

Одержаний при другій перевірці кодовий синдром вказує на відсутність або наявність помилки в інформаційній та перевірочній частинах прийнятої кодової комбінації. Якщо $s_1 = 0$ і $s_2 = 0$, то помилок немає. При $s_1 \neq 0$ помилки є в інформаційній, при $s_2 \neq 0$ - в перевірочній частині прийнятої кодової комбінації НСКППМ, а при $s_1 \neq 0$ та $s_2 \neq 0$ - помилки є в інформаційній та перевірочній частинах прийнятої кодової комбінації або спотворений додатковий перевірочний елемент b_0' .

Таким чином можуть бути виявлені, як було вказано

вище, всі одно-, дво- і трикратні помилки в прийнятій кодовій комбінації.

При використанні НСКППМ для виявлення помилок прийнята комбінація, в якій виявлені декодером хоча б одна помилка, бракується і по зворотному каналу зв'язку СПД передається запит на повторення цієї комбінації.

При використанні НСКППМ тільки для виправлення у прийнятій кодовій комбінації однократних помилок у декодері виконують, крім вищезгаданих операцій, які виконувались декодером у процесі виявлення помилок, операцію виправлення помилки. Для цього, якщо помилка виникла в інформаційній частині прийнятої кодової комбінації, на що вказує значення синдрому $s_1 \neq 0$, виконують або виправлення спотвореного елемента інформаційної частини прийнятої кодової комбінації: $a_i = (a_i' \oplus s_1) \bmod q$ або подають одержувачу повідомлень як інформаційну неспотворену перевірочну частину прийнятої кодової комбінації. У разі спотворення додаткового перевірочного елемента b_0' кодовий синдром $s_1 \neq 0$, $s_2 \neq 0$ виправлення b_0' не виконують, а одержувачу повідомлень видають неспотворену інформаційну частину прийнятої кодової комбінації.

3. НЕДВІЙКОВІ ІТЕРАТИВНІ КОДИ

Недвійкові ітеративні коди мають більш високу здатність виявляти помилки у порівнянні з аналогічними двійковими ітеративними кодами. При $q > 2$ зростає обсяг інформації, що передається, завдяки тому, що кількість інформації, яка міститься в одному елементі кодової комбінації, визначається алфавітом коду q .

Найбільш просто реалізуються двомірні ітеративні коди, у яких виконується кодування попередньо сформованого інформаційного масиву у вигляді блока по рядках та стовпцях недвійковим кодом з перевіркою за $\bmod q$. Такі коди мають мінімальну кодову відстань $d_{\min} = 4$ і можуть виявляти одно-, дво- та трикратні помилки і помилки більшої кратності, за виключенням деяких 4-, 6- та 8-кратних помилок, коли помилки розташовуються у вершинах прямокутників або попарно у визначеному порядку. У режимі виправлення помилок код виправляє будь-які однократні помилки.

Таким чином, перевірочні елементи у таких ітеративних кодах визначаються як результат перевірки за $\bmod q$ по рядках і стовпцях.

У режимі виявлення помилок декодером виконуються операції за аналогією з двійковим кодом - визначення контрольних елементів за допомогою перевірки за $\bmod q$ кожного рядка та стовпця одержаного інформаційного масиву. При відсутності помилок у рядку (стовпці) контрольний елемент буде дорівнювати "0". У разі виникнення помилки у масиві елементів ітеративного коду, що надходить до декодера з каналу, одержані у декодері шляхом обчислень контрольні елементи для рядків (стовпців), де виникають помилки, не дорівнюють "0".

Виправлення спотвореного елемента виконують таким чином. Якщо не виконується перевірка для i -го рядка і j -го стовпця, то елемент, що знаходиться на перетинанні i -го

рядка і j -го стовпця, замінюють елементом, який є сумою за $\text{mod } q$ даного прийнятого елемента (помилкового) та контрольного елемента i -го рядка (або j -го стовпця), який був одержаний у декодері.

При виникненні декількох помилок у одному рядку (стовпці), помилки виправляють послідовно для тих стовпців (рядків), де вони є поодинокими.

Надмірність та швидкість ітеративного коду у значній мірі залежить від обсягу інформаційного блока, який формується у кодері. Рекомендовані розміри блока $k \times k$ дорівнюють 12, 16 та 56. При цьому кількість перевірочних елементів r в ітеративному коді на виході кодера буде дорівнювати відповідно: 8, 9 та 16.

Слід відзначити, що кратність помилок, що виправляються недвійковими кодами вища у порівнянні з двійковими, через те, що виправлення одного елемента q -коду відповідає виправленню $\log_2 q$ двійкових одиниць.

Зазначимо, що недвійкові коди прийнято ділити на дві великі групи: коди з простою основою $q = p$, де $p \in \{3, 5, 7, 11, 13, \dots\}$, і коди з основою q , що розкладається. Найбільший практичний інтерес має тільки окремий випадок таких кодів при $q = 2^h$ символи якого мають інформаційну ємність 1 біт і можуть бути співставленні з усіма h -розрядними двійковими числами. Вибір q впливає на визначення операцій додавання, віднімання, множення і ділення, які використовуються у процедурах кодування і декодування. Якщо основа просте число, зручно використати апарат обчислень за модулем простого числа. Якщо ж $q = 2^h$ то необхідно звернутися до алгебраїчного апарату обчислень за модулем незвідного полінома. Символи коду при цьому ставлять у відповідність елементам скінченного поля порядку q .

Розглянемо на прикладах процес кодування та декодування первинних кодових комбінацій запропонованими недвійковими кодами.

Таблиця 1

| \oplus | 0 | 1 | 2 | 3 | 4 | 5 | 6 | 7 |
|----------|---|---|---|---|---|---|---|---|
| 0 | 0 | 1 | 2 | 3 | 4 | 5 | 6 | 7 |
| 1 | 1 | 0 | 3 | 2 | 5 | 4 | 7 | 6 |
| 2 | 2 | 3 | 0 | 1 | 6 | 7 | 4 | 5 |
| 3 | 3 | 2 | 1 | 0 | 7 | 6 | 5 | 4 |
| 4 | 4 | 5 | 6 | 7 | 0 | 1 | 2 | 3 |
| 5 | 5 | 4 | 7 | 6 | 1 | 0 | 3 | 2 |
| 6 | 6 | 7 | 4 | 5 | 2 | 3 | 0 | 1 |
| 7 | 7 | 6 | 5 | 4 | 3 | 2 | 1 | 0 |

Операції додавання за $\text{mod } 8$ будемо виконувати за допомогою таблиці 1, у якій наведені результати операцій додавання у скінченному полі $GF(2^3) = GF(8)$. Операцію додавання \oplus двох елементів виконано як порозрядне додавання за $\text{mod } 2$ їх двійкових еквівалентів з подальшим

записом у таблицю одержаного результату у вісімковій системі числення.

Наприклад, якщо треба додати $5 \oplus 7$, то виконують додавання їх двійкових еквівалентів за $\text{mod } 2$: $101 \oplus 111 = 010$ і одержаний результат з двійкової системи числення переводять у вісімкову: $010_2 = 2$, тобто $5 \oplus 7 = 2$.

4. ПРИКЛАДИ КОДУВАННЯ ТА ДЕКОДУВАННЯ

Приклад 1. Виконати процес кодування та декодування первинної кодової комбінації вісімкового коду ($q=8$) довжиною $k=5$: $A=32705$ недвійковим систематичним кодом з повторенням та перевіркою за модулем (НСКППМ).

Розв'язання. У процесі кодування у кодері виконується: по-перше, просте повторення первинної комбінації і, по-друге, визначається додатковий перевірочний елемент b_0 :

$$b_0 = \sum_{i=1}^k a_i \text{mod } q = (3 \oplus 2 \oplus 7 \oplus 0 \oplus 5) \text{mod } 8 = 3.$$

Таким чином, кодова комбінація НСКППМ на виході кодера буде мати вигляд:

$$\bar{X} = 32705 \ 32705 \ 3.$$

Надмірність коду $R_{\text{над}} = r/n = 6/11$; швидкість коду $R = k/n = 5/11$.

Комбінація НСКППМ подається у канал зв'язку де діють завади, які можуть викликати однократні помилки. На вхід приймального пристрою СПД надходить, наприклад, кодова комбінація

$$\bar{Y} = 32745 \ 32705 \ 3,$$

в якій спотворений один інформаційний елемент.

Для виявлення і виправлення однократної помилки у декодері виконуються такі операції:

- порозрядно порівнюються елементи інформаційної та перевірочної частин прийнятої кодової комбінації:

$$32745$$

$$32705,$$

у результаті якого маємо незбіг у четвертому розряді прийнятої кодової комбінації;

- для визначення місця помилки, тобто у якій, інформаційній чи перевірочній, частині прийнятої кодової комбінації виникла помилка, визначається кодовий синдром:

$$s_1 = (3 \oplus 2 \oplus 7 \oplus 4 \oplus 5 \oplus 3) \text{mod } 8 = 4;$$

$$s_2 = (3 \oplus 2 \oplus 7 \oplus 0 \oplus 5 \oplus 3) \text{mod } 8 = 0.$$

Значення $s_1 \neq 0$, що говорить про наявність помилки в інформаційній частині прийнятої кодової комбінації. Для виправлення помилки у декодері визначається сума за $\text{mod}8$ одержаного спотвореного елемента a_4' та s_1 : $a_4 = (a_4' \oplus s_1) \text{mod} q = (4 \oplus 4) \text{mod} 8 = 0$. Після чого декодер виправляє спотворений елемент a_4' і виправлена інформаційна частина прийнятої кодової комбінації НСКШПМ подається на вихід декодера.

За другим варіантом інформаційна частина прийнятої кодової комбінації не виправляється і на вихід декодера подається як інформаційна, неспотворена перевірна частина прийнятої кодової комбінації (що цілком справедливо, бо $s_2 = 0$) без додаткового перевірного елемента b_0' .

Приклад 2. Закодувати вісімковим ітеративним кодом, що виправляє однократні помилки, інформаційну послідовність 2407435144670215. Визначити надмірність коду та показати процес виправлення однократної помилки.

Розв'язання. Для того, щоб ітеративний код виправляв однократні помилки, досить для кодування по стовпцям і рядкам використати код з перевіркою за $\text{mod} q$, тобто у даному разі - за $\text{mod} 8$.

Запишемо задану інформаційну послідовність у вигляді матриці 4×4 та закодуємо кожний стовпець та кожний рядок одержаної матриці кодом з перевіркою за $\text{mod} 8$. Додавання виконуємо згідно таблиці 1:

$$\begin{array}{cccc|c} 2 & 4 & 0 & 7 & 1 \\ 4 & 3 & 5 & 1 & 3 \\ 4 & 4 & 6 & 7 & 1 \\ 0 & 2 & 1 & 5 & 6 \\ \hline 2 & 1 & 2 & 4 & 5 \end{array}.$$

Таким чином, кодована послідовність вісімкового ітеративного коду буде мати вигляд: 2 4 0 7 1 4 3 5 1 3 4 4 6 7 1 0 2 1 5 6 2 1 2 4 5.

Надмірність коду $R_{\text{над}} = 9/25$, швидкість $R = 16/25$.

Припустимо, що при передачі по каналу зв'язку у кодованій послідовності виникла одна помилка і до декодера надходить така послідовність: 2 4 0 7 1 4 3 5 1 3 4 7 6 7 1 0 2 1 5 6 2 1 2 4 5. Для виявлення та виправлення помилки у декодері кодована послідовність, що надійшла з каналу, записується у вигляді матриці по 5 елементів у кожному рядку і виконується перевірка кожного рядка та кожного стовпця матриці за $\text{mod} 8$:

$$\begin{array}{cccc|c} 2 & 4 & 0 & 7 & 1 & 0 \\ 4 & 3 & 5 & 1 & 3 & 0 \\ 4 & \boxed{7} & 6 & 7 & 1 & 3 \\ 0 & 2 & 1 & 5 & 6 & 0 \\ 2 & 1 & 2 & 4 & 5 & 0 \\ \hline 0 & 3 & 0 & 0 & 0 & 0 \end{array}.$$

У результаті перевірки бачимо, що для третього рядка та другого стовпця перевірка не дає нульового результату. Це говорить про те, що на перетинанні третього рядка та другого стовпця знаходиться спотворений елемент.

Виправлення спотвореного елемента виконують таким чином. У зв'язку з тим, що не виконується перевірка для 3-го рядка і 2-го стовпця, елемент, що знаходиться на перетинанні 3-го рядка і 2-го стовпця, замінюють елементом, який є сумою за $\text{mod}8$ прийнятого (помилкового) елемента та перевірного елемента 3-го рядка (або 2-го стовпця), який був одержаний у декодері, тобто $(7 \oplus 3) \text{mod} 8 = 4$. Таким чином, виправлена послідовність на виході декодера буде мати вигляд: 2 4 0 7 4 3 5 1 4 4 6 7 0 2 1 5.

5. ВИСНОВОК

Як впливає з вищевикладеного, запропоновані недвійкові систематичні коди з повторенням і перевіркою за модулем q та ітеративний мають деякі переваги над відомими недвійковими кодами, що виправляють однократні помилки. Так у порівнянні з недвійковим кодом з багатократним повторенням запропоновані коди мають значно меншу надмірність та більшу швидкість коду, а у порівнянні з узагальненим кодом Хеммінга мають перевагу у технічній реалізації кодера/декодера завдяки значно меншій кількості арифметичних операцій у полі $GF(q)$.

ПЕРЕЛІК ПОСИЛАНЬ

1. Жураковський Ю. П. Передача информации в ГАП. - К.: Вища шк., 1991. - 216 с.
2. Цымбал В.П. Теория информации и кодирование. - К.: Вища шк., 1992. - 263 с.
3. Кодирование информации (двоичные коды): Справ./Н.Т.Березюк, А.Г.Андрущенко, С.С.Мощицкий и др. - Харьков: Вища шк., 1978. - 252 с.
4. Берлекэмп Э. Алгебраическая теория кодирования. - М.: Мир, 1971. - 477 с.
5. Теория кодирования / Т.Касами, Н.Токура, Е.Иведари, Я.Инагаки; Пер. с яп. - М.: Мир, 1978. - 576 с.

ВИКОРИСТАННЯ МАТЕМАТИЧНИХ МЕТОДІВ ДЛЯ ОПТИМІЗАЦІЇ ТЕХНОЛОГІЧНИХ ПРОЦЕСІВ ПЕРЕРОБКИ СОЇ

В.М.Калініченко

В роботі представлена математическа модель технологического процесса гидротермической переработки сои, в которой выделены входные, выходные факторы и управляющие воздействия. Модель позволяет минимизировать содержание антипитательных веществ в продуктах переработки, повысить качество конечного продукта и оптимизировать потери энергии.

У роботі представлена математична модель технологічного процесу гідротермічної переробки сої, у якій виділені вхідні, вихідні фактори і керуючі впливи. Модель дозволяє мінімізувати зміст антипоживних речовин у продуктах переробки, підвищити якість кінцевого продукту й оптимізувати втрати енергії.

Mathematical model of the technological process of hydrothermic soy-bean processing, in which incoming and outgoing factors and operating influence were determined. The model allows to minimize the content of antinourishing substances in processing products, to increase the quality of the final product and optimize the loss of energy.

Кінцевий продукт, це 5% ідея і на 95% - технологія. В розвинених країнах світу в різних галузях виробництва від 45 до 75% коштів витрачається на розробки новітніх технологій, застосування яких дозволяє, не тільки покращити кінцевий продукт за рахунок значного покращення його якісних показників, але й урізноманітнити продукцію, швидко і без особливих витрат змінювати її за вимогами ринку. Це допомагає гідно конкурувати з аналогічною продукцією інших виробників.

Покращення технології кормовиробництва в сучасних умовах дуже актуальна проблема як для світового співтовариства, так і для незалежної України. Якість кормів дозволяє максимально підвищити ефективність тваринництва і максимально зменшити витрати. Кормові продукти, вироблені з зерна сої з застосуванням різних технологій, є дуже цінними для відгодовування практично всіх видів тварин, птиці й риби за рахунок її великої поживної цінності: високого вмісту протеїну, незамінних амінокислот, жиру. Але використанню повножирної сої заважає вміст у неї антипоживних речовин, які значно зменшують цінність корму.

Зараз існує багато різноманітних технологій первинної обробки і глибокої переробки, які дозволяють максимально підвищити поживну цінність сої, як високоякісного корму для різних видів тварин. Всі вони базуються на тепловій, гідротермічній обробці зерна сої [1].

Оскільки всі етапи процесу обробки зерна починаючи з управлінських, економічних, організаційних і закінчуючи технічними і технологічними мають кількісні характеристики і можуть бути записані у вигляді математичних залежностей, то основою розробки і оптимізації цих етапів повинні стати саме математичні методи і моделі. Поряд з

кількісним аналізом вони дозволяють характеризувати і якісні сторони функціонування досліджуваного об'єкту, що зробити за допомогою традиційних заходів не завжди можливо.

Метою нашої роботи є розробка математичної моделі технологічного процесу переробки сої для одержання повноцінного високоякісного, екологічно чистого корму.

Автором досліджено технологічні процеси і обладнання, котрі дозволяють інактивувати антипоживні речовини й значно збільшити поживну цінність соєвого зерна. Зроблено їх порівняльний аналіз. Вивчені технології переробки традиційних для нашої країни культур [1,2]. Але аналіз літератури показує, що проблемам переробки соєвої сировини у нашій країні приділяється недостатньо уваги.

На основі теорії ймовірності, теорії інформаційних процесів з використанням кореляційного, ентропійного аналізу, а також деяких методів математичного моделювання і програмування [2,3,4] розробляється та проходить апробацію математична модель, яка дозволяє оптимізувати і зробити максимально ефективними процеси вирощування, а також первинної та глибокої переробки сої.

У моделі, що розробляється, враховується вплив великої кількості вхідних внутрішніх і зовнішніх факторів, а також керуючих впливів на кінцевий продукт. Тобто модель є багатфакторною. Схематичне зображення виділених і вивчаємих нами найбільш впливових факторів представлено на рис. 1.

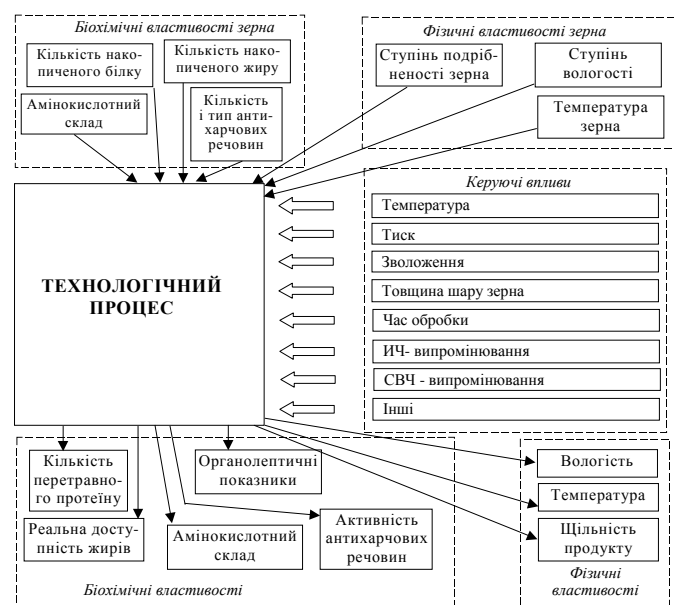


Рисунок 1 - Загальна структура технології переробки

Основні завдання моделі технологічного процесу переробки бобових культур і в першу чергу сої:

- мінімізувати вміст антипоживних речовин у продуктах переробки;
- підвищити якість кінцевого продукту і його перетравність;
- оптимізувати енерговитрати, витрати на обладнання при достатній продуктивності.

В цілому, представлене удосконалення процесу вирощування і переробки сої, дозволяє забезпечити оптимальні режими роботи підприємства по економічним і енерго-економічним показникам, чітко обчислити і забезпечити надійність функціонування апаратури, оптимально керувати резервами сировини, процесами її накопичення і збе-

рігання (задача прогнозування) і що саме головне, отримувати продукцію з максимально можливими якісними показниками. Отримання екологічно чистої продукції дозволяє не тільки вирішити частину екологічних проблем сьогодення, а й зберегти здоров'я нації.

ПЕРЕЛІК ПОСИЛАНЬ

1. Адамень Ф.Ф., Сичкарь В.И., Письменов В.Н., Шерстобитов В.В. Соя. - К: Нора-принт, 1999. - 330 с.
2. Егоров Г.А. Влияние тепла и влаги на процессы переработки и хранения зерна. - Москва: Колос. - 1973 г. - 146 с.
3. Марчук Г.И. Методы вычислительной математики - Москва: Наука, 1989. - 608с.
4. Вергунова І.М. Основи математичного моделювання.- К.: Нора-принт, 2000.-146с.

УДК 621.372.8.01

АЛГОРИТМ МОДЕЛИРОВАНИЯ ФУНКЦИЙ ГРИНА МНОГОСЛОЙНЫХ ДИЭЛЕКТРИЧЕСКИХ СТРУКТУР

Л.М.Карпуков

Рассмотрен алгоритм составления в численно-символьной форме соотношений для расчета функций Грина многослойных диэлектрических структур в краевых задачах электростатики. Основой алгоритма являются предложенные формулы для определения передаточных функций декомпозиционных схем, моделирующих многослойные структуры.

Розглянуто алгоритм складання у чисельно-символьній формі співвідношень для розрахунку функцій Гріна багатослоєвих діелектричних структур у крайових задачах електростатики. Основою алгоритму є запропоновані формули для визначення передатних функцій декомпозиційних схем, що моделюють багатослоєві структури.

The algorithm of compiling in the numerical-symbolical form of ratio for calculations of Green's functions of multilayered dielectric structures in problems of an electrostatic is proposed. Offered formulas for definite of transfer functions of decompositions schemes, modeling multilayered structures are the basis of algorithm.

Использование многослойных диэлектрических структур в конструкциях микроэлектронных устройств обуславливает необходимость совершенствования существующих и разработки новых эффективных алгоритмов моделирования электростатических полей в плоско-слоистых средах.

Для решения краевых задач электростатики предложены разнообразные методы и алгоритмы [1,2]. Среди них наибольшее распространение при моделировании многослойных диэлектрических структур получил метод интегральных уравнений, отличающийся высокой универсальностью и экономичностью.

Эффективность алгоритмов на основе метода интегральных уравнений в значительной степени определяется способом составления их ядер, представляющих собой функции Грина рассматриваемых краевых задач. Существую-

щие методы вычисления функций Грина для многослойных структур обеспечивают составление результатов расчетов в одной из двух форм - в виде явных зависимостей от пространственных координат [3] или в виде Фурье-изображений [4,5], применяемых при решении интегральных уравнений в спектральной области. Недостатками этих методов являются ограничение на сложность исследуемых структур из-за неформализуемости геометрических построений при непосредственном получении явных зависимостей [3] или трудоемкость аналитических преобразований при переходе к оригиналам, обусловленная используемыми способами составления Фурье-изображений [4,5]. Существенно упростить и формализовать процесс вычисления явных зависимостей для функций Грина многослойных структур можно путем формирования изображений в виде отношения экспоненциальных полиномов, по которым легко составляется разностное уравнение для оригинала [6].

Целью настоящей работы является разработка алгоритма расчета коэффициентов экспоненциальных полиномов, определяющих Фурье-изображения функций Грина и соответствующих числителю и знаменателю передаточных функций для параметров рассеяния декомпозиционных схем, моделирующих многослойные структуры [6].

Рассмотрим изображенную на рис. 1,а структуру, составленную из n прилегающих друг к другу пластин однородного и изотропного диэлектрика, размещенных между полупространствами с относительными диэлектрическими проницаемостями ϵ_a , ϵ_b . В плоскости, перпендикулярной оси y , пластины имеют неограниченные размеры.

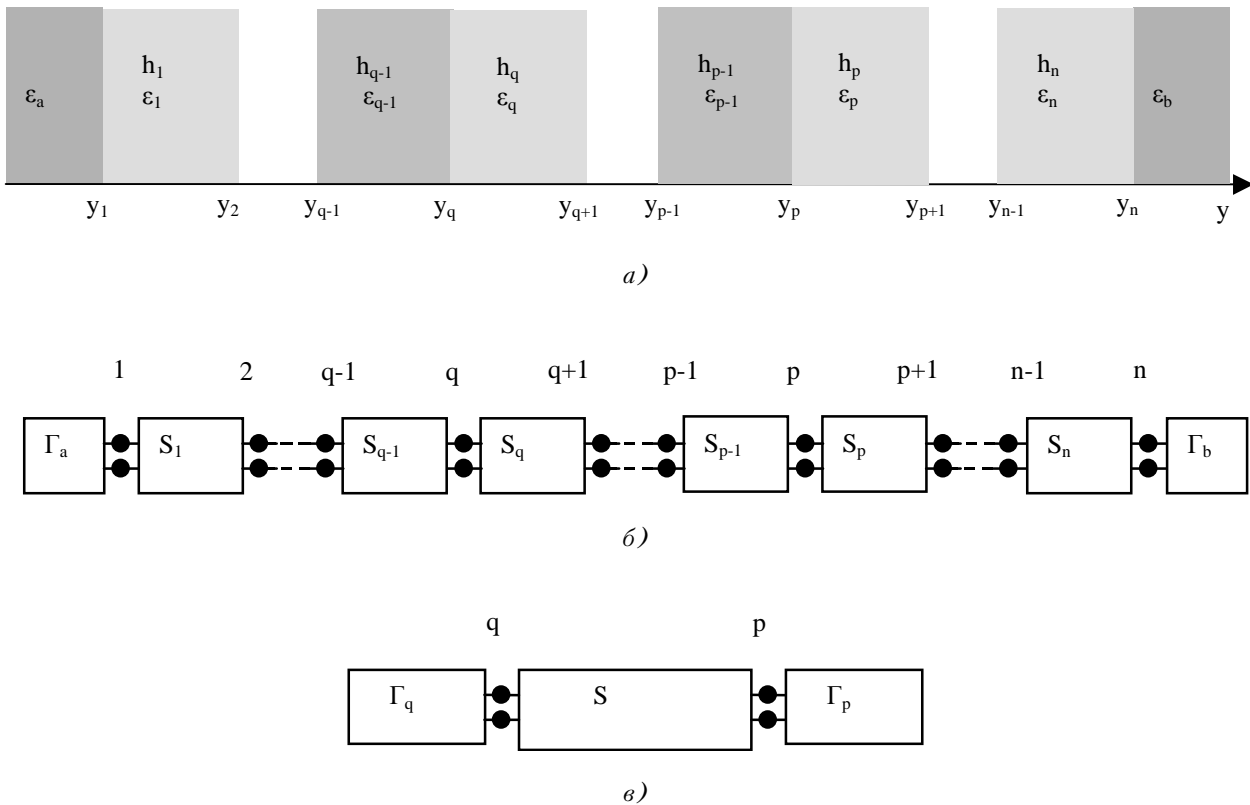


Рисунок 1 - Построение модели многослойной структуры: а) соединение из n диэлектрических пластин, б) декомпозиционная схема соединения, в) эквивалентная схема соединения

Представим рассматриваемую структуру декомпозиционной моделью на рис. 1,б в виде каскадного соединения n четырехполюсников. Элементы декомпозиционной модели будем описывать матрицей рассеяния, а искомое Фурье-изображение функции Грина формировать в виде [6]:

$$G(k_y) = \frac{\hat{A}_{pq}(k_y)}{B(k_y)} u_q(k_y) = \frac{\sum_{i=1}^m a_i e^{-k_y h_{a_i}}}{1 + \sum_{i=1}^m b_i e^{-k_y h_{b_i}}} u_q(k_y). \quad (1)$$

Здесь h_{a_i}, h_{b_i} состояются из толщин диэлектрических слоев; a_i, b_i - выражаются через коэффициенты отражений от плоскостей раздела сред и экранов;

$$u_q = \frac{1}{2k_y \epsilon_0 \epsilon_r} \begin{cases} \exp(jk_x x_0 + jk_z z_0) - \\ \text{для единичного точечного заряда, } k_y = \sqrt{k_x^2 + k_z^2}, \\ \exp(jk_x x_0) - \\ \text{для единичного линейного заряда, } k_y = k_x; \end{cases}$$

индексом 0 отмечены координаты заряда, находящегося в среде с относительной диэлектрической проницаемостью ϵ_r .

Представление (1) удобно для вычислений в спектральной области и одновременно обеспечивает простой переход к оригиналу по рекуррентной формуле [6]

$$G(y + Nh) = \sum_{i=1}^m a_i u_q(y + n_{a_i} h + Nh) - \sum_{i=1}^n b_i G(y + n_{b_i} h + Nh), \quad (2)$$

где h - шаг расчета, кратный толщинам слоев,

$$u_q(y + 2hi) = \begin{cases} \frac{1}{4\pi\epsilon_0\epsilon_r \sqrt{(x-x_0)^2 + (z-z_0)^2 + (y-y_0+2hi)^2}} - \\ \text{для трехмерной задачи,} \\ \frac{1}{4\pi\epsilon_0\epsilon_r \ln[(x-x_0)^2 + (y-y_0+2hi)^2]} - \\ \text{для двумерной задачи.} \end{cases}$$

Для начала вычислений используется нулевое граничное условие для потенциала на бесконечности:

$$G(y + (N + 1)h) = 0, N \rightarrow \infty. \quad (3)$$

Для нахождения передаточных функций $\hat{S}_{pq}(k_y) = \frac{\hat{A}_{pq}(k_y)}{B(k_y)}$, входящих в (1), запишем элементы и определитель мат-

рицы рассеяния i -го четырехполюсника, моделирующего размещенную в свободном пространстве i -ую пластину с относительной проницаемостью ϵ_i и толщиной h_i , следующим образом:

$$\left. \begin{aligned} S_{11_i} = S_{22_i} &= \frac{A_{11_i}}{B_i} = \frac{\Gamma_i(1 - e^{-2k_y h_i})}{1 - \Gamma_i^2 e^{-2k_y h_i}}, \\ S_{21_i} = S_{12_i} &= \frac{A_{21_i}}{B_i} = \frac{(1 - \Gamma_i^2)e^{-k_y h_i}}{1 - \Gamma_i^2 e^{-2k_y h_i}}, \\ \Delta_i &= \frac{A_i}{B_i} = \frac{\Gamma_i^2 - e^{-2k_y h_i}}{1 - \Gamma_i^2 e^{-2k_y h_i}}, \end{aligned} \right\} \quad (4)$$

где $\Gamma_i = (1 - \epsilon_i)/(1 + \epsilon_i)$.

Пусть в q -м сечении находится источник поля, а в p -м сечении точка наблюдения. Тогда по эквивалентной схеме на рис. 1, в нетрудно найти передаточную функцию $\hat{S}_{pq}(k_y) = u_p(k_y)/u_q(k_y)$, определяющую связь между Фурье-изображениями наблюдаемого потенциала u_p и источника поля u_q , располагаемого в среде с проницаемостью ϵ_0 свободного пространства:

$$\hat{S}_{pq} = A_{pq} \frac{C_{pq}}{B} = A_{pq} \frac{[A_q - B_q] \cdot \begin{bmatrix} 1 & 1 \\ -1 & -1 \end{bmatrix} \cdot \begin{bmatrix} -A_p \\ -B_p \end{bmatrix}}{[A_q - B_q] \cdot \begin{bmatrix} A_s & -A_{qq} \\ A_{pp} & -B_s \end{bmatrix} \cdot \begin{bmatrix} -A_p \\ -B_p \end{bmatrix}}. \quad (5)$$

Здесь использовано представление параметров компонентов эквивалентной схемы на рис. 1, в в виде, аналогичном (4): $\Gamma_q = A_q/B_q$, $\Gamma_p = A_p/B_p$, $S_{pq} = A_{pq}/B_s$, $\Delta_s = A_s/B_s$.

Представим члены выражения (5) с помощью формул символического анализа каскадных соединений [7] и с учетом (4) в следующем виде:

$$\left. \begin{aligned} A_{pq} &= (-1)^{p-1-q} \prod_{i=q}^{p-1} (1 - \Gamma_i^2) e^{-k_y h_i}, \\ \hat{B} &= [\Gamma_a - 1] \cdot \prod_{i=1}^n ([V_i] + e^{-2k_y h_i} [W_i]) \cdot \begin{bmatrix} -\Gamma_b \\ -1 \end{bmatrix}, \\ C_{pq} &= [\Gamma_a - 1] \cdot \prod_{i=1}^{q-1} ([V_i] + e^{-2k_y h_i} [W_i]) \cdot \begin{bmatrix} 1 & 1 \\ -1 & -1 \end{bmatrix} \times \\ &\quad \times \prod_{i=p}^n ([V_i] + e^{-2k_y h_i} [W_i]) \cdot \begin{bmatrix} -\Gamma_b \\ -1 \end{bmatrix}. \end{aligned} \right\} \quad (6)$$

$$\text{Здесь } V_i = \begin{bmatrix} \Gamma_i^2 & -\Gamma_i \\ \Gamma_i & -1 \end{bmatrix}, \quad W_i = \begin{bmatrix} -1 & \Gamma_i \\ -\Gamma_i & \Gamma_i^2 \end{bmatrix}.$$

Формулы (5), (6) позволяют легко реализовать процедуру вычисления коэффициентов экспоненциальных полиномов для формирования выражения (1) или (2) в удобном для расчетов численно-символьном виде. Например, для структуры из трех пластин с относительными проницаемостями $\epsilon_1 = 9$, $\epsilon_2 = 5$, $\epsilon_3 = 4$, размещенных между полупространствами с $\epsilon_a = 3$, $\epsilon_b = 100$, и при нахождении точки источник во 2-м, а точки наблюдения в 3-м сечении, расчет по этим формулам дает следующее численно-символьное представление Фурье-изображения функции Грина:

$$u_3 = \frac{A_{32}}{B} \cdot u_2,$$

$$\text{где } A_{32} = 0,153e^{-k_y h_2} + 0,079e^{-k_y(h_2 + 2h_1)} - 0,146e^{-k_y(h_2 + 2h_3)} - 0,073e^{-k_y(h_2 + 2h_1 + 2h_3)},$$

$$B = 1 - 0,143e^{-2k_y h_1} + 0,032e^{-2k_y h_2} - 0,103e^{-2k_y h_3} - 0,056e^{-2k_y(h_1 + h_2)} + 0,015e^{-2k_y(h_1 + h_3)} - 0,264e^{-2k_y(h_2 + h_3)} + 0,461e^{-2k_y(h_1 + h_2 + h_3)}.$$

Предложенный алгоритм расчета функций Грина, отличаясь простотой реализации и низкими вычислительными затратами, позволит существенно повысить эффективность решения сложных краевых задач электростатики при исследовании и проектировании конструкций, содержащих многослойные диэлектрические структуры.

ПЕРЕЧЕНЬ ССЫЛОК

1. Миролюбов Н.Н. и др. Методы расчета электростатических полей. - М.: Высшая школа, 1963. - 415 с.
2. Фуско В. СВЧ цепи. Анализ и автоматизированное проектирование / Пер. с англ. - М.: Радио и связь, 1990. - 288 с.
3. Silvester P. TEM wave properties of microstrip transmission lines // Proc. IEE, -1968. - V.115, -№1. - P. 43-48.
4. Grampagne R., Ahmadpanah M., Guiraud J. A simple method for determining the Green's function for a large class of MIC lines having multilayered dielectric structures // IEEE Trans. MTT. - 1978. - V.26, -№2. - P. 82-87.
5. Li K., Atsuki K., Hasegawa T. General analytical solutions of static Green's functions for shielded and open arbitrarily multilayered media // IEEE Trans. MTT. - 1997. - V.45, -№1. - P. 2-8.
6. Карпуков Л. М. Построение и анализ декомпозиционных моделей микрополосковых структур // Радиоэлектроника. - 1984. - Т. 27. - №9. - С. 32 - 36. (Изв. высш. учеб. заведений).
7. Карпуков Л. М. Символьный анализ устройств СВЧ методом подсхем // Электронное моделирование. - 1984. - Т. 6. - №3. - С. 81 - 84.

ІНШИЙ ВИВІД ОСНОВНИХ СПІВВІДНОШЕНЬ НЕПЕРЕРВНОГО ВАРІАНТУ ДИНАМІЧНОГО ПРОГРАМУВАННЯ

А.Г.Кіку, Т.І.Білоус

Предложенный подход к выводу соотношений непрерывного варианта метода динамического программирования, который основан непосредственно на принципе оптимальности и независимости величины показателя качества от точки деления интервала интегрирования на два подинтервала. Рассмотренный подход натуральным способом делит начальную задачу на две подзадачи. Первая подзадача связана с просмотром решения с конца интервала управления к его началу и этим же вносит ясность в сам метод, как и в случае дискретного варианта. Решение же второй подзадачи позволяет реализовать второй этап решения общей задачи, а именно реализовать движение от начала интервала управления к его концу.

Запропонований підхід до виводу співвідношень неперервного варіанту методу динамічного програмування, який заснований безпосередньо на принципі оптимальності і незалежності величини показника якості від точки розділення інтервалу інтегрування на два підінтервала. Розглянутий підхід натуральним способом поділяє початкову задачу на дві підзадачі. Перша підзадача зв'язана з переглядом розв'язку з кінця інтервалу управління до його початку і тим самим вносить ясність в самий метод, як і у випадку дискретного варіанту. Розв'язання ж другої підзадачі дозволяє реалізувати другий етап рішення загальної задачі, а саме реалізувати рух від початку інтервалу управління до його кінця.

The offered approach to conclusion of continuous variant correlations of dynamical programming, which is applied directly to the principle of optimum and independence of qualitative indice value from the point, separating the interval of integration on two subintervals. The reviewed approach separate the source problem on two subproblems naturally. The first subproblem deals with the review of solution from the end of carrying interval to its beginning and so brings the clearance into the method as in the case of discrete variant. Solving the second subproblem allows us to realize the second step of the general problem's solution, namely to realize the movement from the beginning of carrying interval to its end.

В основі методу динамічного програмування лежить принцип оптимальності, який за своєю суттю є простим і добре узгодженим зі здоровим глуздом. Але тільки після його доведення Беллманом метод отримав широке розповсюдження для розв'язання задач оптимального управління, у тому числі і динамічних задач. Зокрема, він дозволив Беллману вивести диференційне рівняння, на основі якого розв'язання рішення задачі оптимального управління у вигляді аналітичного конструювання регуляторів (АКОР) може бути доведено до кінця.

В статті запропоновано інший підхід до виводу основних співвідношень неперервного варіанту методу динамічного програмування, заснований за аналогією з дискретним варіантом безпосередньо на принципі оптимальності і незалежності величини показника якості від точки розподілу інтервалу інтегрування при поділі останнього на два підінтервали.

Нехай задача оптимального управління сформульована у такій постановці:

$$\min_{u \in U} \{ I[x, u, x(t_2)] = \int_{t_1}^{t_2} f(x, u, t) dt + l[x(t_2)] \} \quad \left\{ \begin{array}{l} \frac{dx_i}{dt} = \varphi_i(x, u, t) \quad i = \overline{1, n}, \\ x(t_i) = x_i^1, \\ u \in U^D, \end{array} \right. \quad (1)$$

де $x = [x_1 \dots x_n]^T$ - вектор змінних стану,

$u = [u_1 \dots u_n]^T$ - вектор управління,

U^D - множина допустимих управлінь,

f, φ_i - інтервально-неперервні функції своїх аргументів.

Розглянемо показник якості при конкретно вибраному допустимому управлінні $u^D(t)$ на інтервалі $[t_1, t_2]$. Виберемо на цьому інтервалі $[t_1, t_2]$ деяку проміжну точку τ і подамо (1) таким чином:

$$I = \int_{t_1}^{t_2} f(x, u^D, t) dt + l[x(t_2)] = \int_{t_1}^{\tau} f(x, u^D, t) dt + \int_{\tau}^{t_2} f(x, u^D, t) dt + l[x(t_2)] = \int_{t_1}^{\tau} f(x, u^D, t) dt + S(x, u^D, \tau). \quad (2)$$

Оскільки показник якості за його властивістю не залежить від обраної проміжної точки, то має місце таке:

$$\frac{dI}{d\tau} = f(x, u^D, \tau) + \frac{dS(x, u^D, \tau)}{d\tau} = 0. \quad (3)$$

Таким чином (3) є необхідною умовою незалежності I від точки розподілу τ .

Подамо $\frac{dS}{d\tau}$ у (3) таким чином:

$$\frac{dS}{d\tau} = \sum_{i=1}^n \frac{dS}{dx_i} \frac{dx_i}{d\tau} + \frac{dS}{d\tau} = S_x^T \varphi(x, u^D, \tau) + S_\tau, \quad (4)$$

де T - індекс операції транспонування,

$$S_{x_i} = \frac{dS}{dx_i}, S_\tau = \frac{dS}{d\tau}.$$

В результаті чого ця умова приймає наступний вигляд:

$$f(x, u^D, \tau) + S_x^T(x, u^D, \tau)\varphi(x, u^D, \tau) + S_\tau = 0, \quad (5)$$

або

$$S_x^T(x, u^D, \tau)\varphi(x, u^D, \tau) + S_\tau = -f(x, u^D, \tau). \quad (6)$$

Тут зазначимо, що на функцію $S(x, u^D, \tau)$ не накладаються обмеження її неперервності диференційованості по x , u та t . Більш того, якщо, наприклад, $f(x, u, t)$ або $\varphi_i(x, u, t)$ будуть розривними в деяких точках x^P , або u^P , або t^P , тоді й S_{x_i} , S_t повинні бути також розривними у відповідних точках, що й має місце, наприклад, в задачах оптимальної швидкодії на лініях перемикання.

Оскільки отримане відношення справедливе при будь-якому $t = \tau$, то воно може бути записане таким чином:

$$S_x^T(x, u^D, \tau)\varphi(x, u^D, \tau) + S_t = -f(x, u^D, \tau). \quad (7)$$

Для визначення функції S необхідно інтегрувати рівняння у часткових похідних (7) при умові на правій межі $S(t_2) = I[x^*(t_2)]$. При цьому одержимо розв'язок S , який залежить від x, u^D, t . Зауважимо, що за оптимального управління $u = u^*$ функція S перетворюється в функцію Беллмана S^* , а рівняння (7) - у диференціальне рівняння Беллмана

$$S_x^{*T}(x^*, u^*, t)\varphi(x^*, u^*, t) + S_t^* = -f(x^*, u^*, t). \quad (8)$$

Таким чином, для визначення функції Беллмана необхідно знайти оптимальне управління.

Знайдемо тепер оптимальне управління безпосередньо на основі принципу оптимальності Беллмана за аналогією з його визначення у дискретному варіанті динамічного програмування.

Нехай задача (1) "переглянута" з кінця інтервалу управління до моменту. При цьому будуть знайдені u^* і S^* у вигляді функцій від $x^*(\tau + dt)$. Але $x^*(\tau + dt)$ залежить від невідомих $x^*(\tau)$ і $u^*(\tau)$, звідки випливає, що потрібен подальший перегляд рішення задачі до початку інтервалу управління. Змістимо час "ліворуч" ще на диференціал часу dt і знайдемо $u^*(\tau)$ згідно з принципом оптимальності шляхом розв'язання задачі:

$$\min_{u(\tau) \in U^D} \left\{ \int_{\tau}^{t_2} f(x^*, u^D, t) dt + I[x^*(t_2)] \right\}, \quad (9)$$

котра приводиться до вигляду:

$$\min_{u(\tau) \in U^D} \left\{ \int_{\tau}^{\tau+dt} f(x^*, u, t) dt + \int_{\tau+dt}^{t_2} f(x^*, u, t) dt + I[x^*(t_2)] \right\} =$$

$$= \min_{u(\tau) \in U^D} \{f(x^*, u, t) dt + S^*(\tau + dt)\}. \quad (10)$$

Відзначимо, що $S^*(\tau + dt)$ потрапляє під знак операції мінімізації по $u(\tau)$, оскільки вона залежить від $x^*(\tau + dt)$, який в свою чергу згідно моделі об'єкта управління залежить від $u^*(\tau)$ і $x^*(\tau)$. Подано в (10) $S^*(\tau + dt)$ у вигляді розкладу в ряд Тейлора в точці τ :

$$\min_{u(\tau) \in U^D} \left\{ f(x^*, u, \tau) dt + S^*(\tau) + S_x^{*T}(x^*, \tau)\varphi(x^*, u, t) dt + S_t^* dt \right\} = S^*(\tau). \quad (11)$$

Оскільки у цій задачі $S^*(\tau)$ представляє число, вже не залежне від $u^*(\tau)$, тоді (11) з точки зору визначення оптимального управління приводиться до вигляду:

$$\min_{u(\tau) \in U^D} \left\{ f(x^*, u, \tau) + S_x^{*T}(x^*, \tau)\varphi(x^*, u, t) \right\} = -S_t^* \quad (12)$$

Співвідношення (12) справедливе для будь-якого $t = \tau \in [t_1, t_2]$, тому воно може бути записано таким чином:

$$\min_{u(\tau)} \left\{ f(x, u, \tau) + S_x^{*T}(x^*, t)\varphi(x^*, u, t) \right\} = -S_t^*. \quad (13)$$

На основі (13) оптимальне управління отримаємо у вигляді функції від S^* , x^* й t , тобто

$$u^* = u^*(S^*, x^*, t), \quad (14)$$

де аргумент t вказує на його нестационарність в загальному випадку алгоритму (14).

Якщо тепер в цьому алгоритмі виразити S^* у вигляді функції від змінних стану x^* і часу t , то отримаємо потрібну залежність оптимального управління від оптимальної траєкторії. Таким чином, якщо б знали поточний оптимальний стан $x^*(t)$ в кожний момент часу, то змогли б знайти оптимальне управління, а по $x^*(t)$ та $u^*(t)$ змогли б знайти $x^*(t+dt)$ й так далі. Проте оптимальний стан $x^*(t)$ траєкторії $x(t)$ відомо тільки для початкового моменту часу $t=t_1$, для якого $x^*(t_1) = x^1$. Звідси випливає, що процес перегляду розв'язання задачі з кінця інтервалу управління повинен бути продовжений до його початку. Але зауважимо, що, якби в (14) функцію S^*

вдалося б виразити у вигляді залежності від x^* , u^* , t , то на його основі можна було б вивести алгоритм розв'язання задачі аналітичного конструювання регуляторів $u^*(x^*, t)$. Отримати вказану залежність можливо інтегруванням рівняння (7), попередньо підставляючи в нього вираз (14) для u^* . Зазначимо, що рівняння (8) розв'язується в зворотньому часі. Цим по суті замінюється процес перегляду задачі з кінця інтервалу управління до його початку при визначенні оптимального управління.

Розв'язуючи (8) зазначеним способом, знаходимо функцію Беллмана у вигляді:

$$S^* = S^*(x^*, u^*, t). \quad (15)$$

Підставляючи тепер (15) та (14), знаходимо алгоритм оптимального управління

$$u^* = u^*(x^*, t). \quad (16)$$

Запропонований підхід виведення основних співвідношень метода динамічного програмування для неперервних задач управління природним способом поділяє початкову задачу на дві підзадачі. Перша з них пов'язана з переглядом розв'язку від кінця інтервалу управління до його початку і тим самим вносить ясність до самого методу, як і у випадку дискретного варіанту. Розв'язання ж другої підзадачі, пов'язаної з визначенням функції Беллмана, дозволяє реалізувати другий етап рішення загальної задачі, а саме реалізувати рух від початку інтервалу управління до його кінця.

У прикінці зазначимо ще раз, що на функцію $S^*(x, u^D, t)$ згідно (6) не накладаються обмеження її неперервної диференційованості по x , u та t .

УДК 519.711: 658.562.3

КОНТРОЛЬ ТЕХНОЛОГИЧЕСКОГО ПРОЦЕССА С ИСПОЛЬЗОВАНИЕМ КАРТЫ ХОТЕЛЛИНГА

В.Н.Клячкин

Рассматривается многомерный статистический контроль технологического процесса с использованием карты Хотеллинга. Анализируется эффективность применения статистики T^2 при многомерном контроле с помощью средней длины серий и интерпретация результатов контроля.

This article presents the multivariate statistical process control (SPC) with Hotelling charts. Its carry out the comparison of sensitivity of T^2 statistic with average run length (ARL) and interpretation of these charts.

ВВЕДЕНИЕ

Статистический контроль технологических процессов - одно из интенсивно развивающихся направлений всеобщего менеджмента качества, его внедрение - необходимое условие сертификации предприятия на соответствие международным стандартам качества ИСО-9000. Основным инструментом такого контроля - контрольная карта процесса - в большей степени пока еще ориентирована на проведение анализа вручную. Между тем использование компьютерной техники позволяет существенно расширить возможности управления процессом производства, в частности, в ситуации, когда качество изготавливаемого изделия определяется множеством показателей.

ПОСТАНОВКА ЗАДАЧИ

При проведении контроля технологического процесса обычно используются контрольные карты Шухарта для одного показателя качества, наиболее важного с точки

зрения результатов процесса. На практике часто процесс характеризуется совокупностью показателей $\mathbf{X} = (X_1, X_2, \dots, X_p)$, имеющих совместное нормальное распределение и коррелированных между собой. В этом случае может быть использована карта Хотеллинга [1]: для каждой выборки рассчитывается статистика

$$T_l^2 = n(\bar{\mathbf{X}}_l - \boldsymbol{\mu})^T \mathbf{S}^{-1} (\bar{\mathbf{X}}_l - \boldsymbol{\mu}), \quad (1)$$

где n - объем выборки, l - ее номер ($l = 1, \dots, m$; m - количество выборок), \mathbf{X}_l - вектор средних мгновенной выборки, $\boldsymbol{\mu}$ - вектор общих средних, \mathbf{S} - оценка ковариационной матрицы $\boldsymbol{\Sigma}$.

Статистика T^2 имеет распределение Хотеллинга [2], а статистика

$$F = [mn - m - p + 1] T^2 / [pm(n - 1)], \quad (2)$$

- нецентральное F -распределение Фишера, поэтому контрольная граница на карте Хотеллинга определяется в зависимости от заданного уровня значимости α с использованием таблицы квантилей F -распределения

$$T_{kp}^2 = [pm(n-1)(mn-m-p+1)] F_{1-\alpha}(p, mn-m-p+1). \quad (3)$$

Если ковариационная матрица $\boldsymbol{\Sigma}$ известна, статистика

Хотеллинга имеет χ^2 -распределение; в этом случае контрольная граница

$$T_{kp}^2 = \chi_{1-\alpha}^2(p). \tag{4}$$

Процесс статистически управляем, если $T_j^2 < T_{kp}^2$.

ОЦЕНКА ЭФФЕКТИВНОСТИ КАРТЫ ХОТЕЛЛИНГА

Эффективность контрольной карты определяется количеством наблюдений от момента нарушения технологического процесса до момента обнаружения этого нарушения; это средняя длина серий. Если нарушение процесса состоит в сдвиге среднего значения одного из показателей качества на некоторую величину δ , то средняя длина серий [3] $L(\delta) = [G(\delta)]^{-1}$, где $G(\delta)$ - вероятность останова процесса (функция мощности).

Для карты Хотеллинга вероятность останова процесса - это вероятность того, что статистика T^2 превысит критическое значение T_{kp}^2 :

$$G(\delta) = P(T^2 \geq T_{kp}^2) = 1 - F(T_{kp}^2), \tag{5}$$

где $F(x)$ - функция распределения Хотеллинга (для ситуации, когда ковариационная матрица Σ неизвестна), с параметром нецентральности, в общем случае равным

$$\lambda = n\delta^T \Sigma^{-1} \delta; \tag{6}$$

вектор $\delta = (\delta_1, \delta_2, \dots, \delta_p)^T$; δ_j - сдвиг по показателю X_j .

Средняя длина серий контрольной карты

$$L(\delta) = \left[\begin{array}{c} T_{kp}^2 \\ 1 - \int f(t) \partial t \\ 0 \end{array} \right]^{-1}, \tag{7}$$

где $f(t)$ - плотность нецентрального распределения Хотеллинга:

$$f(t) = \exp(-\lambda/2) / [(n-1)\Gamma(0, 5(n-p))] \times \tag{8}$$

$$\times \sum_{s=0}^{\infty} [(\lambda/2)^s (t/(n-1))^{0,5p+s-1} \Gamma(0, 5n+s)] /$$

$$/ [s! \Gamma(0, 5p+s) (1+t/(n-1))^{0,5n+s}].$$

На рисунке 1 показаны графические результаты расчета средней длины серий для карты Хотеллинга, приведенные в пакете Mathcad, в зависимости от величины сдвига по одному из p некоррелированных показателей качества для

значений $p = 2, 5, 10$ и для объемов в 50 и 100 наблюдений на уровне значимости $\alpha = 0.005$, соответствующем средней длине серий при отсутствии сдвига, равной 200. Видно, что с увеличением числа показателей эффективность карты снижается: средняя длина серий увеличивается.

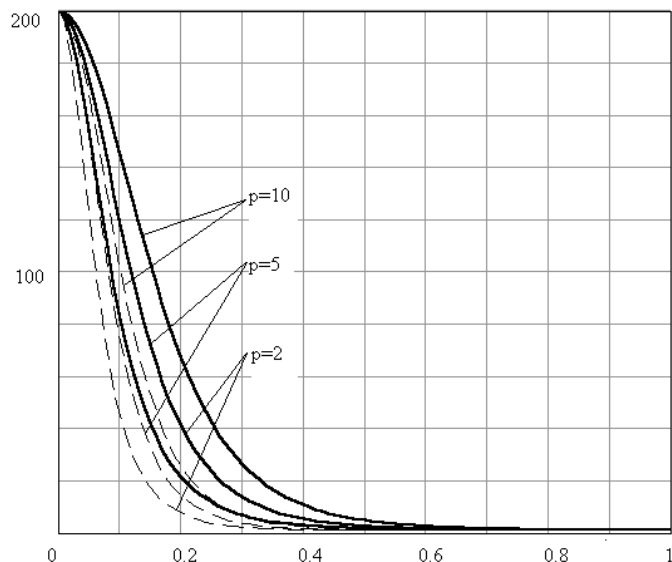


Рисунок 1 - Зависимость средней длины серий карты Хотеллинга от сдвига процесса при различном числе показателей качества p и различном объеме наблюдений (сплошная линия - 50, штриховая - 100)

На рисунке 2 для двух показателей качества показано влияние коэффициента корреляции между показателями на среднюю длину серий. Увеличение степени коррелированности повышает эффективность карты Хотеллинга.

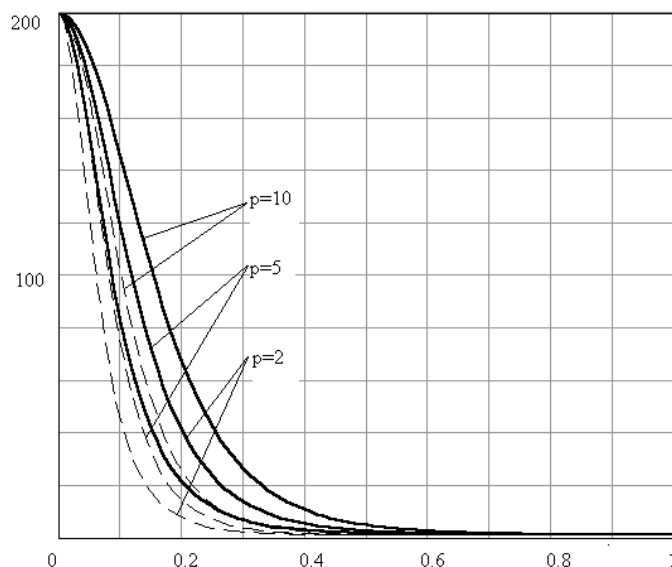


Рисунок 2 - Влияние степени коррелированности показателей качества на среднюю длину серий карты Хотеллинга

ИНТЕРПРЕТАЦИЯ КАРТЫ ХОТЕЛЛИНГА

Пусть при некотором $l = l_0$ оказалось, что $T_{l_0}^2 \geq T_{kp}^2$, то есть на карте зафиксировано нарушение технологического процесса. Карта Хотеллинга не показывает непосредственно, с каким из показателей качества (или совместным влиянием показателей) связано нарушение процесса. В связи с этим встает вопрос об интерпретации результатов контроля. Для проверки гипотезы о том, что ответственность за нарушение процесса лежит на j -м показателе, может быть использован частный критерий Хотеллинга [4]. Гипотеза верна, если

$$T_j^2 = n[\mathbf{c}_j^T(\bar{\mathbf{X}}_{l_0} - \mu)]^2 / [\mathbf{c}_j^T \mathbf{S} \mathbf{c}_j] > T_{kp}^2, \quad (9)$$

где \mathbf{c}_j - вектор-столбец, состоящий из нулей во всех строках, кроме j -й, и единицы в j -й строке.

Если все p гипотез отвергаются, то это означает, что

нарушение процесса вызвано совместным воздействием нескольких показателей. Для проверки таких гипотез, например, о совместном влиянии двух показателей одновременно, столбец \mathbf{c}_j корректируется: в соответствующих двух строках нули заменяются единицами. Таким образом, используя частный критерий Хотеллинга, можно выявить причину, по которой произошло нарушение технологического процесса.

ПЕРЕЧЕНЬ ССЫЛОК

1. Клячкин В.Н. Оценка эффективности многомерного контроля качества технологического процесса // Вестник УлГТУ. Информационные технологии. 1999. №2. С.59-62.
2. Андерсон Т. Введение в многомерный статистический анализ. М.: Физматгиз, 1963. - 500 с.
3. Миттаг Х., Ринне Х. Статистические методы обеспечения качества. - М.: Машиностроение, 1995. - 616 с.
4. Сошникова Л.А., Тамашевич В.Н., Уебе Г., Шефер М. Многомерный статистический анализ в экономике. М.: ЮНИТИ-ДАНА, 1999. - 598 с.

УДК 612.13:537.8

АНАЛИЗ РАБОТЫ СИСТЕМЫ КРОВООБРАЩЕНИЯ, КАК ПРОЦЕССА ПЕРЕДАЧИ ИНФОРМАЦИИ КЛЕТКАМИ КРОВИ, МЕТОДАМИ ВНЕШНЕЙ АЛГЕБРЫ

Г.В.Кузнецов, А.А.Яшин

Исследованы информационные аспекты кровообращения. Получена модель кровообращения на основе кинематических уравнений потока крови.

Dosліджено інформаційні аспекти кровообігу. Отримано модель кровообігу на основі кінематичних рівнянь потоку крові.

The information aspects of blood circulation are investigated. The blood circulation model is obtained on a basis cinematic equations of blood stream.

1. ВВЕДЕНИЕ

Моделирование системы кровообращения с различных точек зрения и по различным причинам рассматривалось в работах многих авторов. Вообще, под системой понимается "упорядоченное целостное множество взаимосвязанных элементов, обладающее собственной организацией и структурой" [1]. Следует отметить, что основным свойством системы является ее целостность, которая выражается в не сводимости свойств всей системы к сумме свойств составляющих ее частей.

Еще в 1939 году Г.Ф. Ланг предложил понятие системы крови, в которое он включил кровь, регулирующий нейрогуморальный аппарат, а также органы, в которых происходит образование клеток крови и их разрушение: костный мозг, вилочковую железу, лимфатические узлы,

селезенку и печень. Компоненты системы крови осуществляют контакт между собой посредством различных кровеносных сосудов или кровяным руслом.

Система кровообращения является, прежде всего, транспортной системой организма. Основная задача, которая решается ей в организме, является перенос различных веществ, с помощью которых происходит защита от воздействий внешней среды, а также регуляция деятельности отдельных органов и систем организма. Тем самым, функции крови можно охарактеризовать, основываясь на химическом составе переносимых веществ, а также на последствиях, вызванных действиями этих веществ. Исходя из этого, кровь выполняет следующие функции: 1) дыхательную, 2) питательную, 3) экскреторную, 4) гомеостатическую, 5) регуляторную, 6) креаторных связей, 7) терморегуляционную, 8) защитную.

Нас интересуют функции крови, которые касаются передачи информации клетками крови. Наиболее ярко эту роль, среди всех функций крови, выполняет функция креаторных связей. Эта функция состоит в переносе плазмой и форменными элементами макромолекул, которые осуществляют в организме информационные связи. Благодаря этому регулируются не только различные внутриклеточные процессы и поддержание постоянства структуры тканей, но и функционирование организма как единой системы, что обеспечивает приспособление орга-

низма к постоянно меняющимся внешним условиям.

По скорости передачи информации в организме можно привести следующие ее пути: электромагнитный (самый быстрый); второй путь - это передача информации посредством биохимических реакций; и, наконец, третий путь - это перенос информации с движущимися частицами. Предложенные пути рассматриваются по мере уменьшения скорости переносимой информации. Оптимальная деятельность организма, как сложной целостной системы, возможна только при условии, что перечисленные выше пути передачи информации дублируют и дополняют друг друга. Передаваемая информация должна доходить до всех клеток и тканей, восприимчивость которых к способам передачи этой информации может существенно различаться. Одни клетки более активно воспринимают информацию, передаваемую первым путем, а другие, например, третьим путем. Следует также иметь в виду, что в различных ситуациях, в которых может находиться организм, важно и дублирование информации различными способами ее передачи.

Тем самым, встает задача моделирования движения крови, как одном из путей передачи информации. Данная работа посвящена данной проблеме. Причем, в области непрерывности потока крови, исследования ведутся методами внешней алгебры, продуктивность которого не раз показывалась авторами данной работы [2,3]. Все рассмотрения ведутся в одном из видов риманова пространства - субпроективном [4], в котором моделируется движение крови и которое является моделью сердечно - сосудистой системы человека.

2. ОСНОВНЫЕ ПОНЯТИЯ И ФОРМУЛЫ

В касательном пространстве к трехмерному субпроективному пространству C^3 рассмотрим репер, определяемый точкой $x \in C^3$ и векторами \vec{e}_i , где $i = 1, 2, 3$. Также, в дальнейшем, все индексы, обозначаемые малыми латинскими буквами, будут принимать эти значения. Уравнения перемещения такого репера имеют вид:

$$d\vec{x} = \omega^i \vec{e}_i, d\vec{e}_i = \omega_i^j \vec{e}_j + \omega^j \vec{e}_{ij}, \quad (1)$$

где \vec{e}_{ij} - векторы, образующие с векторами \vec{e}_i , репер второго порядка, связанный с точкой x [5].

Рассматриваем вихревое движение крови, при котором существует семейство поверхностей, на котором находятся все линии тока крови и все вихревые линии. В этом случае векторы \vec{e}_{ij} симметричны по нижним индексам [6].

Уравнения структуры субпроективного пространства имеют вид:

$$D\omega^i = \omega^j \wedge \omega_j^i, D\omega_j^i = \omega_j^k \wedge \omega_k^i + R_{jkl}^i \omega^k \wedge \omega^l, \quad (2)$$

где тензор $R_{ijkl} = g_{is} R_{jkl}^s$ для данного субпроективного

пространства [4] имеет вид:

$$R_{ijkl} = g_{ik} \alpha_{jl} + g_{jl} \alpha_{ik} - g_{il} \alpha_{jk} - g_{jk} \alpha_{il},$$

а

$$\alpha_{jk} = P(\alpha) g_{jk} + Q(\alpha) \alpha_j,$$

где α - некоторая не постоянная функция, которая связана с коэффициентом конформности, при конформном отображении евклидова пространства на риманово пространство [7].

Для удобства дальнейших вычислений примем следующие обозначения:

$$\begin{aligned} \omega_2^3 &= -\omega_3^2 = p = p_i \omega^i, \quad \omega_3^1 = -\omega_1^3 = q = q_i \omega^i, \\ \omega_1^2 &= -\omega_2^1 = r = r_i \omega^i, \end{aligned} \quad (3)$$

где учтено, что в окрестности Ω точки $x \in C^3$

$$\omega_i^i = 0 \text{ и } \omega_j^i + \omega_i^j = 0. \quad (4)$$

Направим вектор \vec{e}_3 ортогонально семейству поверхностей, на которых находятся линии тока крови и вихревые линии. Тем самым в каждой окрестности любой точки пространства C^3 будет определено поле векторов \vec{e}_3 .

Для произвольного смещения точки x по поверхности семейства имеем:

$$\vec{e}_3 d\vec{x} = \omega^3 = 0. \quad (5)$$

Последнее уравнение является вполне интегрируемым, то для него по теореме Фробениуса получим:

$$p_1 + q_2 = 0. \quad (6)$$

1-форма ω^3 имеет вид:

$$\omega^3 = S ds. \quad (7)$$

Ввиду того, что линии тока крови и вихревые линии лежат на поверхностях, то положим:

$$\vec{v} = v(\cos \sigma \vec{e}_1 + \sin \sigma \vec{e}_2), \quad \vec{v} = \text{rot } \vec{v} = v^1 \vec{e}_1 + v^2 \vec{e}_2, \quad (8)$$

где первая формула представляет вектор скорости крови \vec{v} , а вторая - вихревой вектор \vec{v} .

Уравнения Громеки-Ламба запишем в виде:

$$dH = \vec{v} d\vec{x}, \quad (9)$$

где $H = \frac{1}{2}v^2 + U + \frac{p}{\rho}$, p - давление крови, ρ - ее плотность, а U - потенциал внешних сил, действующих на кровь.

С учетом (8) из уравнения (9) можно заключить, что:

$$H = f(s). \tag{10}$$

Тогда получим:

$$Sv(-v^1 \sin \sigma + v^2 \cos \sigma) = f'(s). \tag{11}$$

Как было получено в [6], запишем:

$$\begin{aligned} \vec{e}_1 \vec{e}_{12} = 0, \quad \vec{e}_2 \vec{e}_{21} = 0, \quad \vec{e}_3 \vec{e}_{31} = 0, \quad \vec{e}_1 \vec{e}_{13} = 0, \quad \vec{e}_2 \vec{e}_{22} = 0, \\ \vec{e}_3 \vec{e}_{32} = 0, \quad \vec{e}_1 \vec{e}_{11} = 0, \quad \vec{e}_2 \vec{e}_{23} = 0, \quad \vec{e}_3 \vec{e}_{33} = 0. \end{aligned} \tag{12}$$

Как было получено в [6], а также ввиду симметричности векторов \vec{e}_{ij} по нижним индексам, примем следующее:

$$\begin{aligned} \vec{e}_{11} = a_{11}^2 \vec{e}_2 + a_{11}^3 \vec{e}_3, \quad \vec{e}_{22} = a_{22}^1 \vec{e}_1 + a_{22}^3 \vec{e}_3, \quad \vec{e}_{12} = a_{12}^3 \vec{e}_3, \\ \vec{e}_{23} = a_{23}^1 \vec{e}_1, \quad \vec{e}_{13} = a_{13}^2 \vec{e}_2, \quad \vec{e}_{33} = a_{33}^1 \vec{e}_1 + a_{33}^2 \vec{e}_2. \end{aligned} \tag{13}$$

В трехмерном субпроективном пространстве C^3 рассмотрим голономное двумерное многообразие, которое назовем двумерной поверхностью P^2 [8], относящуюся к семейству поверхностей, на котором находятся все линии тока крови и все вихревые линии. С поверхностью P^2 свяжем семейство реперов $R(P^2)$ так, что точка x - начало любого репера семейства реперов принадлежит поверхности, а векторы \vec{e}_i лежат в касательном пространстве в точке x к C^3 причем так, что вектор \vec{e}_3 направлен перпендикулярно векторам \vec{e}_1 и \vec{e}_2 . Так смещение $d\vec{x}$ происходит вдоль поверхности P^2 , то $\omega^3 = 0$. Тогда уравнения инфинитезимального перемещения ортогонального репера (1) примут вид:

$$\begin{aligned} d\vec{x} = \omega^1 \vec{e}_1 + \omega^2 \vec{e}_2, \quad d\vec{e}_1 = \omega_1^2 \vec{e}_2 + \omega_1^3 \vec{e}_3 + \omega^1 \vec{e}_{11} + \omega^2 \vec{e}_{12}, \\ d\vec{e}_2 = -\omega_1^2 \vec{e}_1 + \omega_2^3 \vec{e}_3 + \omega^1 \vec{e}_{21} + \omega^2 \vec{e}_{22}, \\ d\vec{e}_3 = -\omega_1^3 \vec{e}_1 - \omega_2^3 \vec{e}_2 + \omega^1 \vec{e}_{31} + \omega^2 \vec{e}_{32}. \end{aligned} \tag{14}$$

3. РОТОР ВЕКТОРА СКОРОСТИ КРОВИ

Выражение для ротора вектора скорости крови \vec{v} найдем, используя формулу:

$$\iiint \text{rot } \vec{v} d\tau = -\oint \llbracket \vec{v}, d\vec{\sigma} \rrbracket, \tag{15}$$

где

$$\begin{aligned} d\tau = d_1 \vec{x} \wedge d_2 \vec{x} \wedge d_3 \vec{x} = \omega^1 \wedge \omega^2 \wedge \omega^3 \vec{e}_1 \vec{e}_2 \vec{e}_3 = \\ = \omega^1 \wedge \omega^2 \wedge \omega^3 \end{aligned}$$

а $d\vec{\sigma}$ - вектор для элемента поверхности. Также элемент поверхности в точке x , который образован векторами $d_2 \vec{x}$ и $d_3 \vec{x}$, обозначим следующим образом: $\overrightarrow{d\sigma_{23}} = [d_2 \vec{x}, d_3 \vec{x}]$, а противоположащий ему элемент поверхности в точке $\vec{x} + d_1 \vec{x}$ обозначим как $\overrightarrow{d\sigma_{23}} + d_1(\overrightarrow{d\sigma_{23}})$.

Следует помнить, что рассуждения ведутся для точки субпроективного пространства, принадлежащей окрестности U и базис репера R_x рассматривается в этой точке.

Пусть элемент объема $d\tau$ ограничивает замкнутая поверхность, для которой

$$\begin{aligned} \overrightarrow{d\sigma_{23}} + d_1(\overrightarrow{d\sigma_{23}}) + \overrightarrow{d\sigma_{32}} + \overrightarrow{d\sigma_{31}} + d_2(\overrightarrow{d\sigma_{31}}) + \overrightarrow{d\sigma_{13}} + \\ + \overrightarrow{d\sigma_{12}} + d_3(\overrightarrow{d\sigma_{12}}) + \overrightarrow{d\sigma_{21}} = 0. \end{aligned}$$

Отсюда

$$d_1(\overrightarrow{d\sigma_{23}}) + d_2(\overrightarrow{d\sigma_{31}}) + d_3(\overrightarrow{d\sigma_{12}}) = 0. \tag{16}$$

Далее, применяя к объему $d\tau$ формулу (15), получим, как и в [9]:

$$\begin{aligned} -\text{rot } \vec{v} d\tau = [(\vec{v} + d_1 \vec{v}), (\overrightarrow{d\sigma_{23}} + d_1 \overrightarrow{d\sigma_{23}})] + [\vec{v}, \overrightarrow{d\sigma_{32}}] + \\ + [(\vec{v} + d_2 \vec{v}), (\overrightarrow{d\sigma_{31}} + d_2 \overrightarrow{d\sigma_{31}})] + [\vec{v}, \overrightarrow{d\sigma_{13}}] + \\ + [(\vec{v} + d_3 \vec{v}), (\overrightarrow{d\sigma_{12}} + d_3 \overrightarrow{d\sigma_{12}})] + [\vec{v}, \overrightarrow{d\sigma_{21}}]. \end{aligned}$$

С учетом (16), последнее примет вид:

$$\begin{aligned} -\text{rot } \vec{v} d\tau = \\ = [d_1 \vec{v}, [d_2 \vec{x}, d_3 \vec{x}]] + [d_2 \vec{v}, [d_3 \vec{x}, d_1 \vec{x}]] + [d_3 \vec{v}, [d_1 \vec{x}, d_2 \vec{x}]]. \end{aligned}$$

После преобразования векторных произведений, запишем:

$$\begin{aligned} -\text{rot } \vec{v} d\tau = d_2 \vec{x} (d_1 \vec{v} d_3 \vec{x}) - d_3 \vec{x} (d_1 \vec{v} d_2 \vec{x}) + \\ + d_3 \vec{x} (d_2 \vec{v} d_1 \vec{x}) - d_1 \vec{x} (d_2 \vec{v} d_3 \vec{x}) + \\ + d_1 \vec{x} (d_3 \vec{v} d_2 \vec{x}) - d_2 \vec{x} (d_3 \vec{v} d_1 \vec{x}), \end{aligned} \tag{17}$$

где круглыми скобками, как всегда, обозначены скалярные произведения.

Вектор скорости крови представим в виде $\vec{v} = v^i \vec{e}_i$.

Тогда

$$\vec{v} = \text{rot } \vec{v} = v^i \vec{e}_i. \quad (21)$$

$$d\vec{v} = (dv^i + v^j \omega_j^i) \vec{e}_i + v^i \omega^j e_{ij}. \quad (18)$$

С учетом (18) равенство (17) перепишем в следующем виде:

$$\text{rot } \vec{v} d\tau = -\vec{e}_i \omega^l \wedge \omega^k \wedge (dv^i + v^j \omega_j^i) (\vec{e}_i \vec{e}_k) - \vec{e}_i \omega^l \wedge \wedge v^i (\omega^k \wedge \omega^j) (\vec{e}_{ij} \vec{e}_k). \quad (19)$$

Для ортогонального репера в точке $x \in U$ получим:

$$\begin{aligned} -\text{rot } \vec{v} d\tau = & \vec{e}_1 (\omega^1 \wedge \omega^2 \wedge (dv^2 + v^1 \omega_1^2 + v^3 \omega_3^2) + \\ & + \omega^1 \wedge \omega^3 \wedge (dv^3 + v^1 \omega_1^3 + v^2 \omega_2^3)) + \vec{e}_2 (\omega^2 \wedge \omega^1 \wedge \\ & \wedge (dv^1 + v^2 \omega_2^1 + v^3 \omega_3^1) + \omega^2 \wedge \omega^3 \wedge (dv^3 + v^1 \omega_1^3 + \\ & + v^2 \omega_2^3)) + \vec{e}_3 (\omega^3 \wedge \omega^1 \wedge (dv^1 + v^2 \omega_2^1 + v^3 \omega_3^1) + \\ & + \omega^3 \wedge \omega^2 \wedge (dv^2 + v^1 \omega_1^2 + v^3 \omega_3^2)) + \vec{e}_1 (d\tau (v^1 \vec{e}_{13} \vec{e}_2 + \\ & + v^3 (e_{33} e_2)) - d\tau (v^1 (e_{12} e_3) + v^2 (e_{22} e_3))) + \\ & + \vec{e}_2 (-d\tau (v^2 (e_{23} e_1) + v^3 (e_{33} e_1)) + d\tau (v^1 (e_{11} e_3) + \\ & + v^2 (e_{21} e_3))) + \vec{e}_3 (d\tau (v^2 (e_{22} e_1) + v^3 (e_{32} e_1)) - \\ & - d\tau (v^1 (e_{11} e_2) + v^3 (e_{31} e_2))). \end{aligned}$$

С учетом (3), (12) и (13), последнее примет вид:

$$\begin{aligned} -\text{rot } \vec{v} d\tau = & \vec{e}_1 (\omega^1 \wedge \omega^2 \wedge (dv^2 + v^1 r - v^3 p) + \\ & + \omega^1 \wedge \omega^3 \wedge (dv^3 - v^1 q + v^2 p) + d\tau (v^1 (a_{13}^2 - a_{12}^3) + \\ & + v^3 a_{33}^2 - v^2 a_{22}^3)) + \vec{e}_2 (\omega^2 \wedge \omega^1 \wedge (dv^1 - v^2 r + v^3 q) + \\ & + \omega^2 \wedge \omega^3 \wedge (dv^3 - v^1 q + v^2 p) + d\tau (v^1 a_{11}^3 + \\ & + v^2 (a_{12}^3 - a_{23}^1) - v^3 a_{33}^1)) + \vec{e}_3 (\omega^3 \wedge \omega^1 \wedge (dv^1 - \\ & - v^2 r + v^3 q) + \omega^3 \wedge \omega^2 \wedge (dv^2 + v^1 r - v^3 p) + \\ & + d\tau (-v^1 a_{11}^2 + v^2 a_{22}^1 + v^3 (a_{23}^1 - a_{13}^2))). \quad (20) \end{aligned}$$

Если все коэффициенты в разложениях векторов \vec{e}_{ij} по векторам \vec{e}_k равны между собой и равны единице, то получим формулы аналогичные формулам для $\text{rot } \vec{v}$ как в [6]. Подход к моделированию движения крови, рассмотренный в [6] является частным случаем разложения векторов \vec{e}_{ij} по векторам репера R_x .

Пусть вектор вихря крови \vec{v} по базисным векторам репера R_x имеет вид:

Тогда с учетом (20) компоненты вихря будут иметь вид:

$$\begin{aligned} -d\tau v^1 = & \omega^1 \wedge \omega^2 \wedge (dv^2 + v^1 r - v^3 p) + \omega^1 \wedge \omega^3 \wedge \\ & \wedge (dv^3 - v^1 q + v^2 p) + (v^1 (a_{13}^2 - a_{12}^3) - v^2 a_{22}^3 + v^3 a_{33}^2) d\tau, \\ -d\tau v^2 = & \omega^2 \wedge \omega^1 \wedge (dv^1 - v^2 r + v^3 q) + \omega^2 \wedge \omega^3 \wedge \\ & \wedge (dv^3 - v^1 q + v^2 p) + (v^1 a_{11}^3 + v^2 (a_{12}^3 - a_{23}^1) - v^3 a_{33}^1) d\tau, \\ -d\tau v^3 = & \omega^3 \wedge \omega^1 \wedge (dv^1 - v^2 r + v^3 q) + \omega^3 \wedge \omega^2 \wedge \\ & \wedge (dv^2 + v^1 r - v^3 p) + (-v^1 a_{11}^2 + v^2 a_{22}^1 + v^3 (a_{23}^1 - a_{13}^2)) d\tau. \quad (22) \end{aligned}$$

Так как на поверхностях полной энергии имеют место формулы (8), то уравнения компонентов вихря (22) будут иметь вид:

$$\begin{aligned} -d\tau \frac{v^1}{v} = & \omega^1 \wedge \omega^2 \wedge \left(\frac{dv}{v} \sin \sigma + \cos \sigma (d\sigma + r) \right) + \omega^1 \wedge \omega^3 \wedge \\ & \wedge (-q \cos \sigma + p \sin \sigma) + (\cos \sigma (a_{13}^2 - a_{12}^3) - \sin \sigma a_{22}^3) d\tau, \\ -d\tau \frac{v^2}{v} = & \omega^2 \wedge \omega^1 \wedge \left(\frac{dv}{v} \cos \sigma - \sin \sigma (d\sigma + r) \right) - \omega^2 \wedge \omega^3 \wedge \\ & \wedge (q \cos \sigma - p \sin \sigma) + (\cos \sigma a_{11}^3 + \sin \sigma (a_{12}^3 - a_{23}^1)) d\tau, \\ 0 = & \omega^3 \wedge \omega^1 \wedge \left(\frac{dv}{v} \cos \sigma - \sin \sigma (d\sigma + r) \right) + \omega^3 \wedge \omega^2 \wedge \\ & \wedge \left(\frac{dv}{v} \sin \sigma + \cos \sigma (d\sigma + r) \right) + (-\cos \sigma a_{11}^2 + \sin \sigma a_{22}^1) d\tau. \quad (23) \end{aligned}$$

4. УРАВНЕНИЕ НЕПРЕРЫВНОСТИ ПОТОКА КРОВИ

Дивергенцию векторного поля \vec{v} можно найти точно так же, как это делали в [6]. Выпишем сразу полученный результат:

$$\begin{aligned} d\tau \text{div } \vec{v} = & (dv^1 + v^j \omega_j^1) \wedge \omega^2 \wedge \omega^3 + (dv^2 + v^j \omega_j^2) \wedge \omega^3 \wedge \omega^1 + \\ & + (dv^3 + v^j \omega_j^3) \wedge \omega^1 \wedge \omega^2 + (v^k a_{ki}^i) d\tau, \quad (24) \end{aligned}$$

где в последнем слагаемом вначале производится суммирование по i , а затем сумма по $k \neq i$.

Уравнение неразрывности потока несжимаемой крови запишется в виде:

$$\begin{aligned} (dv^1 - v^2 r + v^3 q) \wedge \omega^2 \wedge \omega^3 + (dv^2 + v^1 r - v^3 p) \wedge \omega^3 \wedge \omega^1 + \\ + (dv^3 - v^1 q + v^2 p) \wedge \omega^1 \wedge \omega^2 + (v^k a_{k1}^1 + v^k a_{k2}^2 + v^k a_{k3}^3) d\tau = 0. \end{aligned}$$

С учетом (8), последнее равенство запишется в следующем виде:

$$\begin{aligned} & \left(\frac{dv}{v} \cos \sigma - \sin \sigma (d\sigma + r) \right) \wedge \omega^2 \wedge \omega^3 + \\ & + \left(\frac{dv}{v} \sin \sigma + \cos \sigma (d\sigma + r) \right) \wedge \omega^3 \wedge \omega^1 + \\ & + (p \sin \sigma - q \cos \sigma) \wedge \omega^1 \wedge \omega^2 = 0. \end{aligned} \quad (25)$$

5. ОСНОВНЫЕ КИНЕМАТИЧЕСКИЕ УРАВНЕНИЯ

Для дальнейших вычислений введем понятие нормальной кривизны векторного поля в субпроективном пространстве в данной точке, принадлежащей карте U. Поэтому рассмотрим:

$$\begin{aligned} \frac{\vec{k}e_3}{\rho} &= -\frac{d\vec{x}de_3}{ds^2} = \\ &= -\frac{(\omega^1 \vec{e}_1 + \omega^2 \vec{e}_2)(q\vec{e}_1 - p\vec{e}_2 + \omega^1 a_{31}^2 \vec{e}_2 + \omega^2 a_{32}^1 \vec{e}_1)}{ds^2} = \\ &= \frac{-q_1(\omega^1)^2 + (p_1 - q_2 - a_{32}^1 - a_{31}^2)\omega^1 \omega^2 + p_2(\omega^2)^2}{ds^2} = \\ &= p_2 \sin^2 \sigma + (p_1 - q_2 - a_{32}^1 - a_{31}^2) \sin \sigma \cos \sigma - q_1 \cos^2 \sigma, \end{aligned}$$

где $\cos \sigma = \frac{\omega^1}{ds}$, $\sin \sigma = \frac{\omega^2}{ds}$.

Тогда

$$\begin{aligned} \frac{1}{R} &= p_2 \sin^2 \sigma + \\ &+ (p_1 - q_2 - a_{32}^1 - a_{31}^2) \sin \sigma \cos \sigma - q_1 \cos^2 \sigma, \end{aligned} \quad (26)$$

будет кривизной линии, главная нормаль которой совпадает с вектором \vec{e}_3 . Это выражение, по аналогии с [10], назовем нормальной кривизной поля в данной точке карты U. Из (26) видно, что коэффициенты разложения векторов \vec{e}_{23} и \vec{e}_{13} по векторам репера R_x входят в выражение для нормальной кривизны поля.

Рассмотрим конгруэнцию линий, являющихся интегральными линиями [11] или линиями тока векторного поля \vec{e}_3 , то вектор кривизны такой линии будет иметь вид:

$$\begin{aligned} \left(\frac{d\vec{e}_3}{ds} \right)_{\substack{\omega^1=0 \\ \omega^2=0}} &= \frac{\omega_3^1 \vec{e}_1 + \omega_3^2 \vec{e}_2 + \omega_3^3 \vec{e}_3}{ds} = \\ &= (q_3 + a_{33}^1) \vec{e}_1 - (p_3 + a_{33}^2) \vec{e}_2. \end{aligned}$$

Проекция вектора кривизны на направление скорости крови, с учетом первого равенства из (8), примет сле-

дующий вид:

$$\begin{aligned} & ((q_3 + a_{33}^1) \vec{e}_1 - (p_3 - a_{33}^2) \vec{e}_2) \frac{\vec{v}}{v} = \\ & = (q_3 + a_{33}^1) \cos \sigma - (p_3 - a_{33}^2) \sin \sigma. \end{aligned}$$

Примем:

$$\begin{aligned} L &= p_2 \sin^2 \sigma + (p_1 - q_2 - a_{32}^1 - a_{31}^2) \sin \sigma \cos \sigma - q_1 \cos^2 \sigma, \\ N &= (p_3 - a_{33}^2) \sin \sigma - (q_3 + a_{33}^1) \cos \sigma. \end{aligned} \quad (27)$$

Чтобы удовлетворить гемодинамическим уравнениям (23) и (25), примем:

$$\frac{dv}{v} = t\omega^1 + g\omega^2 + \left(L + \frac{f'}{Sv^2} \right) \omega^3,$$

$$d\sigma + r = (g + N \sin \sigma) \omega^1 - (t + N \cos \sigma) \omega^2 + u \omega^3, \quad (28)$$

где t, g, u - функции, которые выбираются таким образом, чтобы выполнялись условия интегрируемости уравнений (28).

Кроме уравнений (28) найдем условия для выполнения гемодинамических уравнений (23) и (25). Вначале проверим выполнимость уравнения (25). Для этого подставим равенства для $\frac{dv}{v}$ и $d\sigma + r$ из (28) в (25) и получим:

$$a_{33}^1 \cos \sigma + a_{33}^2 \sin \sigma = 0. \quad (29)$$

Тем самым, уравнение (25) выполнимо тогда и только тогда, когда выполняется равенство (29).

Для нахождения условия выполнимости третьего уравнения из (23), подставим равенства (28) в это уравнение и после чего получим:

$$-\cos \sigma a_{11}^2 + \sin \sigma a_{22}^1 = 0. \quad (30)$$

И, наконец, найдем условие выполнимости первого и второго уравнений из (23). Для этого подставим равенства (28) в эти уравнения и после преобразований, получим:

$$\begin{aligned} -\frac{v^1}{v} &= \left(L + \frac{f'}{Sv^2} \right) \sin \sigma + u \cos \sigma + q_2 \cos \sigma - p_2 \sin \sigma + \\ &+ \cos \sigma (a_{13}^2 - a_{12}^3) - \sin \sigma a_{22}^3, \\ \frac{v^2}{v} &= \left(L + \frac{f'}{Sv^2} \right) \cos \sigma - u \sin \sigma + q_1 \cos \sigma - p_1 \sin \sigma - \\ &- \cos \sigma a_{11}^3 - \sin \sigma (a_{12}^3 - a_{13}^2). \end{aligned}$$

Далее, первое уравнение умножим на $\sin \sigma$, а второе - на $\cos \sigma$, запишем:

$$-\frac{v^1}{v} \sin \sigma = \left(L + \frac{f'}{Sv^2} \right) \sin^2 \sigma + u \cos \sigma \sin \sigma + q_2 \cos \sigma \sin \sigma - p_2 \sin^2 \sigma + \cos \sigma \sin \sigma (a_{13}^2 - a_{12}^2) - \sin^2 \sigma a_{22}^3,$$

$$\frac{v^2}{v} \cos \sigma = \left(L + \frac{f'}{Sv^2} \right) \cos^2 \sigma - u \sin \sigma \cos \sigma + q_1 \cos^2 \sigma - p_1 \sin \sigma \cos \sigma - \cos^2 \sigma a_{11}^3 - \cos \sigma \sin \sigma (a_{12}^3 - a_{23}^1). \quad (31)$$

Деля обе части равенства (11) на Sv^2 , получим:

$$-\frac{v^1}{v} \sin \sigma + \frac{v^2}{v} \cos \sigma = \frac{f'}{Sv^2}. \quad (32)$$

Сложим первое и второе уравнения из (31), с учетом первого равенства из (27), запишем:

$$-\frac{v^1}{v} \sin \sigma + \frac{v^2}{v} \cos \sigma = \frac{f'}{Sv^2} - 2a_{12}^3 \cos \sigma \sin \sigma - a_{22}^3 \sin^2 \sigma - a_{11}^3 \cos^2 \sigma.$$

Сравнивая последнее равенство с (32), видно, что первое и второе уравнения из (23) будут верны тогда и только тогда, когда

$$2a_{12}^3 \cos \sigma \sin \sigma + a_{22}^3 \sin^2 \sigma + a_{11}^3 \cos^2 \sigma = 0. \quad (33)$$

Для удовлетворения гемодинамических уравнений, наряду с равенствами (28) мы должны принять условия (29), (30), (33). Окончательно имеем:

$$\begin{aligned} \frac{dv}{v} &= t\omega^1 + g\omega^2 + \left(L + \frac{f'}{Sv^2} \right) \omega^3, \\ d\sigma + r &= (g + N \sin \sigma) \omega^1 - (t + N \cos \sigma) \omega^2 + u \omega^3, \\ a_{33}^1 \cos \sigma + a_{33}^2 \sin \sigma &= 0, \quad a_{11}^2 \cos \sigma - a_{22}^1 \sin \sigma = 0, \\ a_{11}^3 \cos^2 \sigma + 2a_{12}^3 \cos \sigma \sin \sigma + a_{22}^3 \sin^2 \sigma &= 0. \end{aligned} \quad (34)$$

ЗАКЛЮЧЕНИЕ

В уравнениях (31) получаем новые неизвестные функ-

ции: t, g, u . Эти функции выбираются таким образом, чтобы выполнялись условия интегрируемости первых двух уравнений (31). Полученные уравнения являются основными кинематическими уравнениями потока крови в субпроективном пространстве и служат исходными уравнениями для изучения потока крови, когда последняя движется с завихрениями, но поток крови и линии вихря располагаются на однопараметрическом семействе поверхностей, которые называются поверхностями полной энергии. Другими словами, репер второго порядка, рассматриваемый в окрестности U точки x субпроективного пространства, содержит симметричные по нижним индексам вектора второго порядка. Точка x связана с частицей движущейся крови. Тем самым, моделью сердечно - сосудистой системы является субпроективное пространство, а сосудам соответствуют геодезические линии данного пространства.

Изучение движения крови ведется на основании изучения геометрии конгруэнции линий тока и вихревых линий. В рассматриваемом случае в окрестности U задано голономное распределение, интегральным многообразием которого являются поверхности полной энергии.

ПЕРЕЧЕНЬ ССЫЛОК

1. Ноздрачев А.Д., Баженов Ю.И. и др. Общий курс физиологии человека и животных: В 2-х книгах - М.: Высшая школа, 1991. - Кн. 2. - 528 с.
2. Афромеев В.И., Протопопов А.А., Фильчакова В.П., Яшин А.А. Математические методы современной биомедицины и экологии. Монография / Под общ. ред. Нефедова Е.И., Хадарцева А.А., Яшина А.А.- Тула: ТулГУ, НИИ НМТ, 1997. - 223 с.
3. Кузнецов Г.В. Особенности и эффективность пространственного подхода к моделированию сердечно-сосудистой системы человека // Вестник новых медицинских технологий. - 2000. - Т. 7, № 2. С. 45 - 47.
4. Каган В.Ф. Субпроективные пространства. - М.: ГИФМА, 1961. - 220 с.
5. Аквис М.А. Многомерная дифференциальная геометрия. - Калинин: КГУ, 1977. - 83 с.
6. Кузнецов Г.В., Яшин А.А. Основы математической теории моделирования сердечно-сосудистой системы человека в субпроективном пространстве // Вестник новых медицинских технологий. - 1999. - Т. 6, № 1. - С. 42 - 45.
7. Кузнецов Г.В. О конформном соответствии между областями евклидова и риманова пространств // Дифференциальная геометрия многообразий фигур. Калининград, 1998. Вып. 29. С. 31 - 35.
8. Постников М.М. Лекции по геометрии. Семинар 5. Риманова геометрия. - М.: Изд-во "Факториал", 1998. 496 с.
9. Бюшгенс С.С. Геометрия стационарного потока идеальной несжимаемой жидкости // Изв. АН СССР. Серия: Математика. - 1948. - Т. 12. - С. 481 - 512.
10. Бюшгенс С.С. Геометрия векторного поля // Изв. АН СССР. Серия: Математика. - 1946. - Т. 10. - С. 73 - 96.
11. Базылев В.Т. Геометрия дифференцируемых многообразий. - М.: Высшая школа, 1989. - 221 с.

ИДЕНТИФИКАЦИЯ СЛОЖНЫХ ДИНАМИЧЕСКИХ СИСТЕМ ПРИ ПОМОЩИ Д-СЕТИ

Е.Н.Литвинов, А.С.Лихоузов, Т.А.Лихоузова

Робота кається проблемі ідентифікації складних динамічних систем. Розроблено алгоритм побудови Д-сеті структурованих моделей для ідентифікації системи з змінною структурою. Для оцінки якості моделей використовується інформаційний критерій.

Робота пов'язана з проблемою ідентифікації складних динамічних систем. Розроблено алгоритм синтезу Д-мережі структурованих моделей для ідентифікації системи зі змінною структурою. Для оцінки якості моделей використовується інформаційний критерій.

The work concerns a problem of identification of complex dynamic systems. The algorithm of construction of a D-set of the structured models for identification of system with variable structure is developed. For an estimation of quality of models the information criterion is used.

При решении задачи структурной идентификации сложной динамической системы необходимо найти множество систем, из некоторого допустимого класса систем, которое является наилучшей моделью для заданной системы данных. Под термином "наилучшая" подразумевается, что искомая модель способна генерировать данные, наиболее близкие к заданным данным, или, говоря другими словами, имеет наилучшее приближение (наименьшее несоответствие) к данным. Модель также должна быть самой простой из множества доступных моделей. Кроме критериев простоты и точности, могут применяться и другие критерии.

Для решения этой задачи необходимо выполнить ряд этапов. Сначала нужно задать исходную систему на основании априорной информации, которая имеется о системе, например, количество переменных, количество учитываемых категорий значений этих переменных и количество измерений, необходимое для формирования системы более высокого уровня - системы данных. Пусть задана система из n переменных

$$X = \{X_i\}, i = \overline{1, n} \quad (1)$$

и последовательность N одновременных измерений X переменных. Тогда можно составить матрицу данных

$$D = X \times N, \quad (2)$$

или матрицу активности, если переменные в матрице упорядочены по времени. Эти данные необходимо отквантовать с целью удовлетворения вычислительного ограничения и ограничения данных [1].

Далее в системе данных выбирается фрагмент, который будет использоваться в качестве обучающей выборки. Для этого фрагмента должно быть априорно известно коли-

чество участков стационарности (т.е. участков с неизменной структурой системы). На основании гладкости функции энтропии определяются границы участков стационарности в точках, где

$$\Delta H \geq \Delta H_{\max} \quad (3)$$

После этого для каждого участка стационарности строится модель порождающей системы, а затем модель структуры системы. Известно [2-5] несколько подходов к построению модели порождающей системы. Если количество переменных в системе достаточно мало (с точки зрения вычислительных ограничений), то для синтеза модели целесообразно использовать алгоритмы, предложенные в [2], в которых маска выбирается одинаковой для всех переменных X системы. Достоинствами этих алгоритмов являются простота и быстроедействие, а недостатком - очень ограниченное количество переменных в системе. Подход, используемый в алгоритмах [3-5], позволяет несколько расширить их область применимости по сравнению с алгоритмами [2], поскольку маска в них выбирается отдельно для каждой переменной X и её мощность последовательно наращивается. Недостатком этих алгоритмов является большая длительность выполнения.

Для построения моделей структурированных систем в настоящее время существует множество методов. В работе [7] дается обобщение и сравнительная характеристика алгоритмов структурной идентификации, разработанных в рамках различных подходов к построению структурной теории систем. На основании этого сравнения делается вывод, что наиболее перспективным выглядит клировский [2,7-11], поскольку он позволяет использовать одни и те же алгоритмы для идентификации систем различных классов. На базе этого подхода разработан ряд алгоритмов структурной идентификации, использующих решетки уточнения моделей структур с оценением их характеристик по специальным критериям [2,5,7-11].

На последнем этапе структурированные модели объединяются в Д-сеть. Для обнаружения необходимости смены модели можно использовать модифицированный вариант предложенного в [12] алгоритма. Пусть в каждый дискретный момент времени измеряется случайный вектор $x(t)$ параметров системы. Путем обработки данных $x(t)$, $x(t-1), \dots, x(t-k)$ определяется значение энтропии $H(t)$ в текущий момент времени и её приращение ΔH (k выбирается в соответствии с ограничением данных [1]), а также распределение вероятностей $f(X)$ на участке $(t-k, t)$. Для описания системы задано конечное множе-

ство моделей с индексами $j = 1, 2, \dots, n$. Предположим, что в некоторый момент времени ΔH превысило ΔH_{\max} . Находим модель, удовлетворяющую условию

$$G(f, f^p) = \min_j G(f, f^j), \quad (4)$$

где G - информационное расстояние между исходной системой и моделью. Поскольку сравниваемые распределения вероятностей являются произвольными, необходима обобщенная мера информационного расстояния [2], определенная формулой

$$G(f, f^p) = D\left(f, \frac{f + f^p}{2}\right) + D\left(\frac{f + f^p}{2}, f^p\right), \quad (5)$$

где f и f^p - произвольные распределения вероятностей, определенные на одном и том же множестве состояний; $(f + f^p)/2$ - распределение вероятностей, полученное взятием среднего для каждой пары соответствующих вероятностей из f и f^p ; D - информационное расстояние, заданное уравнением

$$D(f, f^h) = \frac{1}{\log_2 |C|} \sum_{c \in C} f(c) \log_2 \frac{f(c)}{f^h(c)}, \quad (6)$$

где константа $1/\log_2 |C|$ - нормирующий коэффициент, обеспечивающий выполнение соотношения

$$0 \leq D(f, f^h) \leq 1. \quad (7)$$

Можно также использовать другие меры сравнения моделей, например, предложенные в [2,11].

Введем статистику

$$m(p) = \begin{cases} m(p-1) + 1, & \text{если } j_p(p) \neq j_p^*(p-1) \text{ и } j_p(p) = j_p(p-1), \\ 1, & \text{если } j_p(p) \neq j_p(p-1) \text{ и } j_p^*(p) \neq j_p^*(p-1), \\ 0, & \text{если } j_p(p) = j_p^*(p-1). \end{cases} \quad (8)$$

$m(0) = 0$. Критерием обнаружения вида модели служит достижение статистикой $m(p)$ пороговой величины h , т.е. повторение рекомендаций одного вида модели h раз подряд. Индекс модели, используемой для нахождения текущих оценок, выбирается в соответствии с правилом

$$j^*(p) = \begin{cases} j_p^*(p-1) & \text{при } m(p) < h, \\ j_p(p) & \text{при } m(p) \geq h. \end{cases} \quad (9)$$

После обнаружения изменения вида модели в момент времени $t = p$, ($m(p) = h$) и установления равенства $j^*(p) = j_p(p)$ рекуррентное вычисление статистики (7)

возобновляется при начальном условии $m(p) = 0$.

Значение $j^*(0)$, не оказывающее существенного влияния на оценки при достаточно большом t , определяется видом модели, наиболее подходящим для начала текущего оценивания. Пороговая величина h может быть подобрана по обучающей выборке или оценена в соответствии с рекомендациями, приведенными в [12].

Объединяя вышесказанное, можно сформулировать следующий алгоритм идентификации сложных динамических систем:

1. Формирование исходной системы на основании априорной информации;
2. Формирование системы данных на основании вычислительных ограничений [7];
3. Выделение фрагмента системы данных в качестве обучающей выборки (при этом желательно, чтобы априорно было известно количество участков стационарности);
4. Определение границ участков стационарности по приращению функции энтропии;
5. Выбор наилучшей маски и построение модели порождающей системы [1,5] для каждого участка стационарности;
6. Построение модели структурированной системы [1, 11,15] для каждого участка стационарности;
7. Подбор параметров переключателя Д-сети.

ПЕРЕЧЕНЬ ССЫЛОК

1. Литвинов Е. Н., Лихоузов А. С., Лихоузова Т. А. Модель системы данных в анализе реконструируемости сложных динамических систем//Адаптивные системы автоматического управления, Днепропетровск, 1998
2. Клир, Дж. Системология: автоматизация решения системных задач. - М.: Радио и связь, 1990
3. Литвинов Е. Н., Лихоузов А. С., Лихоузова Т. А. Алгоритм идентификации сложных динамических систем//Проблемы создания новых машин и технологий: Сборник научных трудов КГПИ - выпуск № 1 - Кривой Рог, 1999
4. Литвинов Е. Н., Лихоузов А. С., Лихоузова Т. А. Методика идентификации многосвязных систем//Проблемы создания новых машин и технологий: Сборник научных трудов КГПИ - выпуск № 1 - Кривой Рог, 1999
5. Copant, R. Detection and analysis of dependency structures//International Journal of General Systems, 7, 1981, pp.81-91
6. Литвинов Е. Н., Лихоузов А. С., Лихоузова Т. А. Проблемы идентификации нестационарных объектов и анализ методов их решения//Проблемы создания новых машин и технологий: Сборник научных трудов КГПИ - выпуск № 1 - Кривой Рог, 1999
7. Klir, G.J. Identification of generative structure in empirical data//International Journal of General Systems, 3, N2, 1976, pp.89-104
8. Klir, G.J. On the representation of activity arrays//International Journal of General Systems, 2, N3, 1975, pp.149-168
9. Cavallo R.E., Klir G.J. Reconstructibility analysis of multidimensional relations//International Journal of General Systems, 5, N3, 1979, pp.143-171
10. Gerardy, R. Probabilistic finite-state system identification//International Journal of General Systems, 8, N4, 1982, pp.229-242
11. Литвинов Е.Н., Лихоузов А.С., Лихоузова Т.А. Алгоритм структурной идентификации сложных динамических систем //Адаптивные системы автоматического управления, Днепропетровск, 1998
12. Бывайков М.Е. Алгоритм обнаружения изменения вида модели при текущем оценивании//Автоматика и телемеханика, 1993, №5, С. 82-93.

ОСОБЕННОСТИ ПРОЦЕДУРНО-ПАРАМЕТРИЧЕСКОЙ ПАРАДИГМЫ ПРОГРАММИРОВАНИЯ

А.И.Легалов

Рассматривается новый стиль разработки программ, названный процедурно-параметрической парадигмой программирования. В основе парадигмы лежит параметрический полиморфизм, обеспечивающий расширение и модификацию программ с минимальными изменениями существующего кода. Описываются программные объекты, которые необходимо включить в языки программирования, чтобы обеспечить инструментальную поддержку этой парадигмы.

Procedure-parametric paradigm of programming is proposed. It's the new programming style, which uses parametric polymorphism for expanding and modification of programs and leads to minimal change of existing code. Program objects, which support this paradigm in programming languages, are described.

ВВЕДЕНИЕ

Одним из главных критериев программных продуктов является сложность, а основными требованиями, предъявляемыми к методологиям разработки программного обеспечения, считаются: удобство сопровождения, возможность безболезненного наращивания уже существующей программы, возможность повторного использования разработанных программных объектов [1]. Этим требованиям во многом соответствует объектно-ориентированная методология (ООМ), которая, в настоящее время, получила наибольшее распространение. Составной частью ООМ является объектно-ориентированное программирование (ООП). Развитие ООП практически полностью вытеснило процедурное программирование из областей, связанных с разработкой сложных программных систем. Это обусловлено тем, что, несмотря на возможность написания объектно-ориентированных (ОО) программ с использованием процедурных языков, применение последних ведет к ошибкам, которые можно обнаружить лишь на более поздних этапах построения программы, чаще всего, во время выполнения. Такая ситуация возникает из-за того, что процедурные языки лишь моделируют объектную парадигму структурными средствами, используя при написании кода только динамические проверки состояния объектов и их типа. Применение объектно-ориентированных языков и инструментов зачастую позволяет избавиться от подобных ошибок уже во время компиляции.

Вместе с тем, практическое использование ОО подхода позволило выявить ряд недостатков [2, 3]. Их преодоление осуществляется так же, как и при использовании процедурных языков: моделированием с помощью языковых средств объектов, решающих проблемы, связанные с использованием объектно-ориентированных средств. В качестве одного из успешных примеров такого решения можно привести разработку образцов проектирования [4]. Однако использование моделирования вновь сдвигает

процесс обнаружения ошибок на более поздние этапы. Это показывает, что поиск новых методов и стилей программирования и сейчас является достаточно актуальным.

Одним из возможных путей, предложенным в работе [3], является разработка инструментальных средств поддерживающих процедурно-параметрическое программирование (ППП). Опираясь на процедурный подход, данная парадигма расширяет его механизмом, поддерживающим обработку альтернативных объектов без их явного описания в теле одной процедуры. Этот механизм похож на полиморфизм классов, используемый в ООП. В данной работе рассматриваются общие особенности процедурно-параметрического программирования и способы представления программных объектов, обеспечивающих его поддержку.

1. ОСОБЕННОСТИ ПРОЦЕДУРНО-ПАРАМЕТРИЧЕСКОГО ПОДХОДА

Сравним между собой процедурный, объектно-ориентированный и процедурно-параметрический подходы. Основное отличие проявляется в организации языковых средств, обеспечивающих обработку альтернативных данных. Такая обработка встречается практически в любой задаче и связана с выполнением преобразований, проводимых над коллекциями разнотипных элементов, принадлежащих к одной категории. В качестве примера можно привести процедуру *draw*, осуществляющую рисование одной из геометрических фигур, включенной в состав некоторого изображения. Разные фигуры (круг, прямоугольник и т.д.), образующие изображение, принадлежат к одной категории (*shape*).

1.1. Организация процедурной программы

При процедурном подходе множество структур данных описывают отдельные геометрические фигуры. Для каждой фигуры создается специализированная процедура отображения. Для рисования одной из альтернатив формируется вариантное объединение, которое содержит признак хранимой фигуры и поле для хранения одной из фигур. Процедура отображения осуществляет анализ признака и рисует ту фигуру, которая соответствует этому признаку. Код, реализующий рассмотренные программные объекты, может быть разнообразным. Поэтому, представленный ниже фрагмент достаточно схематично отображает полученный результат. Для иллюстрации используется язык программирования C++.

```
//Геометрические фигуры и специализированные процедуры их обработки
struct circle {int r;}; // Круг радиуса r
void draw_circle (circle &c) {...} // Рисование круга
```

```
struct rectangle {int a,b;}; // Прямоугольник со сторонами a, b
draw_rectangle (rectangle &r) {...} // Рисование прямоугольника
// Обобщенная фигура
enum shape_type {CIRCLE, RECTANGLE};
struct shape {
    shape_type key;
    union {
        circle c;
        rectangle r;
    } value;
};
// Обобщенная процедура рисования
void draw (shape &s) {
    switch(s.key) {
        case CIRCLE: draw(s.value.c); break;
        case RECTANGLE: draw(s.value.r); break;
    }
}
```

Объединительной основой служит обобщающая процедура, в которой обрабатываются данные. Добавление новых методов обработки существующих данных не ведет к модификации написанного кода. Однако изменение числа фигур ведет к переделке всех процедур, обрабатывающих обобщенные фигуры. В реальной ситуации именно данные подвержены наибольшей модификации, а набор процедур, определяющий интерфейс программы, остается достаточно консервативным. Поэтому, процедурный подход не может обеспечить эффективного сопровождения при разработке больших программных систем, когда одно небольшое изменение данных ведет к огромному количеству изменений в процедурах.

1.2. Организация объектно-ориентированной программы

При объектно-ориентированном подходе первоначальный акцент смещается на проектирование (зачастую, абстрактное) сигнатур процедур, обеспечивающих обработку обобщающей фигуры. Создание такого обобщения предоставляет интерфейс и каркас для построения конкретных реализаций фигур на основе наследования. Оболочкой для хранения процедур, обрабатывающих конкретные фигуры, служит класс. Поэтому, появление новых фигур ведет к созданию нового класса и связанных с ним процедур без модификации уже написанного кода. Такое ортогональное проектирование оказывается ближе к реальной разработке больших программных систем. В качестве примера можно привести следующий фрагмент кода.

```
// Обобщенная геометрическая фигура
class shape {
    public: virtual void draw()=0;//Чистый метод, определяющий
    //рисование фигуры
    ... // Прочие члены класса
};
// Добавляемые геометрические фигуры
class circle { // Круг
    public: void draw() {...} // Рисует круг
    ... // Прочие члены класса
};
class rectangle { // Прямоугольник
    public: void draw() {...} // Рисует прямоугольник
    ... // Прочие члены класса
};
```

Пример показывает, что в ОО программе происходит распределение обработчиков отдельных специализаций по классам, представляющим конкретные моделируемые объекты.

1.3. Моделирование процедурно-параметрического механизма

Процедурно-параметрический механизм управления программой использует однозначную зависимость между данными, описывающими конкретные объекты, и процедурами их обработки. Как и при процедурном подходе, каждому из объектов (геометрической фигуре) ставится в соответствие признак, а общая для всех фигур процедура обработки (*draw*) использует этот признак в качестве ключа, определяющего конкретный обработчик. Отличие заключается в применении индексации вместо алгоритмического перебора вариантов, что позволяет обойтись без написания кода, анализирующего альтернативы. Признак, кроме идентификации специализации, выполняет функции переключателя на нужную процедуру. Моделирование процедурно-параметрического механизма можно осуществить с использованием процедурного языка:

```
// Геометрические фигуры и их специализированные
// обработчики
struct circle {int r;}; // Круг радиуса r
void draw_circle (circle &c) {...} // Рисование круга
struct rectangle {int a, b;}; // Прямоугольник со сторонами a, b
draw_rectangle (rectangle &r) {...} // Рисование прямоугольника
// Модель параметрического обобщения
enum shape_type {CIRCLE = 0, RECTANGLE};
struct shape {
    shape_type key; // Признак параметрического обобщения
    void *value; // Указатель на конкретную фигуру
};
//Обобщенная параметрическая процедура рисования строится из:
// - отдельных обработчиков специализаций обобщенной фигуры
void draw_shape_circle(shape &s) {draw_circle(*((circle*)(s.value)));}
void draw_shape_rectangle(shape &s) {draw_rectangle(*((rectangle*)(s.value)));}
// - массива указателей на эти обработчики
int const shape_number = 2; // Количество фигур
void (*draw[shape_number]) (shape&s) = {draw_shape_circle,
draw_shape_rectangle};
```

При выполнении ПП программы осуществляется создание конкретной специализации (например, *circle*) и привязка к указателю *value* объекта *shape*:

```
shape s; s.key = CIRCLE; s.value = new circle; s.value->r = 5;
```

Вызов обработчика специализации осуществляется следующим образом:

```
draw [s.key] (s);
```

Следует отметить, что использование массивов указателей на процедуры не является новым словом в программировании, однако, в работе основной упор делается на создание языковой надстройки, скрывающей параметрический механизм за более эффективными средствами написания программы. Сформированный программный код обобщает ранее разработанные программные объекты (геометрические фигуры и их обработчики). Добавление или изменение данных не ведет к модификации написанных процедур, а затрагивает только массив указателей на обработчики специализаций. Конечно, в процедурном языке и эти модификации можно считать существенными. Поэтому, использование специальных языковых конструк-

ций, определяющих процедурно-параметрический язык программирования, позволит разработчику не менять уже написанный код.

2. ОРГАНИЗАЦИЯ ОБЪЕКТОВ ПРОЦЕДУРНО-ПАРАМЕТРИЧЕСКОГО ЯЗЫКА

2.1. Основные понятия процедурно-параметрической парадигмы

Основными объектами, используемыми в процедурно-параметрическом программировании, являются:

1. *Параметрические обобщения*, поддерживающие однозначное соответствие между признаками и специализациями.

2. *Обобщающие параметрические процедуры*, определяющие сигнатуры обработчиков параметрических обобщений.

3. *Обработчики параметрических специализаций*, обеспечивающие манипуляции с конкретными экземплярами данных, включенными в состав параметрических обобщений.

4. *Экземпляры параметрических обобщений*, представляющие конкретные программные объекты, обрабатываемые в ходе вычислений.

5. *Вызовы параметрических процедур*, обеспечивающие доступ к отдельным обработчикам специализаций через единые сигнатуры, определяемые обобщающими параметрическими процедурами.

Процедурно-параметрический полиморфизм - это способ написания программы, позволяющий процедурам обрабатывать объекты разных типов. При этом процедурно-параметрическое программирование позволяет модифицировать состав типов обрабатываемых объектов без изменения кода уже написанных процедур.

Механизм параметрического обобщения - совокупность понятий и объектов, обеспечивающих поддержку процедурно-параметрического программирования. Различные способы организации механизма параметрического обобщения определяют множество методов процедурно-параметрического полиморфизма.

2.2. Организация параметрических обобщений

Параметрическое обобщение позволяет сгруппировать в единое понятие не связанные ранее программные объекты, однозначно сопоставляя их с признаками. *Признак* - это ключ, идентифицирующий программный объект в рамках рассматриваемого контекста, задающего некоторую категорию. Множество признаков: $T = \{t_1, t_2, \dots, t_3, \dots, t_n\}$, где n - количество признаков, используемых для построения параметрического обобщения. Каждому t_i ставится в соответствие специализация $S_j = \{s_{j1}, s_{j2}, \dots, s_{jq}, \dots, s_{jk}\}$, где S_j - программный объект, выступающий в роли специализации обобщения. $S_j \subset S$, множеству программных объектов, используемых в качестве специализаций обобщения. Множество $\{s_{j1}, s_{j2}, \dots, s_{jq}, \dots, s_{jk}\}$ - это значения,

принимаемые специализацией S_j .

Параметрическое обобщение задается отображением $U: T \times S$, где $u_i = (t_i, S_j)$ определяет отдельный элемент параметрического обобщения. При этом $|T| \geq |S|$, то есть, различным признакам могут ставиться в соответствие одинаковые программные объекты, определяющие специализации. Это может быть связано с различными способами использования одинаковых программных объектов. Следует также отметить, что все признаки, специализации и параметрические обобщения принадлежат множеству различных программных объектов $X: T \subset X$ и $S \subset X$.

2.3. Организация обобщающих параметрических процедур

Обобщающая параметрическая процедура F является отображением:

$$F: U_1 \times U_2 \times \dots \times U_n \times Z_1 \times Z_2 \times \dots \times Z_m \rightarrow Y_1 \times Y_2 \times \dots \times Y_k,$$

где U_1, U_2, \dots, U_n - параметрические аргументы, являющиеся параметрическими обобщениями, Z_1, Z_2, \dots, Z_m - прочие аргументы, Y_1, Y_2, \dots, Y_k - результаты отображения. Аргументы и результаты отображения принадлежат множеству X допустимых программных объектов: $U_i \in X, Z_j \in X, Y_l \in X$.

Игнорируем особенности аргументов Z_1, Z_2, \dots, Z_m и результатов Y_1, Y_2, \dots, Y_k , обозначив их через W . Кроме того, нас будут интересовать не столько отображения, сколько параметрические аргументы. Тогда обобщающую параметрическую процедуру можно будет представить как:

$$\begin{aligned} F: U_1 \times U_2 \times \dots \times U_n \times Z_1 \times Z_2 \times \dots \times Z_m &\rightarrow Y_1 \times Y_2 \times \dots \times Y_k \equiv \\ &\equiv F: U_1 \times U_2 \times \dots \times U_n \times Z_1 \times Z_2 \times \dots \times Z_m \times Y_1 \times Y_2 \times \dots \times Y_k \equiv \\ &\equiv F: U_1 \times U_2 \times \dots \times U_n \times W. \end{aligned}$$

Сигнатуру обобщающей процедуры можно представить в обычной функциональной форме: $F(U_1, U_2, \dots, U_n, W)$, где U_1, U_2, \dots, U_n - параметрические аргументы, а W - прочие аргументы и результаты. Для удобства восприятия параметрических аргументов предлагается записывать их отдельно от прочих аргументов, охватывая квадратными скобками:

$$F[U_1, U_2, \dots, U_n](W) \equiv F(U_1, U_2, \dots, U_n, W).$$

Наряду с разделением аргументов, такая запись подчеркивает один из возможных способов реализации механизма параметрического полиморфизма посредством многомерных массивов указателей на обработчики параметрических специализаций. Конструкцию $[U_1, U_2, \dots, U_n]$

назовем параметрическим списком.

Основной задачей обобщающей параметрической процедуры является предоставление единого интерфейса для множества обработчиков параметрических специализаций. Обобщающая параметрическая процедура поддерживает выполнение следующих функций:

1. Обеспечивает доступ к требуемому обработчику параметрической специализации в зависимости от значения аргументов параметрического списка посредством параметрического полиморфизма. Реализация механизма доступа зависит от реализации параметрического обобщения и организации процесса обработки параметрических специализаций.

2. При необходимости, предоставляет обработчик параметрической специализации по умолчанию. Его можно использовать при отсутствии обработчика соответствующей специализации. Предоставление обработчика по умолчанию не является обязательным. В этом случае отсутствие обработчиков специализаций должно привести к выдаче ошибки во время компиляции, компоновки или выполнения. Момент выдачи такой ошибки определяется периодом окончательного формирования параметрического обобщения.

3. Обобщающая параметрическая процедура может предоставить и обработчик исключения, если отсутствует один из обработчиков специализации. Обработчик исключения можно использовать как во время компиляции, так и во время выполнения. Предполагается, что, при отсутствии обработчиков параметрических специализаций можно выбрать обработчик по умолчанию или обработчик исключения. Присутствие этих обработчиков предполагается определять опционально. Корректность логики функционирования программы при их отсутствии должна определяться средой разработки и исполнительной системой.

2.4. Организация обработчиков параметрических специализаций

Обработчики параметрических специализаций имеют сигнатуру, аналогичную сигнатуре обобщающей параметрической процедуры. Но они предназначены для обработки только определенных комбинаций значений параметрических аргументов. Комбинация, на которую "настроен" конкретный обработчик параметрической специализации, задается значениями признаков при описании его сигнатуры:

$$F[U_1(t_{1i}), U_2(t_{2j}), \dots, U_n(t_{nk})](X_m),$$

где $t_{qr} \in T_q$ - одно из значений признака, принадлежащего множеству признаков T_q параметрического обобщения U_q . Множество обработчиков некоторой обобщающей параметрической процедуры отличаются друг от друга тем, что каждый имеет уникальную комбинацию признаков, определяющую ту специализацию параметрического списка, которую он может обработать. Тело каждого обработчика содержит свой набор инструкций. Максимально возможное число обработчиков

параметрических специализаций определяется как произведение мощностей признаков обобщений, входящих в параметрический список:

$$N = |T_1| \times |T_2| \times \dots \times |T_n|,$$

где T_i - множество признаков параметрического обобщения U_i .

При отсутствии обработчиков параметрических специализаций для некоторых комбинаций, их можно заменить обработчиком по умолчанию или обработчиком исключений.

2.5. Экземпляр параметрического обобщения

Экземпляром параметрического обобщения является размещаемый в памяти компьютера программный объект, для которого определено и зафиксировано одно из конкретных значений элемента отображения $u_i = (t_i, S_j)$.

При этом, во время работы программы, элемент S_j может принимать любое значение $s_{jq} \in S_j$. Значение признака t_i в ходе вычислений должно оставаться неизменным.

2.6. Вызовы параметрических процедур

Вызовы параметрических процедур используются в ходе выполнения программы. Значения, подставляемые вместо списка формальных параметров, должны соответствовать сигнатуре обобщающей параметрической процедуры. В зависимости от значений признаков экземпляров параметрических обобщений, используемых в качестве фактических параметров, осуществляется выполнение тела соответствующей специализации. При отсутствии обработчика специализации для подставляемого параметрического списка происходит вызов обработчика по умолчанию или запускается соответствующий обработчик исключения.

2.7. Возможная реализация процедурно-параметрических понятий

Ниже представлен фрагмент кода рисования геометрических фигур, написанный с использованием воображаемых процедурно-параметрических конструкций, встроенных в язык программирования, похожий на C++.

```
// Реализуемые геометрические фигуры и их обработчики
struct circle {int r;}; // Круг радиуса r
void draw_circle (circle &c) {...} // Рисование круга
struct rectangle {int a, b;}; // Прямоугольник со сторонами a, b
draw_rectangle (rectangle &r) {...} // Рисование прямоугольника
// Параметрическое обобщение
union shape switch(auto) { // Использует автоматическое
// формирование признака
    case CIRCLE: circle;
    case RECTANGLE: rectangle;
};
// Обобщенная параметрическая процедура
void draw[shape&s]( ) = 0; // отсутствуют обработчики
// по умолчанию и исключений
// Обработчики специализаций
```

```
void draw[shape<CIRCLE> &s] {draw(s);} // Нарисует круг
void draw[shape<RECTANGLE>&s] {draw(s);} // Нарисует
// прямоугольник
```

Создать требуемый экземпляр круга можно так:
`shape<CIRCLE> c;`

Дальнейшая работа с этим объектом осуществляется как с обычным кругом:

```
c.r = 5;
```

Вызов обработчика производится следующим образом:

```
draw [c] ();
```

Более подробное описание других аспектов ППП представлено в работе [3].

ЗАКЛЮЧЕНИЕ

Использование объектов, поддерживающих процедурно-параметрическое программирование, позволяет расширить функциональные возможности существующих процедурных языков, сделать их более приспособленными для

разработки больших программных систем. Процедурно-параметрическую парадигму можно использовать в качестве альтернативы объектно-ориентированному подходу. Интересные эффекты могут быть получены и при совместном использовании ООП и ППП [3].

ПЕРЕЧЕНЬ ССЫЛОК

1. Буч Г. Объектно-ориентированный анализ и проектирование с примерами приложений на C++, 2-е изд./Пер. с англ. - М.: Издательства Бином, СПб: Невский диалект, 1998 г. - 560 с., ил.
2. Moessenboeck H., Wirth N. The Programming Language Oberon-2. Institut fur Computersysteme, ETH Zurich, July, 1996.
3. Легалов А.И. Процедурно-параметрическая парадигма программирования. Возможна ли альтернатива объектно-ориентированному стилю? - Красноярск: 2000. Деп. рук. № 622-В00 Деп. в ВИНТИ 13.03.2000. - 43 с.
4. Gamma E., Helm R., Johnson R., Vlissides J. Design Patterns: Elements of Reusable Object-Oriented Software. - ISBN 0-201-63442-2 Hardback, 416 p. 1995.

УДК 519.72: 616.8 - 091.81

АНАЛИЗ ФОРМАЛЬНО-ЛОГИЧЕСКОЙ МОДЕЛИ НЕЙРОНА С ПОМОЩЬЮ ДИСКРЕТНОГО ПРОСТРАНСТВА СОСТОЯНИЙ

В.Н.Лопин

Рассмотрен метод комплексного анализа формально-логической модели нейрона с позиций его надежности и сложности, использующий дискретное пространство состояний. Полученные результаты позволяют подойти к решению задачи синтеза надежных искусственных нейронных сетей и объяснить возможный механизм надежности биологических нейронных сетей.

Considered method of complex analysis formally-logical neuron model from positions of its reliability and difficulty, using discrete condition space. Tinned results allow to decide the problem of syntheses reliable artificial neuronal nets and explain a possible mechanism of reliability biological neuronal nets.

ВВЕДЕНИЕ

Одним из важнейших достоинств биологических нейронных сетей является их высокая надежность. Этим можно объяснять постоянный интерес к изучению данной проблемы [1-4]. Исследования в этом направлении позволяют лучше понять высокоэффективный механизм функционирования нейронных структур и использовать полученные знания при создании искусственных информационных систем принципиально нового поколения. Можно отметить, что к настоящему времени еще не сложился единый подход к объяснению феномена высокой надежности и целостного механизма организации биологических нейронных сетей. Несомненно, ввиду чрезвычайной сложности этой проблемы, исследования надежности необходимо проводить комплексно, с учетом других взаимосвязанных характеристик нейронных сетей.

В данной работе делается попытка комплексного подхода к объяснению механизма высокой надежности нейрона с позиций его пространства состояний [5], а также к установлению связи надежности нейрона с его сложностью и внутренним состоянием.

АНАЛИЗ ФОРМАЛЬНО-ЛОГИЧЕСКОЙ МОДЕЛИ НЕЙРОНА

Введем в рассмотрение, без потери общности, простейшую формально-логическую модель нейрона, реализующую все булевы функции одной переменной:

$$y = \text{Sign}(B \cdot x + T), \quad (1)$$

где y - выходной сигнал, x - входной сигнал, B - вес входа, T - порог,

$$\text{Sign}(A) = \begin{cases} 0, & A < 0 \\ 1, & A \geq 0 \end{cases}.$$

Состоянием нейрона назовем пару действительных чисел (B, T) , тогда пространство состояний нейрона определится как:

$$R = \{(B, T)\}. \quad (2)$$

Известно, что все булевы функции одной переменной являются пороговыми, обозначим их совокупность через Y . Это означает, что всегда существует такое состояние

$$(B, T) \subset R, \quad (3)$$

что любую булеву функцию

$$y_r \subset Y, \quad r = 1, 2, 3, 4 \quad (4)$$

можно представить в виде:

$$y_r = \text{Sign}(B \cdot x + T). \quad (5)$$

Следовательно, модель (1) можно рассматривать как оператор L , осуществляющий отображение пространства состояний в множество функций:

$$R \xrightarrow{L} Y. \quad (6)$$

Обратное отображение функции y_r , т.е.

$$L^{-1}y_r = \{(B, T): \text{Sign}(B \cdot x + T) = y_r\} = R_r, \quad (7)$$

является подпространством:

$$R_r \subset R. \quad (8)$$

Легко проверить, что

$$\begin{aligned} R_1 &= \{(B, T): B \geq -T, T < 0\}; \\ R_2 &= \{(B, T): B < -T, T \geq 0\}; \\ R_3 &= \{(B, T): B \geq -T, T \geq 0\}; \\ R_4 &= \{(B, T): B < -T, T < 0\}. \end{aligned} \quad (9)$$

Следовательно, пространство R разбивается на четыре подпространства, являющихся также бесконечными множествами. Однако, наиболее подходящим для анализа является дискретное пространство состояний. При этом считается, что нейрон содержит n двоичных элементов памяти для задания B и k двоичных элементов памяти для задания T . В этом случае, его дискретное пространство состояний $R_{n,k}$ определится выражением:

$$\begin{aligned} R_{n,k} &= \\ &= \{(i, j): n - i = 2\nu, k - j = 2\mu, \nu \in \{0, \dots, n\}, \mu \in \{0, \dots, k\}\}. \end{aligned} \quad (10)$$

Очевидно, что прообразы функций в $R_{n,k}$ будут конечными множествами.

Если из n двоичных элементов памяти, задающих B , $n(+)$ находятся в состоянии $+1$ и $n(-)$ находятся в

состоянии -1 , а из k двоичных элементов памяти, задающих T , $k(+)$ находятся в состоянии $+1$ и $k(-)$ находятся в состоянии -1 , то считается, что нейрон находится в состоянии:

$$E_{(n^*, k^*)} = (n(+), n(-), k(+), k(-)) = (n^*, k^*). \quad (11)$$

Будем понимать под отказом двоичного элемента памяти, находящегося в состоянии $+1$, его переход в состояние -1 , а под отказом двоичного элемента памяти, находящегося в состоянии -1 его переход в состояние $+1$. Надежность P элемента памяти при хранении им значений $+1, -1$ предполагается одинаковой. Пусть в начальный момент нейрон находится в состоянии:

$$E_0 = (n^*, k^*), \quad (12)$$

где $|n^*| = n$, $|k^*| = k$. Тогда, при отказе ν из n двоичных элементов памяти, задающих B , и μ из k двоичных элементов памяти, задающих T , нейрон переходит в состояние:

$$E_{(n^*, k^*)} = \left(n^* - 2\nu \frac{n^*}{|n^*|}, k^* - 2\mu \frac{k^*}{|k^*|} \right). \quad (13)$$

Если при этом булева функция, заданная начальным состоянием, не искажается, т.е. выполняется условие:

$$L(E_0) = L(E_{(n^*, k^*)}), \quad (14)$$

то считается, что нейрон не отказал. При независимых элементах памяти, вероятность состояния $E_{(n^*, k^*)}$ определится выражением:

$$P_{(n^*, k^*)} = C_n^{n-\nu} \cdot P^{n-\nu} \cdot (1-P)^\nu \cdot C_k^{k-\mu} \cdot P^{k-\mu} \cdot (1-P)^\mu, \quad (15)$$

где $C_n^{n-\nu}$, $C_k^{k-\mu}$ - биномиальные коэффициенты. Надежность нейрона при реализации им некоторой булевой функции $y_r \subset R$ определится суммой вероятностей состояний, принадлежащих подпространству R_r :

$$P_r = \sum_{R_r} P_{(n^*, k^*)}. \quad (16)$$

Под надежностью нейрона на функциональном множестве Y будем понимать его минимальную надежность:

$$P_{(n,k)} = \min P_{(n^*, k^*)}. \quad (17)$$

Можно показать, что:

$$P_{(n,k)} = \sum_{R_2} C_n^{n-v} \cdot P^{n-v} \cdot (1-P)^v \cdot C_k^{k-\mu} \cdot P^{k-\mu} \cdot (1-P)^\mu. \quad (18)$$

Введем количественную оценку дискретного пространства с помощью параметра:

$$S = n + k. \quad (19)$$

Этот параметр характеризует сложность нейрона и задает конечное множество дискретных пространств, определяемое выражением:

$$F_s = \{(R_{n,k}): n + k = S\}. \quad (20)$$

Можно показать, что для каждого S существует $(S + 1)$ дискретных пространств. Отсюда, естественна следующая задача.

Для фиксированных P и S найти подпространство $R_{\alpha, \beta}$, удовлетворяющее условию:

$$P_{(\alpha, \beta)} = \max_{n+k=S} P_{(n,k)}. \quad (21)$$

Эта задача была решена на компьютере для некоторых диапазонов S и P . Результаты счета показали, что надежность нейрона возрастает с увеличением S . При фиксированном значении S максимальная надежность соответствует подпространству $R_{\alpha, \beta}$, для которого приближенно выполняется условие

$$\alpha/\beta \approx 0,5. \quad (22)$$

УДК 518.5: 532.54

НОВЫЕ ИНФОРМАЦИОННО-ВЫЧИСЛИТЕЛЬНЫЕ ТЕХНОЛОГИИ ДЛЯ РАСЧЕТА И АНАЛИЗА РЕЖИМОВ ТЕПЛОСНАБЖАЮЩИХ СИСТЕМ

Н.Н.Новицкий, В.В.Токарев, З.И.Шалагинова

Приводится описание программно-вычислительного комплекса (ПВК) "АРМ ТТС" (Автоматизированное Рабочее Место Технолога Тепловых Сетей), ориентированного на персональные компьютеры. ПВК максимально учитывает потребности практики в автоматизации гидравлических, тепловых и наладочных расчетов теплоснабжающих систем, предназначен для решения задач проектирования, эксплуатации и диспетчерского управления и может быть использован в научно-исследовательских, учебных, проектных и эксплуатационных организациях, занимающихся вопросами теплоснабжения.

The paper describes the software package "ARM TTS" (Work Station of Technologist of Heat Networks) intended for personal computers. The software package takes into account practical demands for automation of hydraulic, heat and adjustment calcu-

ЗАКЛЮЧЕНИЕ

Рассмотренная формально-логическая модель нейрона позволяет подойти к решению задачи синтеза надежных искусственных нейронных сетей с ограничениями на их сложность и функциональную мощность, а также сделать предположение о возможном механизме надежности биологической нейронной сети. Действительно, введенные при анализе состояния элементов памяти модели можно сопоставить с состояниями синаптических связей биологического нейрона в сети. В этом случае параметр S модели соответствует числу этих связей, а параметр P соответствует вероятности безотказной работы каждой связи. Тогда, для любого нейрона можно указать подпространства состояний синаптических связей, определяющие функции нейрона. Если флуктуации нейронов и связей, участвующих в формировании этих состояний, не выводят нейрон из некоторого допустимого подмножества состояний, то нейрон правильно реализует соответствующую функцию. При данной интерпретации любой нейрон сети может участвовать в формировании состояний синаптических связей для нейрона следующего уровня. Однако, вопросы анализа и синтеза нейронных сетей являются предметом отдельного исследования.

ПЕРЕЧЕНЬ ССЫЛОК

1. *Нейман Дж.* Вероятностная логика и синтез надежных организмов из ненадежных компонент. - В кн.: Автоматы /- М.: ИЛ, 1956, С. 68.
2. *Шенон К.* Работы по теории информации и кибернетике. М.: ИЛ, 1968, 750 с.
3. *Дунин - Барковский В.Л.* Информационные процессы в нейронных структурах. М.: Наука, 1978, 166 с.
4. *Горбань А.Н., Россиев Д.А.* Нейронные сети на персональном компьютере. - Новосибирск: Наука, 1996, 276 с.
5. *Дуруссо П., Рой Р., Клоуз Ч.* Пространство состояний в теории управления. М.: Наука, 1970, 619 с.

lations of heat supply systems to the greatest extent, is aimed at solving the problems of designing, operation and dispatching control and can be used at research, educational and operational organizations involved with heat supply.

ВВЕДЕНИЕ

Современные теплоснабжающие системы (ТСС) крупных городов и промышленных центров представляют собой сложные пространственные объекты большой размерности, развивающиеся во времени и объединяющие множество разнородных элементов. Наличие в системах источников теплоты разного типа (ТЭЦ, районные и

пиковые котельные), работающих на общие тепловые сети, множества потребителей, имеющих различный состав нагрузок, рассредоточенных на значительной территории и соединенных с источниками сложными протяженными и закольцованными тепловыми сетями, обуславливают сложность задач эксплуатации, управления, организации и планирования режимов ТСС. По мере развития ТСС взаимосвязи между отдельными подсистемами: источниками, тепловыми сетями и абонентскими установками еще более усиливаются. Это связано с насыщением структуры системы различного рода регулирующими элементами - регуляторами расхода, давления и т.п., присутствием промежуточных ступеней управления - контрольно-распределительных (КРП), центральных (ЦТП) и индивидуальных тепловых пунктов (ИТП) и т.д. С другой стороны, в связи с постоянно растущими ценами на тепловую энергию, потребители тепла начинают активно устанавливать теплосчетчики и средства автономного регулирования. Это приводит к переменности гидравлических режимов работы системы в целом и заставляет эксплуатационные службы пересчитывать и организовывать новые режимы.

Для принятия решений и оценки их реализуемости необходимо осуществлять расчеты гидравлических и тепловых режимов, оперативно пересчитывать и строить температурные и пьезометрические графики, осуществлять наладочные расчеты и т.д.

Актуальность расчетов тепловых и гидравлических режимов возрастает не только в связи с усложнением и ростом масштабов ТСС, но и с учетом все более усиливающихся современных требований к качеству, комфортности и надежности теплоснабжения потребителей.

Компьютерное моделирование нормальных, утяжеленных и аварийных гидравлических и тепловых режимов является одним из наиболее мощных средств отыскания и количественного обоснования разнообразных проектных, эксплуатационных и наладочных решений, эффективность которого прямо пропорциональна масштабам и сложности систем теплоснабжения, как объектов инженерного анализа.

В институте систем энергетики им. Л. А. Мелентьева (ИСЭМ) СО РАН (г. Иркутск) на протяжении целого ряда лет ведутся теоретические и прикладные исследования в области моделирования, анализа и оптимизации трубопроводных систем [1-5]. Сформулированная в ИСЭМ в начале 60-х годов *Теория гидравлических цепей (ТГЦ)*, синтезирующая на межотраслевом физико-математическом уровне общие результаты, справедливые для любых трубопроводных и гидравлических систем (ТП и ГС), стала научно-методической базой для моделирования, расчета, оценивания и оптимизации тепло-снабжающих систем и используется при разработке моделей для проектирования, управления развитием и функционированием ТСС.

ЭТАПЫ РАЗВИТИЯ ПРОГРАММНОГО ОБЕСПЕЧЕНИЯ ДЛЯ РАСЧЕТА ТСС

Программное обеспечение (ПО) для решения задач потокораспределения и расчета тепловых режимов претерпело уже несколько стадий своего развития: отдельные

программы, реализующие тот или иной метод расчета; пакеты прикладных программ (ППП), оснащенные формализованными процедурами автоматизированной обработки и анализа входной и выходной информации; программно-вычислительные комплексы (ПВК), позволяющие решать несколько сопряженных задач, а также одну и ту же задачу разными методами; диалоговые вычислительные системы (ДВС), обеспечивающие возможность активной работы с исходной информацией и оперативного управления процессом счета. Существенное развитие претерпели и собственно методы расчета потокораспределения как в отношении их универсальности, так и с точки зрения учета большого числа разнообразных факторов, в зависимости от целей моделирования.

НОВЫЕ ТРЕБОВАНИЯ К УРОВНЮ РЕАЛИЗАЦИИ

Разработанное программное обеспечение для решения задач расчета гидравлических и тепловых режимов нашло широкое практическое применение в десятках проектных, эксплуатационных и научных организаций. Вместе с тем в настоящее время возникла необходимость в разработке новых компьютерных технологий моделирования режимов в гидравлических цепях, что обусловлено следующими основными причинами:

- качественно новым уровнем вычислительной техники, связанным с появлением относительно дешевых и доступных ПЭВМ, и, как следствие, интенсивной компьютеризацией проектных и эксплуатационных организаций;
- разработкой высокоэффективных средств поддержки диалогового взаимодействия пользователя и ЭВМ, информационного обеспечения прикладных программ, управления взаимодействием между программами и др.;
- созданием новых методов расчета, системного моделирования, решения задач линейной алгебры, нелинейного программирования, оптимального управления и т.п.;
- появлением или актуализацией новых свойств и качеств собственно ТСС как объектов моделирования (таких, как масштабность, структурная и функциональная иерархичность, наличие множества неформализуемых ограничений, объективная неопределенность информации и др.), которые, с одной стороны, обусловили появление повышенных требований к традиционным методам расчета (таких как быстродействие, надежность, размерность решаемых задач и др.), а с другой, выявили и их ограниченность.

Таким образом, в качестве одного из основных принципов компьютерной реализации нового поколения программного обеспечения для решения данного класса задач должно закладываться оптимальное сочетание возможностей формализованного и эвристического моделирования.

Неподдающиеся полной формализации процедуры имеют место практически на всех этапах моделирования:

- при подготовке исходных данных к ним относятся процессы составления расчетных схем, выбора рациональной степени агрегирования отдельных фрагментов и узлов, выделения существенных ограничений и прочих факторов, назначения численных значений параметров и т.д.;
- на этапе численного решения задач это связано с

неоднозначностью выбора методов решения, эффективность и сходимость которых зависят от сочетаний исходных данных, конфигурации расчетной схемы, характеристик элементов, степени "тяжести" моделируемого режима и т.д.;

- на этапе анализа и интерпретации результатов, получаемые выводы практически полностью зависят от инженерного опыта, уровня квалификации и знаний пользователя, а также от целей моделирования. Роль машинных методов здесь сводится лишь к обеспечению наглядности и разнообразия форм представления расчетных данных.

Современные условия развития ТСС требуют адекватного подхода к созданию соответствующих информационных технологий системных исследований ТСС, включающих организацию сложных вычислительных экспериментов и интегрированных баз данных (БД) в локальных вычислительных сетях, реализацию пакетов прикладных программ, средств иллюстративной графики и инструментальных средств интеграции этих разнотипных компонентов.

НАПРАВЛЕНИЯ РАЗВИТИЯ ПО

Преодоление отмеченных выше трудностей и противоречий становится возможным на основе использования наряду с имеющимися ППП диалоговых вычислительных систем, баз данных и геоинформационных технологий с их интеграцией в информационно-вычислительные системы (ИВС), с возложением на них функций:

- обеспечения взаимосвязей математических моделей, описывающих различные процессы в ТСС;
- создания единой информационной базы для различных задач, доступной для непрограммирующего специалиста (инженера-теплоэнергетика);
- объединения и обмена разнородными информационными продуктами и технологиями;
- представления, накопления, хранения и активного использования исходных и расчетных данных, а также данных телеизмерений;
- обмена информацией между отдельными задачами;
- графического представления расчетных схем на плане городской застройки;
- графического представления результатов расчета (в виде пьезометрических и температурных графиков и др.);
- возможности отображения результатов расчета на схеме сети.

Вместе с тем принципы реализации и архитектура подобных систем в значительной степени зависят от профессиональной специфики их предполагаемого использования: при проектировании ТСС, например, в рамках САПР; при управлении состоянием ТСС - при решении задач АСДУ; при эксплуатации - в составе тех или иных АРМов (автоматизированных рабочих мест) инженера технолога; в научных исследованиях.

При проектировании целями моделирования являются задачи анализа допустимости режимов, работоспособности, управляемости и надежности функционирования ТСС с учетом: неопределенности информации; иерархичности моделирования (из-за уточнения модели объекта при сниже-

нии многообразия условий); многовариантности проводимых расчетов, обуславливающих повышенные требования к сервису, уровню автоматизации информационных обменов с БД; обобщения результатов расчетов и формирования технической документации.

В рамках АСДУ подобные системы могут использоваться при решении задач планирования режимов, моделирования и анализа последствий принимаемых решений по управлению, а также возможного выхода из строя отдельных элементов или оборудования ТСС. Специфическими требованиями здесь являются: обеспечение решения задач в реальном масштабе времени; учет многоуровневых информационных обменов вследствие распределенности технических систем сбора, передачи и хранения первичной информации, высокая надежность и робастность алгоритмов по отношению к возможному "пропаданию" или зашумлению отдельных данных и т.д. Вместе с тем эксплуатация программного обеспечения применительно к конкретному объекту позволяет в значительной степени формализовать многочисленные типичные ситуации, что не исключает необходимости их обобщения, учета опыта пользователя, обучения программ в процессе их работы, накопления базы знаний.

При проведении научных исследований такие вычислительные системы могут быть эффективно использованы в целях подготовки заключений, экспертиз проектов, выработки рекомендаций по наладке и рациональной реконструкции и эксплуатации существующих ТСС. Однако главным их назначением здесь является генерация качественно нового методического и программного обеспечения. При этом основными требованиями являются: наличие и доступность информационного обеспечения в виде соответствующих баз данных и баз программных модулей; наличие дружественной интерфейсной оболочки, обеспечивающей взаимодействие пользователя с ИВС, с ППП и между отдельными ППП.

КРАТКАЯ ХАРАКТЕРИСТИКА АРМ ТЕХНОЛОГА ТЕПЛОВЫХ СЕТЕЙ

Одна из последних разработок ИСЭМ СО РАН - программно-вычислительный комплекс (ПВК) "АРМ ТТС" (Автоматизированное Рабочее Место Технолога Тепловых Сетей), ориентированный на персональные компьютеры типа IBM PC-386 и выше. ПВК максимально приближен к потребностям практики в автоматизации гидравлических, тепловых и наладочных расчетов ТСС и выполнен в комплексе с интегрированной графической системой (ИГС) "АРМТЕСТ", разработанной Военным инженерным техническим университетом (ВИТУ, г. Санкт-Петербург).

ПВК "АРМ ТТС" предназначен для решения важных для практики проектирования, эксплуатации и диспетчерского управления задач и может быть использован в научно-исследовательских, учебных, проектных и эксплуатационных организациях, занимающихся вопросами теплоснабжения.

Перечень задач, решаемых комплексом:

- 1) расчет установившегося гидравлического режима (потокораспределения);
- 2) построение пьезометрических графиков;

- 3) расчет установившегося "температурного поля" по узлам сети;
- 4) расчет температур воздуха в отапливаемых помещениях, а также нагрузок и количества тепла, полученного потребителями по всем видам присоединенной нагрузки: отопления, вентиляции, горячего водоснабжения в текущем режиме;
- 5) определение суммарных потерь тепла по участкам за расчетный период;
- 6) построение температурных графиков;
- 7) определение параметров дросселирующих устройств для наладки тепловой сети;
- 8) проверка нарушений параметров режимов ТСС.



Рисунок 1 - Схема взаимодействия модулей ПКВ "АРМ ТТС"

ПКВ "АРМ ТТС" можно разбить (рис.1) на четыре части: 1) ПО для формирования схемы тепловой сети на плане города, занесения исходной информации в базу данных и просмотра результатов расчета; 2) система управления базой данных (СУБД); 3) ПО для связи базы данных с расчетными модулями, фильтрации ошибок исходных данных, анализа разрешимости, управления расчетом и выводом результатов в базу данных; 4) ПО для проведения технологических расчетов.

Расчетные модули ПКВ обладают:

- возможностью решения задач для сетей произвольной конфигурации (разветвленных, многоконтурных) и структуры (с произвольным числом и размещением насосных станций, источников, потребителей, регуляторов давления или расхода), с присоединением потребителей теплоты,

имеющих разнородные нагрузки;

- высоким быстродействием (8-15сек. для многоконтурных сетей, содержащих порядка 1500-2000 ветвей расчетной схемы при расчете на Pentium/166) и теоретически гарантированной сходимостью вычислительного процесса, обеспечивающей получение решения с наперед заданной точностью;

- возможностью решения задач большой размерности (в данной версии до 3000 участков расчетной схемы одного уровня в двухлинейном изображении и практически неограниченной размерности при иерархическом представлении схемы);

- повышенной надежностью за счет детальной автоматизированной экспертизы корректности задания исходных данных;

- возможностью графической визуализации исходной информации (характеристик насосных станций и источников) и интерпретации результатов расчетов (автоматизированное построение, отображение, спецификация и вывод на печать пьезометрических и температурных графиков);

- графическим представлением расчетной схемы на фоне городской застройки (с возможностью изменения режимов работы различных элементов системы).

ПКВ "АРМ ТТС" работает на IBM PC совместимых компьютерах с процессором не ниже 386. Требуемый объем ОЗУ - 560 Кб основной и 500 Кб верхней памяти. Для инсталляции ПКВ требуется 5.8 Мб пространства на HD (2.3 Мб занимает ИГС "АРМТЕСТ" и 3.5 Мб - расчетные модули).

ПКВ позволяет решать следующие практические задачи:

- количественно обосновать решения по организации режимов при эксплуатации и диспетчерском управлении;

- осуществлять теплогидравлическую экспертизу проекта при реконструкции и расширении системы, а также выявлять слабые места системы на стадиях наладки и эксплуатации режимов;

- правильно планировать последовательность наладочных мероприятий;

- обосновать ведение диспетчером графика центрального регулирования отпуска тепла;

- выявить факты нарушения обеспечения потребителей требуемым количеством тепла, а также технологических ограничений в расчетном режиме;

- определить допустимое время снижения подачи тепла в аварийных режимах из условия допустимого снижения температуры воздуха в отапливаемых помещениях;

- определить возможные последствия для потребителей при вынужденном отклонении параметров теплоносителя от требуемых и оценить мероприятия по нормализации режимов;

- оценить последствия от переключений в сети, изменения напоров насосов и подключения дополнительных потребителей и т.д.

Кроме того, применение ПКВ "АРМ ТТС" на эксплуатационных предприятиях тепловых сетей также позволяет реализовывать следующие информационные функции:

- создание графических, иерархически связанных баз данных по системам теплоснабжения;

- получать быстрый доступ к схемно-параметрической и

топологической информации, а также бумажную копию любого фрагмента тепловой сети с соответствующими данными;

- вести объектно-ориентированный дневник событий при диспетчерском управлении;
- создавать выборки данных о режимах работы ТСС;
- организовывать информационную поддержку при проведении испытаний тепловых сетей;

ПВК "АРМТСС" внедрен на предприятиях тепловых сетей АОЭиЭ "Иркутскэнерго" для городов Ангарска, Братска и Иркутска.

ПВК "АРМ ТТС" кроме расчетных модулей включает в себя диалоговую графическую оболочку; программу выбора состава и структуры базы данных по уровням сетей для возможности осуществления одно- и много-уровневых расчетов; программу ведения архива, а также сервисные утилиты.

ПРИНЦИПЫ ОРГАНИЗАЦИИ БД

Создание расчетной схемы на фоне городской застройки и формирование базы данных по элементам теплоснабжающей системы осуществляется с помощью ИГС "АРМТЕСТ", которая состоит из редакторов плана и сети, а также самой базы данных и позволяет: отображать расчетную схему на плане города в нужном масштабе; отображать результаты расчета на схеме сети; создавать принципиальные схемы объектов системы теплоснабжения (источников теплоты, насосных станций, тепловых пунктов и др.); создавать гибкие структуры базы данных в соответствии с требованиями конкретных расчетных задач;

Принципы организации структуры БД в значительной степени зависят от размера рассчитываемой схемы теплоснабжения, целей расчета, а также специфики и набора решаемых задач.

Для решения задач большой размерности, увеличения быстродействия и привязки задач к ИГС "АРМТЕСТ", в ИСЭМ СО РАН разработана иерархическая структура базы данных. Разбиение тепловых сетей (ТС) на уровни обусловлено большим количеством информации и разной степенью детализации объектов ТСС крупных городов и промышленных центров. Многоуровневая структура позволяет отдельно рассчитывать магистральные и распределительные сети, а также осуществлять совместные расчеты системы в целом во взаимной увязке сетей разного уровня. Структурно тепловые сети подразделяются на магистральные и распределительные по следующему принципу. К магистральным тепловым сетям относятся ТС от источника теплоснабжения до вводов в микрорайоны или квартала. Кроме того, к ним относятся насосные станции и участки распределительной тепловой сети, если они совместно с магистральными сетями образуют кольца в однолинейном изображении. К распределительным тепловым сетям относятся тупиковые ответвления от ввода в микрорайон или квартал до потребителей теплоты (домов или промышленных предприятий). Распределительная сеть одного микрорайона может состоять из нескольких несвязанных между собой фрагментов, число которых равно количеству вводов в микрорайон. Магистральные и распределительные тепловые сети связаны между собой

через общий элемент сети - обобщенный потребитель (на уровне магистральных сетей), являющийся обобщенным источником на уровне распределительных сетей, через который осуществляется пересылка данных с одного уровня на другой.

Для решения задач оперативного управления режимами ТСС, представляется целесообразным применение эквивалентирования и агрегирования теплоснабжающих систем, поскольку производить расчеты режимов крупной ТСС с детальным анализом режимов всех ее составляющих (источники тепла, протяженные тепловые сети, потребители с различным сочетанием и видом присоединения разнородной нагрузки) в реальном времени практически невозможно. Эквивалентная модель должна обеспечить: 1) возможность применения для оперативных расчетов существующих программных разработок по расчету и анализу установившихся тепло-гидравлических режимов; 2) значительное сокращение времени расчета, позволяющее в приемлемые сроки исследовать режимы систем большой размерности.

Укрупнение реальной схемы теплоснабжения осуществляется по следующим принципам. Все узлы разветвления схемы можно рассматривать как обобщенные потребители, предшествующие по потоку определенным группам реальных потребителей. Это позволяет свертывать тупиковые ответвления внутри-квартальных тепловых сетей с присоединенными к ним однотипными потребителями в один узел с обобщенной нагрузкой. Например, если все потребители ответвления являются жилами зданиями типовой застройки с одинаковым составом и соотношением нагрузок

В магистральных сетях возможно объединение нескольких участков с одинаковыми конструктивными характеристиками (диаметры, способ прокладки, тип и толщина изоляции и т.д.) в один участок, если на них нет отборов теплоносителя.

Надо отметить, что при укрупнении схемы из нее не должны выпадать узлы, в которых установлены регулирующие органы, перекачивающие или подкачивающие насосные станции, узлы рассечки, или любые другие элементы тепловой сети, предназначенные для регулирования.

Второй принцип, позволяющий существенно упростить расчетную схему - это эквивалентирование, как обобщенных, так и реальных потребителей, участками с заданными сопротивлениями и перепадами температур между подающим и обратным трубопроводами. Естественно, что такой подход приемлем в случае наличия достоверной информации о градиенте давлений и температур, которую можно получить с помощью предварительных многовариантных расчетов различных режимов по модели, подробно учитывающей специфику всех потребителей [6]. Осуществив все необходимые расчеты, можно получить значения температурного перепада для всех потребителей. Учитывая большую инерционность зданий по температурному тракту, аккумулирующая способность которых составляет от двух до пяти суток (в зависимости от типа постройки - материала и толщины стен, степени остекления и т.д.), такой подход для оперативных расчетов представляется вполне приемлемым.

Для категорий потребителей, у которых необходимо

контролировать температуру воздуха внутри помещений, применяется расчет теплогидравлических режимов потребителей через относительные характеристики (по отношению к расчетному режиму) [7], что позволяет существенно снизить объем исходной информации, необходимой при расчете ТСС в целом и в несколько раз сокращает время расчета потребителя.

Изложенные принципы эквивалентирования и декомпозиции таким образом диктуют соответствующие принципы организации БД (разбиение на уровни тепловых сетей, набор объектов, из которых формируется расчетная схема и т.п.).

ВЫВОДЫ

1. Сформулированы новые требования и направления развития прикладного ПО для компьютерного моделирования ТСС при их проектировании, эксплуатации и диспетчерском управлении.

2. Дана характеристика реализации нового поколения ПО, обеспечивающего решение основных режимно-технологических задач для систем теплоснабжения.

УДК 681.32:007.58

ИМИТАЦИОННОЕ МОДЕЛИРОВАНИЕ НЕЙРОСЕТЕЙ ДЛЯ РЕШЕНИЯ УРАВНЕНИЙ МАТЕМАТИЧЕСКОЙ ФИЗИКИ ЛОКАЛЬНО-АСИНХРОННЫМИ МЕТОДАМИ

М.А.Новотарский

В статье дано краткое изложение краевой задачи математической физики и изложены особенности локально-асинхронного метода для ее решения, отражены основополагающие принципы реализации данного метода на основе нейросети. Обсуждается структура нейросети и алгоритм ее функционирования, даны основные подходы к формализации исходной задачи с помощью PRO-сетей, формулируются принципы построения нейросетевого имитатора.

В статті дається короткий виклад крайової задачі математичної фізики і викладені особливості локально-асинхронного методу для її вирішення, відображені основні принципи реалізації даного методу на нейромережі. Обговорюється структура нейромережі і алгоритм її функціонування, дані основні підходи до формалізації вихідної задачі за допомогою PRO-мереж, формулюються принципи побудови нейромережевого імітатора.

In clause the brief statement of a boundary value problem of mathematical physics is given and the features of a local-asynchronous method for its solving are stated, the basic principles of realization of the given method on a neural network reflected. The neuronet structure and algorithm for it's functioning is discussed, the basic approaches to formalization of an initial task with the help of PRO-nets are given, and the principles of construction of the neurosimulator are formulated.

ПЕРЕЧЕНЬ ССЫЛОК

1. Методы и алгоритмы расчета тепловых сетей/ В.Я. Хасилев, А.П. Меренков, Б.М. Каганович М.: Энергия, 1978.
2. Меренков А. П., Хасилев В. Я., Теория гидравлических цепей, М.: Наука, 1985.
3. Сеннова Е.В., Сидлер В.Г. Математическое моделирование и оптимизация развивающихся теплоснабжающих систем. - Новосибирск: Наука, СО, 1987.- 221с
4. Математическое моделирование и оптимизация систем тепло-, водо-, нефте-, газоснабжения. / А.П.Меренков, Е.В.Сеннова, С.В.Сумароков, и др.- Новосибирск: ВО "Наука", -Сиб. Изд. фирма, 1992. - 407 с.
5. Методы управления физико-техническими системами энергетики в новых условиях / Н.И.Воропай, Н.Н.Новицкий, Е.В.Сеннова и др. - Новосибирск: Наука, Сибирская издательская фирма РАН, 1995.-334с.
6. Сидлер В.Г., Шалагинова З.И., Математическая модель теплогидравлических режимов абонентских вводов // Методы анализа и оптимального синтеза трубопроводных систем. - Иркутск: СЭИ СО РАН, 1991.- С.111-124.
7. Шалагинова З.И. Разработка и применение методов теплогидравлического расчета для задач эксплуатации теплоснабжающих систем // Международная научно-практическая конференция "Человек. Среда. Вселенная, том 1, Иркутск, 1997.- 198с.

ВВЕДЕНИЕ

Прошло десятилетие с тех пор, как мир охватила "нейронная лихорадка". Число компьютеров и программных продуктов, использующих нейросетевую технологию, растет с каждым днем. Объем мирового рынка нейронных сетей ежегодно увеличивается на 40% и составляет около 1 млрд. долларов. К настоящему времени уже сформировался круг проблем, традиционно решаемых с применением данной технологии. Наиболее широко используется она для решения всевозможных задач прогнозирования и распознавания, преимущественно в финансовой сфере и медицине. Однако такая специализация нейросетей ни в коей мере не может рассматриваться как окончательная.

Это универсальная технология, позволяющая решать задачи в самых неожиданных областях. Строгим обоснованием такого подхода следует считать знаменитую теорему Колмогорова, перефразированную в терминах нейросетей.

1. ТЕОРЕМА КОЛМОГОРОВА

Если известны функции h_{ij} , то любую непрерывную функцию от n переменных можно точно реализовать с помощью простой нейросети на основе трехслойного перцептрона. Для этого достаточно подобрать $2n + 1$ передаточных функций g_j нейронов скрытого слоя.

$$F(x_1, x_2, \dots, x_n) = \sum_{j=1}^{2n+1} g_j \left(\sum_{i=1}^n h_{ij}(x_i) \right), \quad (1)$$

где g_j и h_{ij} - непрерывные функции, причем h_{ij} не зависят от функции F .

К сожалению, практическое использование теоремы Колмогорова сильно затруднено тем, что функции h_{ij} - негладкие и трудно вычисляемые, а для выбора функций g_j не существует сколь-нибудь эффективных методик. Поэтому для каждого нового класса задач, решение которых предполагается получать с помощью нейросети, разрабатываются свои концепции построения ее структуры и алгоритма функционирования. Решение дифференциальных уравнений - это сравнительно молодая, но бурно развивающаяся область применения нейросетей. Известны примеры успешных разработок в этой области, как в ближнем, так и в дальнем зарубежье [1,2]. Специфика исследований, проводимых в Институте математики НАН Украины, заключается в том, что нейросети ориентированы на решения уравнений математической физики локально-асинхронными многосеточными методами. Как раз особенности данных методов и являются побудительным мотивом к использованию нейросетей для их реализации.

2. ЛОКАЛЬНО-АСИНХРОННЫЕ МЕТОДЫ

Локально-асинхронные многосеточные методы для решения краевых задач - это методы, которые предполагают получение решения глобальной задачи путем разбиения ее на множество локальных задач, решаемых параллельно [3]. Решение каждой из таких задач понимается как самостоятельный локальный процесс, который может взаимодействовать с другими подобными локальными процессами с целью приема данных и передачи результатов вычислений. Разбиение локальной задачи может осуществляться с учетом пространственно-временных аспектов решения общей задачи. Самый простой подход предполагает отождествление нейронов нейросети с узлами сеточной области. В этом случае форма и размерность нейросети соответствует конфигурации границ краевой задачи.

Исходная краевая задача задается в общем виде операторным уравнением

$$L_U = F, \quad (2)$$

Заменяя, оператор L разностным оператором L_h , правую часть f - некоторой сеточной функцией f_h , а точное

решение u разностным решением u_h , получим разностную краевую задачу

$$L_h u_h = f_h, \quad (3)$$

Для решения данной разностной краевой задачи на основе нейросети предложен локально-асинхронный многосеточный метод [3]. Его отличительной чертой есть то, что при реализации итерационного процесса $u_h(n+1) = \text{Iter}(u_h(n), L_h, f_h, v_h)$ вычисление v_h осуществляется исходя из точного решения уравнения $L_h v_h = d_h$ и формирования невязки d_h для каждого узла сеточной области путем операций проецирования и интерполирования невязок, взятых с узлов с различным шагом сетки. Подлинная асинхронность этого метода состоит в том, что для него доказана сходимость итерационного процесса к решению при использовании хаотической последовательности формирования координат соседних узлов для переменного точечного шаблона. Изменение шага сетки, а, следовательно, и координат точечного шаблона, принято выбирать кратным двойке. Отсюда следует выбор структуры нейросети, реализующей обработку значений функции u_h и невязки d_h , полученных на предыдущей сетке. Количество входов нейрона и размерность сети будет определяться размерностью исходной задачи. На рис.1 представлена простейшая нейросеть, позволяющая рассмотреть одномерный случай.

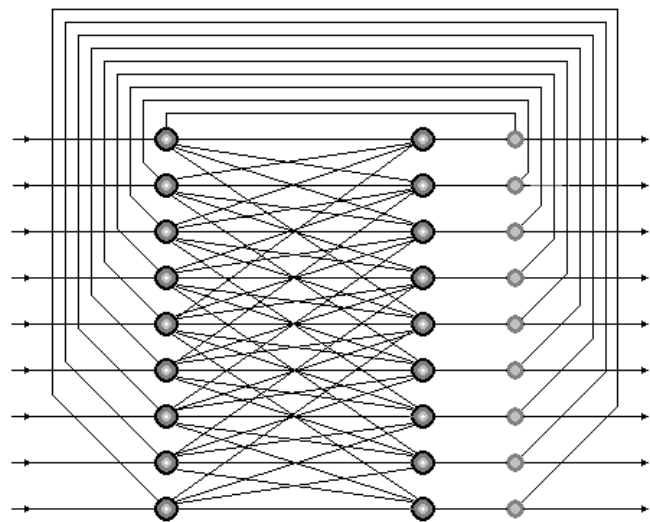


Рисунок 1 - Нейросеть для решения уравнений математической физики локально-асинхронным многосеточным методом

Основной особенностью такой сети можно считать ее рекуррентность, что диктуется необходимостью реализации итерационного процесса. Известно, что использование рекуррентных нейросетей сопряжено с рядом проблем, одна из которых заключается в трудности определения момента останова. В данном случае эти трудности реша-

ются путем введения послойной специализации. Сеть содержит три слоя нейронов, каждый из которых ориентирован на реализацию специфических функций. Первый слой предназначен для выбора текущего канала приема информации и предварительной ее подготовки к участию в итерационном процессе. Дендриты нейронов второго (скрытого) слоя управляются синапсами, значения весов которых могут принимать значения 0 или 1.

Формирующийся таким образом вектор весов синапсов для каждого нейрона второго слоя (рис.2) задает требуемый согласно алгоритма точечный шаблон окружения.

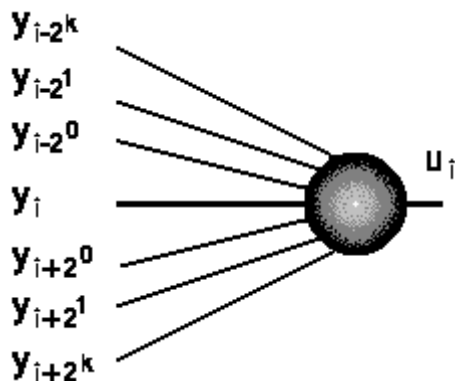


Рисунок - 2 Нейрон второго слоя

$$u_i(n) = \Psi_i(Y),$$

где

$$\Psi = \{\Psi_i\}, i = (0, 2^{k+1}),$$

$$Y = \sum(w_{i-2^k}y_{i-2^k} + \dots + w_i y_i + w_{i+2^k}y_{i+2^k}),$$

$$w_i = (0, 1), y_i = \{d_p, u_i(n-1)\}.$$

Функция Ψ_i для данного нейрона представляет собой итерационную формулу разностной схемы, определяемую способом перехода от (2) к (3).

Третий слой нейронов - это анализаторы завершения итерационного процесса, в функции которых входит активация синапсов первого слоя до момента получения сходимости. При получении признака сходимости для i -й точки нейрон третьего слоя блокирует соответствующий синапс первого слоя, что приводит к блокировке рекуррентной связи и прекращению итерационного процесса.

3. PRO-СЕТЕВАЯ МОДЕЛЬ

Для исследования данной нейросети и методов решения на ней уравнений математической физики создан программный имитатор. В основу его положена PRO-сетевая имитационная модель. Такая модель есть формальное описание структуры и процессов, протекающих в нейро-

сети в ходе ее функционирования. Инструментом формализации являются PRO-сети [4], которые фиксируют уровень исследования нейросети путем задания ее имитационной модели в виде набора абстракций

$$\Phi = (S, T, G, V). \quad (4)$$

При этом множество позиций S отражает степень детализации компонент модели и, тем самым, задает статическое состояние нейросети. Множество переходов T отражает динамику функционирования модели, которая реализуется посредством алгоритмов переработки информации, представленной элементами множества S . Структура путей перемещения информации в модели фиксируется множеством G , а результаты исследований и текущее состояние модели заносятся в глобальные переменные, составляющие множество V .

Множество позиций

$$S = \{(Q_0, M_0, B_0), (Q_1, M_1, B_1), \dots, (Q_i, M_i, B_i), \dots, (Q_N, M_N, B_N)\}$$

содержит набор элементов, каждый из которых есть триада: Q_i - номер позиции S_i , $M_i = \{m_1, m_2, \dots, m_k\}$, m_k - k -й атрибут метки, $B_i = \{b_1, b_2, \dots, b_k\}$, b_k - k -й буфер данных.

Множество переходов содержит набор элементов t_i , каждый из которых является сложной структурой данных:

$$T = \{t_0, t_2, \dots, t_p, \dots, t_L\},$$

где L - количество t -элементов в PRO-сети; t_i - переход или t - элемент, $t_i \in T$, $t = (A, C, \rho, \pi, \tau, d_{\min}, d_{\max})$, A, C - конечные непустые множества входных и выходных ребер; ρ - управляющая процедура; π - вычислительная процедура; τ - время срабатывания перехода; d_{\min}, d_{\max} - соответственно минимальное и максимальное время задержки срабатывания перехода.

Множество ребер сетевого двудольного графа определяет связи между позициями и переходами: $G: S \times T \cup T \times S$. Структура связей между переходами и позициями может быть задана произвольным образом.

Наиболее простой способ - это матрица инцидентности $I: S \times T \Rightarrow Z$, где $Z = \{-1, 1, 0\}$,

$$I(s, t) = \begin{cases} -1, & \text{если } (s, t) \in G; \\ 1, & \text{если } (s, t) \in G^{-1}; \\ 0, & \text{остальные случаи.} \end{cases}$$

Ввиду громоздкости и большой разреженности матрица инцидентности не является эффективным способом задания структуры связи. Более эффективным подходом могут служить различные динамические списковые и/

или иерархические структуры.

Множество глобальных переменных $V = \{v_1, v_2, v_3\}$ включает три вида подмножеств, отличающихся по функциональному назначению. Подмножество v_1 содержит глобальные переменные, описывающие временные характеристики имитационной модели (состояние глобального счетчика времени, время реакции различных ресурсов и др.). Подмножество v_2 - это служебные переменные, отражающие текущее состояние компонент модели. Подмножество v_3 включает глобальные переменные статистического характера.

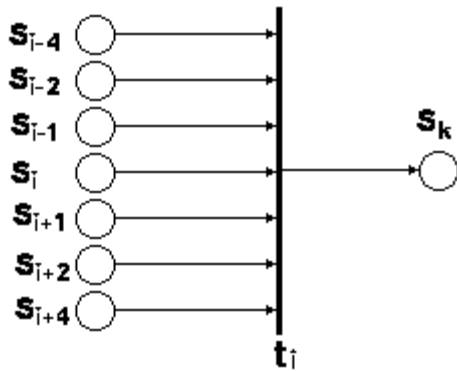


Рисунок 3 - PRO-сетевая модель нейрона

PRO-сетевая модель имеет также свое графическое представление. Элементы множества S изображаются в виде кружков, связанных ребрами, входящими во множество G , с переходами, представленными в виде вертикальных отрезков и входящими во множество T . PRO-сетевой эквивалент нейрона скрытого слоя исследуемой нейросети изображен на рис.3. Активизация процедуры π перехода осуществляется посредством процедуры ρ , основной задачей которой является анализ состояния меток на входных и выходных позициях. В данной модели наличие метки на входной позиции сети интерпретируется как включенное состояние синапса нейрона. При наличии метки на входной позиции процедура ρ имитирует случайно распределенную в диапазоне d_{\min}, d_{\max} аппаратную задержку и запускает процедуру π , параметрами которой выступают элементы множества B . По результатам вычисления формируется выходной вектор B ,

который в сопровождении соответствующей метки устанавливается на выходную позицию. Затем глобальный счетчик системного времени продвигается на величину τ .

По PRO-сетевому формальному описанию в среде визуального программирования Delphi построен программный имитатор PROSimulator, позволяющий в диалоговом режиме задавать большое количество параметров модели. Основная задача его калибровки сводится к обучению моделируемой нейросети по методу обучения без учителя. Как уже упоминалось, локально-асинхронный метод гарантирует сходимость к решению при использовании хаотического выбора координат точечного шаблона. Однако скорость такой сходимости далеко не всегда является приемлемой. Поэтому обучение сети заключается в запоминании каждым нейроном наиболее эффективной последовательности активации синапсов для каждого класса краевых задач. Практическая работа показала высокую скорость обучаемости сети при использовании специальных методик обучения на типовых задачах с небольшим количеством точек дискретизации и перенесением этих результатов на задачи с большим объемом вычислений.

ЗАКЛЮЧЕНИЕ

В заключение следует отметить, что использование несколько нетрадиционного понимания работы нейрона дало возможность объединить идеи параллельной обработки с преимуществами нейросетевых структур. Такой подход позволил расширить круг задач, для решения которых используются нейросети и, по мнению автора, немного сблизить позиции представителей точных наук и биологов, которые в ответ на предложенный им перцептрон Розенблата постоянно напоминают о том, что не мозг, а, скорее всего каждый нейрон можно сравнивать с вычислительной машиной.

ПЕРЕЧЕНЬ ССЫЛОК

1. Федорова Н.Н., Терехов С.А. Нейросетевой аппроксиматор для численного решения дифференциальных уравнений. Тезисы III рабочего семинара-совещания "Нейронные сети в информационных технологиях", Снежинск, 1-3 апреля, 1998. Интернет: <http://www.vniitf.ru/~nimfa/conf/nnit98/abstracts.html>.
2. Lagaris I.E., Likas A., Fotiadis D.I. Artificial Neural Networks for solving Ordinary and Partial Differential Equations. Submitted to IEEE TNN, 1997 Интернет: <http://xyz.lanl.gov/physics/9705023>
3. Марчук В.А., Нестеренко Б.Б. Асинхронные методы параллельных вычислений // Труды Института математика НАН Украины.-1999.-22.-308с.
4. Нестеренко Б.Б., Новотарский М.А. Мультипроцессорные системы. Институт математики.-Киев:1995.-408с.

ИМИТАЦИОННЫЙ ПРОЕКТ АВТОМОБИЛЬНОГО ДОРОЖНОГО ДВИЖЕНИЯ

В.Н.Томашевский, Д.С.Печенежский

Предложена концепция создания имитационного проекта для управления движением автотранспортных средств на основе визуального моделирования. Описаны архитектура и принципы функционирования основных компонентов, и способы их реализации. Имитатор модели движения транспортных средств использует сетевую распределенную обработку и технологию Java.

Запропоновано концепцію створення імітаційного проекту для управління рухом автотранспортних засобів на основі візуального моделювання. Описано архітектуру і принципи функціонування основних компонентів, і засоби їхньої реалізації. Імітатор моделі руху транспортних засобів використовує розподілену обробку у мережі і технологію Java.

The concept of creation of the imitative project for traffic control of vehicles is offered on the basis of visual simulation. Are described the architecture and principles of operation of the basic components, and methods of their realization. The simulator of model of movement of means of transport uses the network distributed processing and technology Java.

ВВЕДЕНИЕ

Интеграция Украины в европейские структуры ставит новые задачи по улучшению дорожного движения. С этой целью разработана "Государственная программа развития дорожного движения на автомобильных дорогах, улицах...". Одна из задач этой программы заключается в разработке имитационных моделей дорожного движения для анализа пропускной способности и мест концентрации дорожно-транспортных происшествий. Пропускная способность зависит от множества факторов: конфигурации дорог, их полосности, количества и типов перекрестков, переходов и др. Основной причиной снижения пропускной способности дорог являются так называемые "узкие места", поиск которых и является главной целью исследований. Применение аналитических моделей для решения такой задачи затруднено из-за ее сложности.

Ранее созданные программы подобного назначения имитировали дорожное движение достаточно условно. Основное их назначение было нахождение приблизительных оценок пропускной способности и коэффициентов загрузки участков городской сети при разных вариантах планировочных решений.

Основным недостатком этих моделей является их неточность. Участки дорог в них рассматриваются как многоканальные приборы, а перекрестки не моделировались вообще, т.е. предполагалось, что любой автомобиль может свободно пройти перекресток в любой момент времени вне зависимости от ситуации на нем. Данное допущение уместно при грубых оценках и при небольшом коэффициенте загрузки участков дорог прилегающих к перекрестку. Однако при увеличении коэффициента загрузки дорог

автомобили начинают оказывать все большее влияние на движение друг друга. Также усиливается влияние конфигурации перекрестка, режима его регулирования и т.д. Аналитически просчитать влияние конфигурации на его пропускную способность очень сложно, так как увеличение коэффициента загрузки дорог ведет к снижению пропускной способности перекрестка, что в свою очередь приводит к еще большему увеличению коэффициентов загрузки дорог и т.д.

Еще одним существенным недостатком уже существующих моделей является их неразвитый интерактивный интерфейс, что делает трудоемкой работу с ними.

Исходя из вышесказанного, можно сделать вывод, что возникла необходимость в создании новых имитационных систем для моделирования движения автомобильного транспорта. Подобные системы имитации дорожного движения разработаны в странах Западной Европы и Америки, такие как MITSIM, FRESIM, PHAROS (США), AUTOBAHN, PLANSIM-T (Германия), DRACULA, PADSIM (Великобритания), ANATOLL, SIMDAC (Франция) и др. Они используются, в первую очередь, для улучшения условий движения городского и общественного транспорта, а также для оценки его экономической эффективности. Европейским сообществом разработан перечень требований для систем подобного рода. Однако в Украине пока нет подобных систем.

ЦЕЛЬ И КОНЦЕПЦИЯ

Целью создания имитационного проекта автомобильного дорожного движения является моделирование сети дорожного движения, определение пропускной способности участков дорог, а также нахождение оптимальных маршрутов для специальных транспортных средств, таких как скорая помощь, милиция, пожарная служба.

Исходя из принципов построения визуальных объектных систем, необходимо провести статическую декомпозицию объектов для модели дорожного движения. Основным компонентом верхнего уровня визуального размещения имитационной модели является карта основных автомагистралей Украины, которая в свою очередь декомпозируется на множество других компонентов, каждый из которых представляет собой карту города либо карту загородных дорог. Каждый компонент городского размещения может быть расчленен на компоненты, представляющие всевозможные фрагменты городской сети дорог (районов). Те в свою очередь декомпозированы на различные участки дорог: пересечения, линейные участки, участки с круговым движением, наклоны, подъемы, участки с поворотами, участки с туннелями, участки с ограни-

чением по высоте (проезды под мостами), участки с остановками общественного транспорта, участки с ограничениями по грузоподъемности (мосты, эстакады).

Отличием разрабатываемой модели от других ей подобных аналогов, перечисленных, выше, является то, что имитационная модель создается автоматически для того участка дорожного движения, который отображается на экране дисплея. Такой подход не требует построения полной исчерпывающей имитационной модели и ее хранения в памяти компьютера, т.е. используется принцип "что отображаю, то и моделирую". В качестве наименьшего отображаемого элемента моделирования был взят *домен* [1], который представляет собой однополосный регулируемый или нерегулируемый участок дороги. Предполагается, что ширина полосы достаточная для размещения автомобиля любой ширины, но не более одного автомобиля по ширине. Участки дорог разделяются на домены так, чтобы светофоры, знаки, переходы, разветвления, консоли дорог были на концах доменов. Таким образом, *домен* - это элементарный однополосный участок дороги, на протяжении которого автомобили не встречают препятствий кроме других автомобилей. В конце такого участка по ходу движения автомобиля могут находиться препятствия: светофор, знак, переход, пересекающийся участок дороги, разветвление участков дорог. Очевидно, что любой участок дороги может быть разбит на домены, причем на рассматриваемом участке реальной дороги могут находиться одновременно несколько доменов, как показано на рис. 1.

В зависимости от участков дорог, на которые накладываются домены, динамический объект домен характеризуется различной длиной и количеством автомобилей на нем.

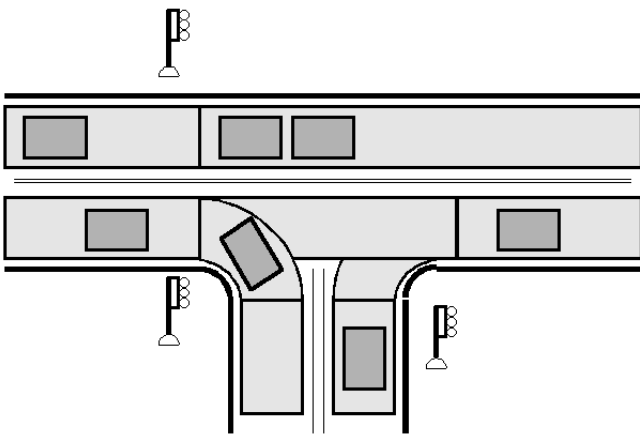


Рисунок 1 - Наложение доменов на участки дорог и расположение автомобилей на доменах

Домены включают в себя динамические объекты *автомобили*. Объект *автомобиль* служит для моделирования движения автотранспортного средства. Он создается на участке дороги для обозначения конкретного транспортного средства с его характеристиками. Транспортное средство может принадлежать к одному из следующих типов:

мотоцикл, легковой или грузовой автомобиль, грузопассажирский транспорт, специальный транспорт (милиция, пожарная служба, скорая помощь) и т.д.

Динамический объект *автомобиль* в зависимости от транспортного средства, которое он отображает, имеет различные характеристики, включающие: отрицательное минимальное ускорение торможения, положительное максимальное ускорение разгона, длину, высоту, ширину и полную массу автомобиля (масса снаряженного транспортного средства с грузом и пассажирами), приоритет, который может устанавливаться для специального транспорта и общественного транспорта.

Объект *автомобиль* может входить в объект домен, передвигаться в нем, останавливаться, становиться в очередь и покидать его. Попадая в какой-либо определенный домен, "поведение" автомобиля определяют параметры домена (максимальная разрешенная скорость, знаки и т.п.).

В модели должны отражаться особые ситуации, которые могут быть активизированы в любое время. Под особыми ситуациями подразумевается затор на трассе, авария, эффект поперечного сужения и т.п. Инцидент может быть сгенерирован стохастическим либо детерминированным способом. Детерминированные инциденты должны определяться пользователем. Стохастические инциденты создаются для каждого участка дороги согласно заданному вероятностному распределению.

ПРЕДСТАВЛЕНИЕ СТАТИЧЕСКИХ ОБЪЕКТОВ

Для имитации дорожного движения необходимо сетевое представление участков дорог (граф), организация управления потоком автомобилей, воспроизведение особых ситуаций, маршрутизация транспортных средств и перемещение транспортных средств.

Детализация имитационной модели должна меняться в зависимости от уровней визуального отображения, определяемых пользователем, что дает возможность моделировать как отдельные транспортные средства, так и потоки автомобилей.

Для имитатора дороги представляются связями, узлами сети и доменами. Данные, которые описывают имитируемую сеть, читаются из сетевой базы данных, которая может быть легко создана, используя интерактивный графический редактор сети дорог, разработанный для имитатора. Сетевая база данных включает описание всех сетевых объектов, соединение участков дорог, приоритет использования трассы, регулирование разворотов и перемещений на пересечениях (например, запрет левых поворотов и разворотов) и управляющих устройств (светофоров).

Узел может быть или пересечением отдельных дорог, или источником, или стоком, где потоки движения вводятся или покидают имитируемую сеть. Каждый узел представляется следующими элементами данных:

- тип узла, который определяет, является ли узел пересечением или источником и/или стоком;
- код узла: это - уникальный опознавательный номер для узла;
- имя узла - может использоваться произвольное имя узла, чтобы маркировать узел в графическом окне и отчете моделирования.

Связи - протяженные участки дорог без пересечений, которые соединяют узлы. Каждая связь может состоять из одного или большего количества участков дороги и характеризуется:

- типом связи, которая может быть автострадой, уклоном, городским проездом, туннелем;
- кодом связи: код связи - опознавательный номер, назначенный для каждой связи;
- стартовыми и конечными узлами для соединения связей;
- число доменов на участке;
- число полос на участке дороги (количество параллельных полос одного направления движения).

Входные и выходные связи узлов соединяются, если существуют, по крайней мере, одно соединение пути между этими двумя связями. Ограничения перестроений из одной связи в другую, представляется в таблице. Эта таблица содержит список ограничений на выполнение перестроений, для каждой определенной пары.

Имитатор представляет сеть на уровне доменов, накладывающихся на полосы участков дорог. Каждый домен описан двумя элементами данных: 1) кодом домена, т.е. уникальным опознавательным кодом; и 2) маршрутом, который изменяет направление и код приоритета использования трассы, определяемый по некоторым правилам.

Правилом изменения полосы:

- разрешенным поворотом налево;
- разрешенным поворотом направо;
- разрешенным поворотом и налево и направо.

Правилом приоритета использования полосы:

1. Запрещением движения грузовикам;
2. Ограничением проезда по высоте;
3. Ограничение по массе.
4. Только для обслуживающих транспортных средств.
5. Только для общественного транспорта.

Кодом приоритета использования полосы для рассматриваемого участка дороги может быть любая непротиворечивая комбинация из пяти типов, определенных выше.

Каждый участок дороги может быть состыкован с последующим и предыдущим участком дороги. Для соединения участков дорог использует список пар кодов участков дорог в сетевой базе данных, чтобы представить соединения между предыдущим и последующим участком дороги.

Домен имеет следующие атрибуты: номер; длину; логическую функцию условий разрешения покинуть автомобиль домен; максимально разрешенную скорость на домене, которая зависит от кривизны участка дороги или ограничивается знаками; упорядоченный список автомобилей, находящихся на домене; время последнего обновления атрибутов домена.

Домен моделирует состояние участка дороги и имеет метод обработчика домена, который вычисляет предикат разрешения автомобилю покинуть домен, и, в зависимости от результата, планирует поведение автомобилей на данном домене. Функция-обработчик определяет время наступления следующего события. Под событием понимается изменение логического условия, изменение ускорения любым автомобилем (торможение, остановка, начало движения, достижение максимальной или заданной скорости),

прибытие или убытие автомобиля на (из) домен. Особо рассматривается первый в списке автомобиль. Только он может покинуть домен за время обработки одного события. Возможность покинуть первым автомобилем домен зависит от условий: достиг ли он конца участка, от его скорости и ускорения; разрешено ли ему покинуть домен; каково положение, скорость и ускорение последнего автомобиля на следующем участке.

Движение последующих автомобилей зависит только от их положения на домене, скорости, ускорения и таких же характеристик для предыдущего автомобиля.

В конце обработки из всех возможных событий выбирается ближайшее и заносится в таблицу событий, которая предназначена для управления процессом моделирования движения автомобилей в имитационной модели.

Данные относительно сетевой геометрии необязательные и необходимы только, если требуется графический интерфейс для пользователя. Эти данные определяют геометрию из двух основных линий (левые и правые полосы с ограничениями) для каждого участка дороги.

Пересечение путей с круговым движением функционирует как однонаправленная циркулирующая система вокруг центрального изолированного участка, где управление происходит по маркировке и приоритету движению на пути с круговым движением.

ПРЕДСТАВЛЕНИЕ ДИНАМИЧЕСКИХ ОБЪЕКТОВ

Обработка динамических объектов имитационной модели движения состоит из цикла обращения к набору модулей в определенные интервалы времени или при создании некоторых событий (дискретный событийный подход).

Моделирование начинается по определенному сценарию с загрузки параметров моделирования и загрузки статических компонент сети дорог. Затем инициализируется итерационная процедура с заданным размером шага. Задачи, которые выполняются внутри каждой итерации, включают:

1. Модификацию состояния сигналов движения, признаков и особых ситуаций.
2. Модификацию самых коротких путей и таблиц маршрутов.
3. Чтение новых мест назначения, начала координат таблицы маршрутов и соответствующее место транспортных средств в виртуальных очередях.
4. Загрузка транспортных средств из виртуальных очередей в сеть.
5. Модификация возможностей ускорения транспортных средств и проверка, должны ли они изменять полосу или допустимые интервалы для желательного изменения полосы.
6. В соответствии с новыми позициями транспортных средств модифицируются их скорости. Соответствующие данные (скорость, размещение и др.) записываются модулем системы наблюдения. В конце участка дороги транспортное средство или удаляется из сети (если оно достигает своего места назначения) или помещается на находящийся далее по ходу участок дороги.
7. Модифицируется отображение на дисплей, если допускается графический интерфейс пользователя.

8. Вычисляются меры эффективности (МЭ) или посылается состояние сети внешнему графическому интерфейсу пользователя (ГИП) или модулям.

9. Модифицируется время моделирования и возвращение на следующую итерацию.

Имитатор использует подход моделирования с временным разделением для обработки перемещений транспортных средств. В ответ на происходящие изменение полосы и получение информации от предыдущих автомобилей, связанные с изменением полосы, т.е. вызываются функции расчета динамических параметров для каждого транспортного средства каждые ω секунд. Скорости и позиции транспортных средств модифицируются каждый τ секунд. Размер шага должен быть в расчетном модуле большим, чем ω или равен размеру шага τ . Этот фиксированный размер шага принят для общего управления потоком данных. Для индивидуальных транспортных средств размер шага для вызова определенных функций моделирования может изменяться как время реакции водителя, встречающего специфические условия, например, резкое сокращение расстояния между транспортными средствами и др.

Модель выбора маршрута используется, чтобы фиксировать решения о выборе маршрута водителем в ответ на информацию о движении. В зависимости от доступа к информации в имитаторе приняты два типа водителей: информированные и не информированные. Поскольку водитель выбирает маршрут, то вычисляется вероятность выбора исходящей связи для каждого пересечения дорог. В настоящее время имитатор использует модель логического многочлена и рассматривает только задержку во времени движения при выборе маршрута (см. рис.2):

$$p(l|j, t) = \frac{\exp(u_l(t))}{\sum_{m \in L_j} \exp(u_m(t))}, \quad (1)$$

где $p(l|j, t)$ - вероятность выбора связи l для транспортного средства, которое ожидает прибытия в узел j во время t ; L_j - множество выходов из связи узла j ; $u_l(t)$ - выгода выбора следующей связи маршрута после связи l , которая определяется следующим образом:

$$u_l = \alpha \cdot c_l(t) + \beta \cdot C_k(t + c_l(t)), \quad (2)$$

где $c_l(t)$ - ожидаемое время для транспортного средства, чтобы пересечь связь l , которое вводит эту связь во время t ; $C_k(t + c_l(t))$ - ожидаемое время перемещения для самого короткого пути из узла k (находящегося до узла связи l) к месту назначения транспортного средства, которое достигает узла j во время t ; α , β - модельные параметры.

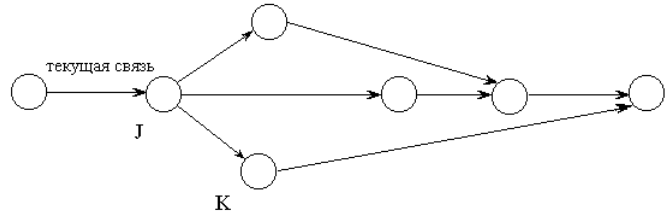


Рисунок 2 - Выбор маршрута

Множество выбора L_j , включает все соединения исходящих связей, находящихся далее по ходу узла текущей связи, которые могут выбрать транспортные средства поближе к его месту назначения. Таким образом, связь должна удовлетворить ограничению $C_k(t + c_l(t)) \leq C_j(t)$, чтобы быть включенной в набор выбора. Это ограничение также предотвращает транспортные средства от выбора следующего пути, который содержит циклы.

Могут также использоваться более сложные модели, которые включают другие алгоритмы выбора маршрута. Модульная конструкция имитатора учитывает простую замену заданных по умолчанию моделей.

ПРОГРАММНАЯ РЕАЛИЗАЦИЯ МИКРОИМИТАТОРА

Имитатор дорожного движения нижнего уровня (далее LLS - Low Level Simulator) состоит из следующих трех основных компонентов (рис. 3):

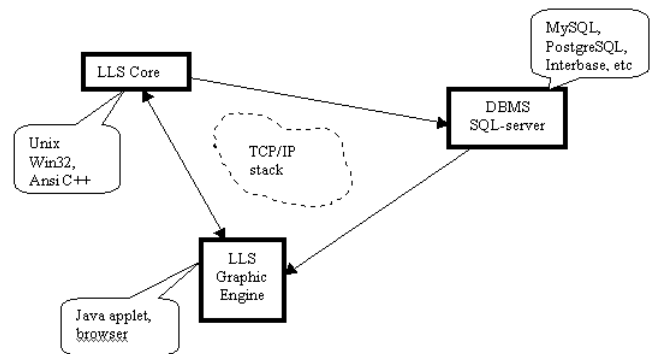


Рисунок 3 - Структура и взаимодействие компонентов системы имитатора

1. Ядро имитатора (LLS Core) - основной компонент системы, производящий все вычисления и управляющий самим процессом моделирования. Графическая подсистема (LLS Graphic Engine) предназначена для визуального отображения процесса моделирования. Она реализована в виде Java приложения, исполняющийся Виртуальной Ява Машиной, входящий в состав современных программ-

броузеров (Netscape Navigator 4 и Internet Explorer 4 и выше). Использование технологии Java наделяет имитационный проект несколькими немаловажными качествами: продвинутые графические возможности (прорисовка объектов в трехмерном пространстве) и независимость от платформы.

2. Сетевая реляционная система управления базами данных, поддерживающая структурированный язык запросов SQL (DBMS with SQL-server), например, Interbase, MySQL, PostgreSQL. СУБД служит для связи ядра и графической подсистемы имитатора.

Компоненты системы взаимодействуют между собой через локальную или глобальную компьютерную сеть с использованием протоколов семейства TCP/IP. Благодаря такой структуре компоненты системы могут размещаться как на одном компьютере, так и взаимодействовать через Internet. В первом случае логика работы имитатора остается неизменной, используя в качестве сетевого интерфейса устройство lo0 (network loopback, фиктивный интерфейс "петля"). По прошествии каждого цикла (tick) моделирования ядро системы заносит данные о координатах объектов (автомобилей) в таблицу базы данных откуда эти данные в свою очередь считываются графической подсистемой. Каждая строка в таблице, соответствует положению некоего объекта в конкретный момент модельного времени. Очевидно, что в случае, когда в системе находится несколько объектов, то с конкретным моментом модельного времени связан набор записей в таблице с информацией о состоянии объектов. В качестве идентификаторов таких наборов выступает определенное поле таблицы, в которое заносится значение модельного времени.

Такой алгоритм связи выбран из соображений модульности, и возможности распределенной обработки (на нескольких компьютерах) имитационной модели. Модульность системы заключается в том, что выбор другого продукта базы данных приведет к потребности изменить код модуля связи с СУБД и не повлияет на всю систему в целом.

Распределенная обработка заключается в том, что объекты системы являются сетевыми приложениями и не привязаны к какому-либо физическому компьютеру. Например, ядро и СУБД могут быть установлены на сервере, а Вы можете наблюдать за процессом моделирования с любой рабочей станции.

Так как все перемещения объектов заносятся и хранятся с СУБД, то процесс моделирования может быть воспроизведен в любое время с любого момента внутреннего модельного времени.

Каждому запущенному процессу ядра назначается идентификатор сессии (Session ID). Назовем сессией множество записей, занесенных в СУБД в процессе работы одного конкретного экземпляра ядра. Другими словами сессия - это результаты моделирования конкретного участка дороги. Информация о запущенных и отработавших сессиях заносится в СУБД. Также в СУБД заносятся данные (IP-адрес и номер порта) всех запущенных процессов ядра.

Временная синхронизация между графической подсистемой и ядром осуществляется также по сети. Каждый

экземпляр графической подсистемы после запуска, связывается с процессом ядра, сообщая ему свой IP-адрес и номер порта, и начинает наблюдать за сессией, которую оно исполняет. После занесения очередного набора записей с информацией об объектах в СУБД, ядро посылает сетевой пакет зарегистрировавшимся у него клиентам - экземплярам графической подсистемы. Формат пакета определяется специальным протоколом синхронизации SSP (Simulator synchronization protocol). В пакет включается текущее значение модельного времени ядра. Графическая подсистема фиксирует у себя значение модельного времени, для последнего отображенного набора. Такой механизм позволит добиться последовательного отображения информации об объектах даже в перегруженных сетях, где высока вероятность потери синхронизационных пакетов. К тому же такой механизм позволит очередному клиенту включиться в наблюдение в любой момент работы ядра.

Ядро имитатора нижнего уровня (LSS Core) написано на языке C++. Так как сетевые механизмы протокола TCP/IP большинства современных операционных систем (FreeBSD, Linux, Win NT, Win 98) основываются на коде BSD Sockets, разработанном в университете Berkeley, то код ядра может быть перекомпилирован без особых усилий под любую из ОС, в которой используется механизм Sockets.

Рассмотрим структуру ядра имитатора нижнего уровня (рис. 4).

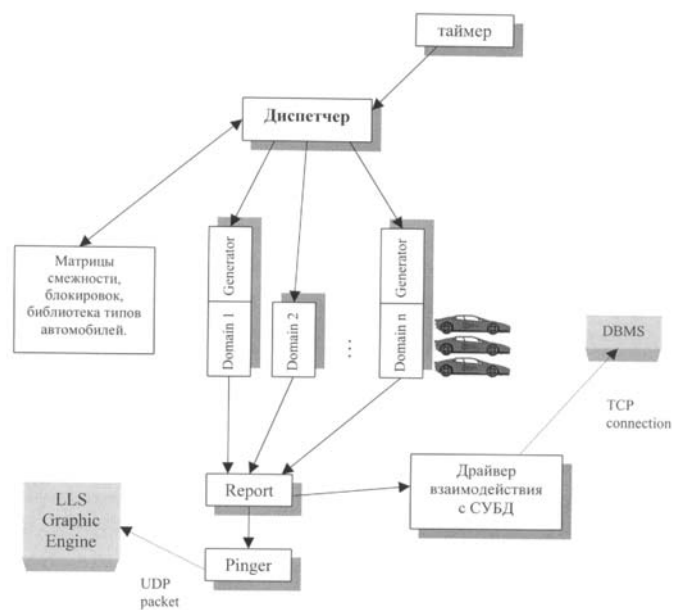


Рисунок 4

Основным компонентом ядра является класс-диспетчер. В его функции входит синхронизация работы всех компонентов ядра. Диспетчер содержит список указателей на домены. Класс таймер осуществляет контроль и учет модельного времени, запускает циклы моделирования.

Объект класса Домен, представляет основные параметры участка дороги. Такие как тип покрытия, длина

домена. Класс содержит указатель на список машин, находящихся на этом домене. Также класс содержит ряд служебных функций:

- drive - основная логика управления автомобилями. На каждом цикле пересчитывает координаты автомобилей, а также изменяет (по некоторому алгоритму) параметры автомобилей (ускорение) в зависимости от обстановки.

- car_in - функция вводящая в домен новый автомобиль.
- domain_exit_blocked - функция определяющая возможность покинуть домен.

```
class domain
{
    uint len;
    int uklon;
    int cover_type; // тип покрытия (асфальт, грунт, ямы)
    int code; // номер домена
public:
    int domain_exit_blocked();
    clist *car_list; // список машин
    void car_in( car * ); // новый автомобиль
}

```

По приходу сигнала от таймера функция-диспетчер запускает функцию drive для каждого домена из списка.

Алгоритм работы функции drive можно свести к следующему:

1. Пересчет координат и скорости автомобилей, на основании их свойств.

2. Изменение параметра ускорения, на основании дорожной обстановки.

Для входных доменов могут быть созданы экземпляры класса Generator, осуществляющие по определенной закономерности создание объектов-автомобилей с определенными характеристиками и включение этих автомобилей в домены.

Класс автомобиль представляет прообраз реальных автомобилей. Объект класса содержит указатель на домен, которому он принадлежит, основные текущие характеристики автомобиля (скорость, ускорение, позиция в домене), указатель на объект класса car_type, который определяет некоторые общие свойства, характерные для автомобиля данного типа.

```
class car
{
    int type;
    car_type *characteristics;
    domain *parent;
    int speed, acc; // текущие характеристики
    int x_position; // координаты передней части машины в домене
}

```

Класс car_type определяет основные свойства автомобиля. Набор классов типа car_type составляет библиотеку типов автомобилей. Основные свойства, такие как максимальная скорость, ускорение разгона и торможения, предпочитаемая скорость и ускорение торможения, ширина, длина, масса, цвет и тип кузова.

```
class car_type
{
    uint width, len;
    int max_speed, max_acc, max_deacc;
    int prf_speed, prf_deacc; // предпочитаемая скорость
    char color; // :-)
    char body_type;
}

```

Служебный класс ctype представляет собой интерфейс, для создания упорядоченных списков объектов. Является аналогом компонента Tlist из системы Delphi.

```
class clist;
{
    car *first;
    int item_number;
public:
    int max_car_num ();
    car *get_by_num ( int index );
}

```

ЗАКЛЮЧЕНИЕ

Имитационный проект - удобное средство моделирования процесса автомобильного движения на различных участках дорог. Он достаточно просто может быть привязан к топографической карте района, города или области, так как имитатор подробно отображает сети дорог и моделирует индивидуальные перемещения транспортных средств с использованием изменения полос и логики сигналов движения. Наличие средств визуализации позволяет, как строить саму имитационную модель участков дорог, так и отображать процесс моделирования перемещения транспортных средств.

ПЕРЕЧЕНЬ ССЫЛОК

1. Томашевский В.Н., Печенежский Д.С. Концептуальные основы имитационного моделирования автомобильного дорожного движения / Праці П'ятої Української конференції з автоматичного управління "Автоматика-98"; - ч. III - Київ: НТТУ "КПІ", 1998. С. 317 - 323.

КУБИЧЕСКОЕ МОДЕЛИРОВАНИЕ НЕИСПРАВНОСТЕЙ ЦИФРОВЫХ ПРОЕКТОВ НА ОСНОВЕ FPGA, CPLD

В.И.Хаханов, И.Ю.Сысенко, Хак Х.М.Джахирул, Масуд М.Д.Мехеди

Представлена кубическая технология анализа цифровых схем для генерации тестов и оценки их качества. Предложены кубический метод обработки неисправностей модели примитива и метод генерации тестов для константных неисправностей, использующий кубические покрытия списков неисправностей для построения пути активизации.

Представлена кубична технологія аналізу цифрових схем для генерації тестів та оцінювання їх якості. Запропоновані кубичний метод обробки несправностей моделі примітива та метод генерації тестів для константних несправностей, який використовує кубичні покриття списки несправностей для створення шляхів активізації.

Cubic methodology of analysis of digital devices for stuck-fault detection and test generation is represented. Cubic method of primitive models fault evaluation and test generation method for stuck-at-faults, which uses fault list cubic covering for path activization building, are offered.

1. ВВЕДЕНИЕ

Программируемые логические интегральные схемы (ПЛИС) FPGA (Field Programable Gate Array), CPLD (Complex Programable Logic Device) достойно конкурируют с БМК (базовыми матричными кристаллами), СБИС, сигнальными процессорами. Вместе с преимуществами ПЛИС существуют проблемы их тестирования, определяемые необходимостью разработки:

1) быстродействующих тестовых генераторов для конфигурации проектов, заданных в форме графов переходов конечных автоматов, булевых уравнений, многоуровневой иерархии;

2) программ моделирования неисправностей, предназначенных для проверки полноты тестов цифровых проектов вентильного, функционального и алгоритмического уровней описания;

3) средств верификации и диагностирования синтезированных цифровых устройств на основе FPGA и CPLD.

Возникновение качественно нового объекта диагностирования в виде ПЛИС стимулируют развитие новых методов и алгоритмов генерации тестов и моделирования неисправностей. Эти методы должны обладать следующими свойствами:

1) проектировать тест для одиночных константных неисправностей, с полнотой, близкой к 100%;

2) покрывать все линии объекта одномерными путями активизации;

3) иметь приемлемое быстродействие алгоритмов генерации тестов и моделирования неисправностей для обработки схем большой размерности;

4) иметь возможность параллельного выполнения векторных операций логического анализа для генерации те-

стов и оценки их качества;

5) обрабатывать устройства вентильного, функционального алгоритмического уровней описания;

6) поддерживать стандарт VHDL для описания цифрового устройства и полученного теста;

7) иметь возможность интегрирования в существующие системы автоматизированного проектирования ведущих фирм мира.

Процесс проектирования программируемых интегральных схем предполагает наличие эффективных средств их анализа. Для этой цели используется в том числе и моделирование неисправностей, которое предназначено для:

- оценки качества тестов;

- проектирования тестов;

- построения алгоритмов поиска дефектов;

- верификации проектов цифровых автоматов.

Объектами применения алгоритмов моделирования неисправностей являются цифровые автоматы, реализуемые в структуру FPGA, CPLD. Наличие нескольких миллионов вентилей на кристалле требует разработки быстродействующих методов, алгоритмов и программ моделирования неисправностей.

Теоретические источники: многозначная алгебра [1] - Hayes J.P., двухтактная кубическая алгебра [2], дедуктивный метод моделирования неисправностей [3-6] - Ермилов В.А., Armstrong D.B., Биргер А.Г., Levendel Y.H., методы генерации тестов [7] - Breuer M.A., Abramovici M.

Предпочтительным по быстродействию является дедуктивный метод моделирования неисправностей. Он позволяет за одну итерацию обработки схемы определить все константные дефекты, проверяемые на входном тестовом векторе. Но данный метод ориентирован на вентильный уровень описания цифровых схем. Это связано со сложностью проблемы генерирования выходных списков неисправностей (output fault list generation) для невентильных примитивов. Предлагаемое решение проблемы позволяет моделировать дефекты цифровых схем, описанных на функциональном и алгоритмическом уровнях. Обратной стороной решения упомянутой проблемы является метод генерации тестов для константных неисправностей, использующий кубические покрытия примитивов для построения заказанного пути активизации.

2. МАТЕМАТИЧЕСКИЙ АППАРАТ АНАЛИЗА ПРИМИТИВА

Автоматная модель последовательностного примитива представлена в виде:

$$M = \langle X, Y, Z, f, g \rangle,$$

где $X = (X_1, X_2, \dots, X_p, \dots, X_m)$,

$Y = (Y_1, Y_2, \dots, Y_p, \dots, Y_h)$, $Z = (Z_1, Z_2, \dots, Z_p, \dots, Z_k)$ - множества входных, внутренних и выходных автоматных переменных, отношения между которыми описываются характеристическими уравнениями:

$$Y(t) = f[X(t-1), X(t), Y(t-1), Z(t-1)];$$

$$Z(t) = g[X(t-1), X(t), Y(t-1), Y(t), Z(t-1)]. \quad (1)$$

Переменные $Z(t)$ отличаются от $Y(t)$ тем, что первые наблюдаемы по выходным линиям, а $Y(t)$ в этом смысле есть внутренние. Формат автоматных переменных для записи кубического покрытия, соответствующий (1), имеет следующий вид:

| | | |
|----------|----------|----------|
| $X(t-1)$ | $Y(t-1)$ | $Z(t-1)$ |
| $X(t)$ | $Y(t)$ | $Z(t)$ |

который может быть представлен автоматной моделью, изображенной на рис. 1.

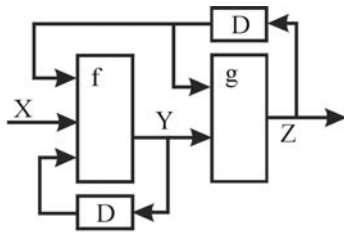


Рисунок 1 - Автоматная модель примитива

Функциональный последовательностный примитив задается компонентами:

$$F^2 = \langle (t-1, t), (X, Z, Y), \{A^2\} \rangle,$$

где $(t-1, t)$ - два автоматных соседних такта в описании функции; (X, Z, Y) - векторы входных, внутренних и выходных переменных; $\{A^2\}$ - двухтактный алфавит описания состояний (переходов) автоматных переменных [1,2]:

$$A^2 = \{Q = 00, E = 01, H = 10, J = 11, O = \{Q, H\}, I = \{E, J\}, A = \{Q, E\}, B = \{H, J\}, S = \{Q, J\}, P = \{E, H\}, C = \{E, H, J\}, F = \{Q, H, J\}, L = \{Q, E, J\}, V = \{Q, E, H\}, Y = \{Q, E, H, J\}, A^1 = \{0, 1, X = \{0, 1\}\}, (U)\}.$$

Примитив описывается кубическим покрытием

$$C = (C_1, C_2, \dots, C_p, \dots, C_n),$$

где $C_i = (C_{i1}, C_{i2}, \dots, C_{ip}, \dots, C_{iq})$ - куб, включающий входные, внутренние, выходные координаты $C_i = (C_i^X, C_i^Y, C_i^Z)$, $q = m + h + k$. Для комбинационного автомата формат описания кубического покрытия $F^1 = \langle (t), (X, Z), \{A^1\} \rangle$ определяется отношениями на $(q = m + k)$ -мерном векторе переменных $C_i = (C_i^X, C_i^Z)$. Формат задает многовыходовой комбинационный примитивный элемент с m входами и k выходами. Двоичная булева функция от m переменных $Z = f(X_1, X_2, \dots, X_m)$ определяется при $k = 1$.

Процедура получения двухтактного кубического покрытия:

1. Выполнение операции конкатенации (#):

| | | | | |
|--------------------|-----|-----|-----|-----|
| $C_{t-1} \# C_t =$ | # | 0 | 1 | X |
| | 0 | Q | E | A |
| | 1 | H | J | B |
| | X | O | I | Y |

для преобразования пары соседних во времени однотактных символов в один двухтактный.

2. Итеративная минимизация полученного покрытия по правилу - два куба объединяются в один, если они отличаются по одной переменной.

3. Верификация полученной модели примитива с позиции описания возможных состязаний.

Пример 1. Применение упомянутой процедуры к таблице переходов двухразрядного счетчика приводит к существенному уменьшению кубического покрытия, записанного в алфавите A^2 :

$$\begin{bmatrix} V & C & A_{t-1} & B_{t-1} & A_t & B_t \\ 0 & X & X & X & 0 & 0 \\ 1 & E & 0 & 0 & 0 & 1 \\ 1 & E & 0 & 1 & 0 & 0 \\ 1 & E & 1 & 0 & 1 & 1 \\ 1 & E & 1 & 1 & 1 & 0 \\ 1 & F & 0 & 0 & 0 & 0 \\ 1 & F & 0 & 1 & 0 & 1 \\ 1 & F & 1 & 0 & 1 & 0 \\ 1 & F & 1 & 1 & 1 & 1 \end{bmatrix} = \begin{bmatrix} V & V & A & B \\ 0 & X & 0 & 0 \\ 1 & E & Q & E \\ 1 & E & E & H \\ 1 & E & J & E \\ 1 & E & H & H \\ 1 & F & Q & Q \\ 1 & F & Q & J \\ 1 & F & J & Q \\ 1 & F & J & J \end{bmatrix} = \begin{bmatrix} V & C & A & B \\ 0 & X & 0 & 0 \\ 1 & E & S & E \\ 1 & E & P & H \\ 1 & E & S & S \end{bmatrix}.$$

3. ФОРМУЛИРОВКА ПРОБЛЕМЫ ТЕСТИРОВАНИЯ

Проблемы тестирования цифрового устройства определяются на множестве $\langle M - модель, L - кубическое покрытие списков неисправностей (КПСН) (Fault Lists Cubic Covering - FLCC), T - тест \rangle$ в условиях неопределенности

одного из компонентов. Будем считать, что модель M задается кубическим покрытием $C = (C_1, C_2, \dots, C_i, \dots, C_n)$. Покрытие неисправностей для примитива или цифровой схемы задается в виде:

$$L = (L_1, L_2, \dots, L_i, \dots, L_n),$$

где $L_i = (L_{i1}, L_{i2}, \dots, L_{ij}, \dots, L_{iq})$ - куб, включающий входные, внутренние, выходные координаты: $L_i = (L_i^X, L_i^Y, L_i^Z)$, $q = m + h + k$; $(L_{ij}^Y, L_{ij}^Z) = \{0, 1, X\}$; 0 - определяет вычитание (дополнение) списка L_j ; 1 - задает пересечение L_j ; $X = \{0, 1\}$ - идентифицирует несущественность списка неисправностей L_j . Если L_{ir}^Z - выходная наблюдаемая переменная, то (0) 1 - есть идентификатор (не-) проверки неисправностей куба L_j на выходе r , $X = \{0, 1\}$ - идентифицирует неопределенное состояние выходной координаты L_{ir}^Z , которое можно интерпретировать как 0, так и 1.

Проблема 1. Кубическое покрытие списков неисправностей L для вектора T и покрытия примитива C вычисляется по линейному уравнению

$$T \oplus C = L, \quad (2)$$

где \oplus - бинарная координатная операция XOR, определяющая взаимодействие компонентов T, C, L в троичном алфавите:

| | | | | |
|-----------------------|----------|-----|-----|-----|
| $T_j \oplus C_{ij} =$ | \oplus | 0 | 1 | X |
| | 0 | 0 | 1 | X |
| | 1 | 1 | 0 | X |
| | X | X | X | X |

(3)

Универсальная формула анализа КПСН, полученного в результате применения (3) к тест-вектору T и покрытию многовыходового примитива C для определения по выходу r списка проверяемых неисправностей L_r , имеет следующий вид:

$$L = \bigcup_{\forall i(T_i \oplus C_{i..} = 1)} \bigcap_{i=1}^n L_j^{T_i \oplus C_{ij}}, \quad (4)$$

где $L_j^{T_i \oplus C_{ij}} = \begin{cases} L_j \leftarrow T_j \oplus C_{ij} = 1; \\ \bar{L}_j \leftarrow T_j \oplus C_{ij} = 0, \end{cases}$

\bar{L}_j - рассматривается как список дефектов, относящийся к линии j , который следует вычитать из неисправностей,

проверяемых по невыходным линиям примитива; L_j - неисправности, которые необходимо пересекать с невыходными списками.

Замечание 1. Если тест-вектор определен в троичном алфавите $T_j = \{0, 1, X\}$, то после получения покрытия списков неисправностей, каждый куб следует верифицировать по правилу

$$L = L - L_i \Leftarrow (L_j \oplus T \neq C_i), i = \overline{1, n}. \quad (5)$$

Пример 2. Задано кубическое покрытие функции

$$C(\bar{X}_2 \bar{X}_3 \vee X_1 X_2 X_3) = \begin{array}{c|ccc|c} X_1 & X_2 & X_3 & Y \\ \hline X & 0 & 0 & 1 \\ 1 & 1 & 1 & 1 \\ X & 0 & 1 & 0 \\ X & 1 & 0 & 0 \\ 0 & 1 & X & 0 \end{array}. \quad (6)$$

Списки идентификаторов неисправностей для трех входов определены подмножествами:

$$\begin{aligned} L_1 &= \{1, 2, 4, 5\}; L_2 = \{1, 2, 3, 6\}; \\ L_3 &= \{1, 3, 4, 7\}. \end{aligned} \quad (7)$$

Определить выходной список неисправностей для тест-вектора 1111.

Решение сводится к выполнению \oplus -операции между вектором T и покрытием C :

$$T(1111) \oplus C = \begin{array}{c|ccc|c} X_1 & X_2 & X_3 & Y \\ \hline X & 0 & 0 & 1 \\ 1 & 1 & 1 & 1 \\ X & 0 & 1 & 0 \\ X & 1 & 0 & 0 \\ 0 & 1 & X & 0 \end{array} = L = \begin{array}{c|ccc|c} X_1 & X_2 & X_3 & Y \\ \hline X & 1 & 1 & 0 \\ 0 & 0 & 0 & 0 \\ X & 1 & 0 & 1 \\ X & 0 & 1 & 1 \\ 1 & 0 & X & 1 \end{array}.$$

Здесь L - кубическое покрытие списков неисправностей, по которому можно записать дизъюнктивную нормальную форму или теоретико-множественное уравнение для вычисления выходного списка дефектов на векторе $T(1111)$. Далее по единичным значениям выхода Y в покрытии L записывается дизъюнктивная нормальная форма (теоретико-множественное уравнение) проверяемых входных списков:

$$\begin{aligned} L &= X_2 \bar{X}_3 \vee \bar{X}_2 X_3 \vee X_1 \bar{X}_2 = \\ &= (X_2 - X_3) \cup (X_3 - X_2) \cup (X_1 - X_2). \end{aligned} \quad (8)$$

Подставив вместо переменных X_i соответствующие списки L_i , получим

$$L = (\{1, 2, 3, 6\} - \{1, 3, 4, 7\}) \cup (\{1, 3, 4, 7\} - \{1, 2, 3, 6\}) \cup \cup (\{1, 2, 4, 5\} - \{1, 2, 3, 6\}) = \{2, 4, 5, 6, 7\}. \quad (9)$$

Проблема 2. Тест проверки дефектов примитива, задаваемых L_i^1 -кубом покрытия списка неисправностей

$$L = (L_i^0, L_i^1), \text{ определяется по уравнению:}$$

$$L_i^1 \oplus C = T^k, \quad (10)$$

где L_i^0, L_i^1 - кубы, имеющие нулевые или единичные значения на выходной координате r . При этом формируются векторы-кандидаты в тест $T_i^k \in T^k$. Из множества

T^k формируется тест T , где в качестве элемента рассматривается набор $T_i \in T$. Каждый $T_i \in T$ должен удовлетворять условиям:

$$T_i = C_i \cap T_i^k \Leftarrow \exists i(C_i \cap T_i^k \neq \emptyset). \quad (11)$$

Пример 3. Построить тест проверки всех одиночных константных неисправностей для функции, заданной покрытием (6).

Сгенерируем L^1 -покрытие списка дефектов и для каждой строки L_i^1 выполним \oplus -операцию по (10):

$$L^1 \left(\begin{array}{c|ccc|c} X_1 & X_2 & X_3 & Y \\ \hline 1 & 0 & 0 & 1 \\ 0 & 1 & 0 & 1 \\ 0 & 0 & 1 & 1 \end{array} \right) \oplus C \left(\begin{array}{c|ccc|c} X_1 & X_2 & X_3 & Y \\ \hline X & 0 & 0 & 1 \\ 1 & 1 & 1 & 1 \\ X & 0 & 1 & 0 \\ X & 1 & 0 & 0 \\ 0 & 1 & X & 0 \end{array} \right) = T^k(L_i^1), \quad (12)$$

$$i = \overline{1, 3}.$$

К полученным подмножествам кандидатов в тест:

$$\begin{array}{c} T^k(L_1^1) \quad T^k(L_2^1) \quad T^k(L_3^1) \\ \left(\begin{array}{c|ccc|c} X_1 & X_2 & X_3 & Y \\ \hline X & 0 & 0 & 0 \\ 0 & 1 & 1 & 0 \\ X & 0 & 1 & 1 \\ X & 1 & 0 & 1 \\ 1 & 1 & X & 1 \end{array} \right); \quad \left(\begin{array}{c|ccc|c} X_1 & X_2 & X_3 & Y \\ \hline X & 1 & 0 & 0 \\ 1 & 0 & 1 & 0 \\ X & 1 & 1 & 1 \\ X & 0 & 0 & 1 \\ 0 & 0 & X & 1 \end{array} \right); \quad \left(\begin{array}{c|ccc|c} X_1 & X_2 & X_3 & Y \\ \hline X & 0 & 1 & 0 \\ 1 & 1 & 0 & 0 \\ X & 0 & 0 & 1 \\ X & 1 & 1 & 1 \\ 0 & 1 & X & 1 \end{array} \right) \end{array}$$

применяется процедура (11), которая определяет в данном случае уже минимальный тест, гарантировано проверяющий схемную структуру, которая соответствует двухуровневой реализации булевой функции (6):

$$T = \left(\begin{array}{c|ccc|c} X_1 & X_2 & X_3 & Y \\ \hline 0 & 1 & 1 & 0 \\ 1 & 1 & X & 1 \\ \hline X & 1 & 0 & 0 \\ 1 & 0 & 1 & 0 \\ X & 1 & 1 & 1 \\ X & 0 & 0 & 1 \\ 0 & 0 & X & 1 \\ \hline X & 0 & 1 & 0 \\ 1 & 1 & 0 & 0 \\ X & 0 & 0 & 1 \\ X & 1 & 1 & 1 \end{array} \right). \quad (13)$$

Утверждение 1. Тест проверки одиночных константных неисправностей одновыходового примитива, задаваемых L^1 -покрытием списка дефектов

$$L^1 = \left(\begin{array}{c|cccc|c} X_1 & X_2 & \dots & X_j & \dots & X_m & Z \\ \hline 1 & 0 & \dots & 0 & \dots & 0 & 1 \\ 0 & 1 & \dots & 0 & \dots & 0 & 1 \\ \dots & \dots & \dots & \dots & \dots & \dots & \dots \\ 0 & 0 & \dots & 1 & \dots & 1 & 1 \\ \dots & \dots & \dots & \dots & \dots & \dots & \dots \\ 0 & 0 & \dots & 0 & \dots & 1 & 1 \end{array} \right), \quad (14)$$

определяется процедурой:

1. Формирование кандидатов в тест путем выполнения операций над векторами из L^1, C :

$$T_i^k \in T^k, \quad (15)$$

2. Получение теста ($T_i \in T$) из множества кандидатов T^k на основе анализа выражения (11).

3. Минимизация множества тест-векторов T путем применения операции поглощения. Если $T_i \cap T_r = T_r$, то набор удаляется T_r из теста T .

Проблема 3. Определяет возможность верификации результатов, полученных при решении первых двух проблем

$$T \oplus L = C. \quad (16)$$

Важность данной процедуры связана с многочисленными ошибками, возникающими в процессе ручного или автоматизированного проектирования тестов. Кроме того, согласно замечанию 1, решение проблемы 3 необходимо для верификации покрытия неисправностей при наличии в тесте символов X . Для полученных наборов (13) верификация путем выполнения операций $L_i^1 \oplus T (L_i^1)$,

$i = \overline{1, 3}$ дает положительный результат:

$$L^1 \left(\begin{array}{c|ccc|c} X_1 & X_2 & X_3 & Y \\ \hline 1 & 0 & 0 & 1 \\ 0 & 1 & 0 & 1 \\ 0 & 0 & 1 & 1 \end{array} \right) \oplus T \left(\begin{array}{c|ccc|c} X_1 & X_2 & X_3 & Y \\ \hline 0 & 1 & 1 & 0 \\ 1 & 1 & X & 1 \\ X & 1 & 0 & 0 \\ 1 & 0 & 1 & 0 \\ X & 1 & 1 & 1 \\ X & 0 & 0 & 1 \\ 0 & 0 & X & 1 \\ X & 0 & 1 & 0 \\ 1 & 1 & 0 & 0 \\ X & 0 & 0 & 1 \\ X & 1 & 1 & 1 \end{array} \right) = C \left(\begin{array}{c|ccc|c} X_1 & X_2 & X_3 & Y \\ \hline X & 0 & 0 & 1 \\ 1 & 1 & 1 & 1 \\ X & 0 & 1 & 0 \\ X & 1 & 0 & 0 \\ 0 & 1 & X & 0 \end{array} \right).$$

Это означает, что любой вектор $L_i^1 \oplus T(L_i^1) \in C$ является кубом исходного покрытия функции. Следовательно, тест проверки неисправностей не имеет ошибок проектирования.

4. АЛГОРИТМ КУБИЧЕСКОГО МОДЕЛИРОВАНИЯ НЕИСПРАВНОСТЕЙ КОМБИНАЦИОННЫХ СХЕМ

На основе процедуры (4), определим алгоритм кубического дедуктивного моделирования неисправностей цифрового устройства:

1. Моделирование исправного поведения очередного примитива P_i ($i = \overline{1, M}$) на тест-векторе T_t ($t = \overline{1, N}$).

Если $t = N$ - формируется список $L(T)$ проверенных неисправностей на тесте T . Конец моделирования. Иначе, $t < N$ - переход к п.2.

2. Если все элементы схемы обработаны ($i = M$), выполняется сравнение векторов исправных состояний линий в двух соседних итерациях. Если векторы идентичны $T_t^r = T_t^{r-1}$, конец моделирования T_t и переход к п. 3. Иначе - переход к п. 1.

3. Определение списков неисправностей внешних входов в виде дополнения к их исправному состоянию $L_j = \{j^{\bar{T}_j}\}$.

4. Моделирование неисправностей примитива P_i ($i = \overline{1, M}$) по процедуре (4). Дополнение полученного списка неисправностью выходной линии примитива, идентифицируемой в виде $\{j^{\bar{T}_j}\}$.

5. Если ($i = M$), выполняется формирование списка проверяемых дефектов $L(T_i)$, и переход к п.1. Иначе, $i < M$ - переход к п. 4.

Быстродействие алгоритма сквозного интерпретативного синхронного моделирования неисправностей для одной итерации имеет оценку

$$W = \sum_{i=1}^M (n_i \times q_i),$$

n_i, q_i - число кубов и переменных в покрытии C_i .

Пример 6. Выполнить моделирование одиночных константных неисправностей схемы, представленной на рис.2.

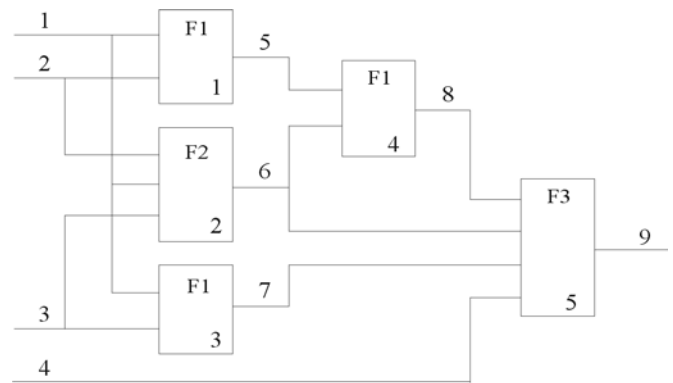


Рисунок 2 - Цифровой функциональный модуль

Поведение примитивов задается кубическими покрытиями:

$$\begin{array}{c} C(F1) \\ \hline 1 \ 2 \ 5 \\ 1 \ 3 \ 7 \\ 5 \ 6 \ 8 \\ \hline \left[\begin{array}{c|c} 1 & X \\ \hline 0 & 1 \\ 0 & 0 \end{array} \right] \end{array}; \quad \begin{array}{c} C(F2) \\ \hline 2 \ 1 \ 3 \ 6 \\ \hline \left[\begin{array}{c|c} X & 0 \\ \hline 1 & 1 \\ X & 0 \end{array} \right] \end{array}; \quad \begin{array}{c} C(F3) \\ \hline 8 \ 6 \ 7 \ 4 \ 9 \\ \hline \left[\begin{array}{c|ccc|c} X & 0 & 0 & X & 1 \\ \hline X & 1 & 0 & 1 & 1 \\ 0 & 1 & 1 & X & 1 \\ X & 0 & 1 & X & 0 \\ 1 & X & 1 & X & 0 \\ X & 1 & 0 & 0 & 0 \end{array} \right] \end{array}.$$

Рассматриваются одиночные константные неисправности линий, идентифицируемые как j^σ , где $\sigma = \{0, 1\}$ - знак константной неисправности, j - номер линии схемы. Для входного набора $T = 1111$ моделируется исправное поведение объекта, которое дает вектор (111101011). Затем генерируется список $L_j = \{j^{\bar{T}_j}\}$ входных дефектов цифрового устройства:

$$L_1 = \{1^0\}; L_2 = \{2^0\}; L_3 = \{3^0\}; L_4 = \{4^0\}.$$

Выполняется последовательная обработка всех элементов схемы для вычисления выходных списков неисправностей с учетом (4). Для первого примитива результат моделирования имеет вид

$$L_5 = T(110) \oplus C_{F1} \begin{bmatrix} 1 & 2 & 5 \\ 1 & X & 0 \\ 0 & 1 & 1 \\ 0 & 0 & 0 \end{bmatrix} = L \begin{bmatrix} 1 & 2 & 5 \\ 0 & X & 0 \\ 1 & 0 & 1 \\ 1 & 1 & 0 \end{bmatrix} = L_1 - L_2 = \{1^0\}.$$

Затем к полученному списку следует прибавить неисправность выходной линии, которая инверсна по отношению к значению координаты тест-вектора. В данном случае получаем список $L_5 = \{1^0, 5^0\}$.

Последовательная обработка оставшихся примитивов в схеме дает следующие результаты:

$$L_6 = \{1^0, 2^0, 3^0, 6^0\}; L_7 = \{1^0, 7^1\}; L_8 = \{1^0, 2^0, 3^0, 5^1, 6^0, 8^0\}; L_9 = \{1^0, 4^0, 7^1, 9^0\}.$$

Количество проверяемых дефектов F по q линиям формирует оценку качества теста (вектора) $Q = [F/(2q)]100\%$. Для тест-вектора 1111 получаем: $Q = [4/29]100\% = 22\%$.

5. МОДЕЛИРОВАНИЕ НЕИСПРАВНОСТЕЙ В ПОСЛЕДОВАТЕЛЬНОСТНЫХ ПРИМИТИВНЫХ ЭЛЕМЕНТАХ

Функциональная зависимость списка дефектов по наблюдаемому выходу Z_r согласно (1) определяется уравнением

$$L_r^t = f(T), (C), (L_X^{t-1}, L_X^t, L_Y^{t-1}, L_Z^{t-1}), \quad (17)$$

где $T_t \in T = (T_1, \dots, T_p, \dots, T_p)$ - пара соседних входных воздействий, где каждая координата определена в следующих сочетаниях:

$$T_t = \begin{bmatrix} 0 & 0 & 1 & 1 & 0 & 1 & X & X & X \\ 0 & 1 & 0 & 1 & X & X & 0 & 1 & X \end{bmatrix}. \quad (18)$$

Двухфреймовый формат входного набора ориентирован на анализ последовательностного автомата, поскольку его покрытие в общем случае задано в двухтактном алфавите A^2 . Для технологичного выполнения операции XOR между координатами тест-вектора и кубического покрытия $T_{ij} \oplus C_{ij}$ определим ее в виде табл.1.

Каждая координата таблицы есть сокращенная форма записи списков неисправностей

$$L_j = T_{ij} \oplus C_{ij}.$$

Например, если $T_{ij} = (01)$, $C_{ij} = P$, то по входной

координате j в соответствии с табл. 1 получается: $T_{ij}(01) \oplus P\{(01), (10)\} = (\bar{L}_j^{t-1} \bar{L}_j^t) \vee (L_j^{t-1} L_j^t)$. Таблица 2 предназначена для доопределения выходных координат куба покрытия списков неисправностей в моменты $(t-1, t)$. Например, для координаты $L_{ij} = V$ ее значение в такте $t-1$ равно 0, а в момент t равно 1, что в табл.2 определяется буквой E. Исключение составляют преобразования символов S, P . Здесь существует различие интерпретации в зависимости от того, рассматривается ли выходная переменная функцией или аргументом к выходу, для которого строится список проверяемых неисправностей. В первом случае доопределение упомянутых символов дает (J, E) , во втором - (X, X) , что свидетельствует об отсутствии списков неисправностей для данного выхода в момент $t-1$.

Таблица 1

| | | | | | | | | | | | | | | | | | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| ⊕ | 0 | 1 | X | Z | G | T | K | N | Q | E | H | J | O | I | A | B | S | P | C | F | L | V | Y | | |
| 0 | 0 | 1 | X | Z | G | T | K | N | Q | E | H | J | O | I | A | B | S | P | C | F | L | V | Y | | |
| 0 | 1 | 0 | X | Z | T | G | K | N | E | Q | J | H | I | O | A | B | P | S | F | C | V | L | Y | | |
| 1 | 0 | 0 | 1 | X | Z | G | T | K | N | H | J | Q | E | O | I | B | A | P | S | L | V | C | F | Y | |
| 1 | 1 | 0 | 0 | 1 | X | Z | T | G | K | N | J | H | E | Q | I | O | B | A | S | P | V | L | F | C | Y |
| 0 | X | X | X | X | G | T | K | N | A | A | B | V | Y | Y | A | B | Y | Y | B | B | A | A | Y | | |
| 1 | X | X | X | X | T | G | K | N | B | B | A | A | Y | Y | B | A | Y | Y | A | A | B | B | Y | | |
| X | 0 | 1 | X | Z | X | X | X | X | O | I | O | I | O | I | Y | Y | Y | Y | I | O | I | O | Y | | |
| X | 1 | 0 | X | Z | X | X | X | X | I | O | I | O | I | O | Y | Y | Y | Y | O | I | O | I | Y | | |
| X | X | X | X | X | X | X | X | X | Y | Y | Y | Y | Y | Y | Y | Y | Y | Y | Y | Y | Y | Y | Y | | |

Таблица 2

| | | | | | | | | | | | | | | | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| 0 | 1 | X | Z | G | T | K | N | Q | E | H | J | O | I | A | B | S | P | C | F | L | V | Y | |
| Z | 1 | I | Z | G | T | Z | Z | G | E | T | J | Z | I | E | J | X | E | X | I | J | I | E | I |

Для вычисления собственных неисправностей триггера следует учитывать следующие правила:

- 1) терм, имеющий более одной переменной без инверсии или все переменные с инверсией вычеркивается;
- 2) оставшиеся термы должны иметь одну переменную без инверсии, которая формирует собственную проверяемую неисправность $j^{\bar{T}_j}$, инверсную состоянию переменной на тесте.

Пример 9. Определить списки проверяемых неисправностей счетчика на двух тест-векторах. Моделирование первого набора дает результат:

$$T_1 \begin{bmatrix} 1 & 0 & 0 & 0 \\ 1 & 1 & 0 & 1 \end{bmatrix} \oplus C_{count} \begin{bmatrix} V & C & A & B \\ 0 & X & 0 & 0 \\ 1 & E & S & E \\ 1 & E & P & H \\ 1 & F & S & S \end{bmatrix} = L \begin{bmatrix} V & C & A & B \\ 1 & X & 0 & 1 \\ 0 & Q & S & Q \\ 0 & Q & P & J \\ 0 & C & S & P \end{bmatrix},$$

$$L_A(T_1) = \bar{V}_t \bar{C}_{t-1} \bar{C}_t A_{t-1} \bar{B}_{t-1} \vee \bar{V}_t \bar{C}_{t-1} \bar{C}_t \bar{A}_{t-1} B_{t-1} \vee \bar{V}_t (C_{t-1} \vee C_t) A_{t-1} \vee A_t = A_{t-1} \vee B_{t-1} \vee A_t;$$

$$L_B(T_1) = V_t \vee \bar{V}_t \bar{C}_{t-1} \bar{C}_t B_{t-1} \vee \bar{V}_t (C_{t-1} \vee C_t) \bar{B}_{t-1} \vee B_t = V_t \vee B_{t-1} \vee C_{t-1} \vee C_t \vee B_t;$$

$$L(T_1) = V_t^0 \vee C_{t-1}^1 \vee C_t^0 \vee A_{t-1}^1 \vee B_{t-1}^1 \vee A_t^1 \vee B_t^0,$$

где инверсия переменной означает дополнение (вычитание) списка неисправностей.

Моделирование второго набора представлено ниже:

$$T_2 \begin{bmatrix} 1 & 1 & 0 & 1 \\ 1 & 0 & 0 & 1 \end{bmatrix} \oplus C_{\text{count}} \begin{bmatrix} V & C & A & B \\ 0 & X & 0 & 0 \\ 1 & E & S & E \\ 1 & E & P & H \\ 1 & F & S & S \end{bmatrix} = L \begin{bmatrix} V & C & A & B \\ 1 & X & 0 & 1 \\ 0 & J & S & H \\ 0 & J & P & E \\ 0 & V & S & S \end{bmatrix}.$$

$$L_{Q_1}(T_2) = \bar{V}_t C_{t-1} C_t A_{t-1} \bar{B}_{t-1} \vee \bar{V}_t C_{t-1} C_t \bar{A}_{t-1} \bar{B}_{t-1} \vee \bar{V}_t (\bar{C}_{t-1} \vee \bar{C}_t) A_{t-1} \vee A_t = A_{t-1} \vee A_t;$$

$$L_{Q_2}(T_2) = V_t \vee \bar{V}_t C_{t-1} C_t \bar{B}_{t-1} \vee \bar{V}_t (\bar{C}_{t-1} \vee \bar{C}_t) B_{t-1} \vee B_t = V_t \vee B_t;$$

$$L(T_2) = V_t^0 \vee A_{t-1}^1 \vee A_t^1 \vee B_t^0.$$

Таким образом, проверяемые неисправности относительно выходов счетчика составляют множество $L(T_1)$ и $L(T_2)$, где константная неисправность отмечена верхним индексом.

ЗАКЛЮЧЕНИЕ

Метод кубического моделирования неисправностей является новой технологией обработки цифровых схем табличного (вентильного, функционального, алгоритмического) уровня представления. Он позволяет за одну итерацию моделировать все одиночные константные неисправности цифрового устройства, проверяемые тест-вектором. Усло-

вие применения заключается в наличии табличной формы описания примитивов цифрового устройства. Предложенный метод также эффективно обрабатывает и последовательностные примитивы цифровых автоматов, описанные двухтактными кубическими покрытиями [2]. Последние формализуют описания алгоритмов в виде примитивов, соответствующих графам переходов, граф-схемам, таблицам переходов цифровых автоматов.

Предложенная технология тестирования по уравнению $T \oplus C = L$ дает возможность: моделировать неисправности на основе анализа кубического покрытия; получать дедуктивные формулы для любых типовых функциональных элементов; проектировать компилятивные симуляторы для обработки цифровых устройств произвольного уровня описания; генерировать тесты для цифровых систем на основе использования КПСН; верифицировать результаты моделирования неисправностей и генерации тестов, проектировать аппаратные быстродействующие симуляторы.

Модели и методы реализованы в виде программных приложений. Последние используются при генерации тестов для проектов на основе FPGA и CPLD. Класс обрабатываемых структур: конечные автоматы в виде графов переходов, а также булевы уравнения с триггерными схемами. Входной язык описания цифровых систем - VHDL. Программы генерации тестов и моделирования неисправностей ориентированы на интеграцию в системы проектирования фирм: Aldec, Xilinx.

ПЕРЕЧЕНЬ ССЫЛОК

1. Hayes J.P. A systematic approach to multivalued digital simulation // ICCD-84: Proc. IEEE Int. Conf. Comput. 1984. No. 4. P. 177-182.
2. Хаханов В.И. Техническая диагностика элементов и узлов персональных компьютеров. К.: ИЗМН. 1997. 308 с.
3. Ермилов В.А. Метод отбора существенных неисправностей для диагностики цифровых схем. Общие выражения для неисправностей, возможных при эксперименте // Автоматика и телемеханика, 1971, № 1.- С. 159-167.
4. Armstrong D.B. A deductive method of simulating faults in logic circuits. IEEE Trans. on Computers. Vol. C-21. No. 5. 1972. P. 464-471.
5. Биргер А.Г. Многозначное дедуктивное моделирование цифровых устройств // Автоматика и вычислительная техника. - 1982.- №4.- С.77-82.
6. Levendel Y.H. and Menon P.R. Comparison of fault simulation methods - Treatment of unknown signal values.- Journal of digital system.- Vol.4.- 1980.- P.443-459.
7. Abramovici M., Breuer M.A. and Friedman A.D., Digital System Testing and Testable Design, Computer Science Press, 1998. 652 p.

ТОРСИОННЫЕ ПОЛЯ И НЕЙРОКОМПЬЮТЕР

В.Д.Цыганков

В работе показана связь нейрокомпьютерных технологий с торсионными полями и теорией физического вакуума. В качестве основы используется нейрокомпьютер "Эмбрион".

In this work the link of neural computer technologies with toric fields and theory of physical vacuum is shown. As the basis of work the neural computer Embrion is used.

ВВЕДЕНИЕ

Настоящее время характеризуется появлением новых, в том числе и нетрадиционных, технологий. Среди них следует особо выделить:

- нанотехнологию,
- новые информационные технологии,
- торсионные технологии.

Эти технологии возникли как следствие получения новых, более глубоких знаний о законах Природы, о внутренних механизмах отдельных процессов и явлений, происходящих в ней.

Данная работа посвящена попытке показать связь новых информационных технологий, в частности, нейрокомпьютерных технологий, с торсионными полями (ТП) и теорией физического вакуума (ФВ).

Авторы теории ФВ и ТП до сих пор не могут убедить большинство физиков в реальности ТП и вскрыть их природу [1]. Почему, например, плотность ФВ на 95 порядков превышает плотность воды? А почему не на 98 или 45, и как измерить плотность вакуума? В работе «"Колдуны" в законе» [2] прямо ставится под сомнение исходный постулат и, как следствие, вся теория ФВ, а именно, тождества типа $0=0$, описывающие уровень "Абсолютное НИЧТО" или 7-й уровень субфизики по Шипову [3], из которого возникают, создаются Творцом ТП, частицы, вещество, весь наш мир.

Однако я хочу встать на сторону концепции теории ФВ и ТП. (Эта теория достаточно сложна и доступна лишь очень подготовленному в теоретической физике специалисту.) Это желание, а скорее потребность, возникли на основании исследования вида и свойств виртуальных дискретных информационных полей и структур, которые возникают при возбуждении квазинеуронных сетей в разработанном мною с сотрудниками нейрокомпьютере (НК) "ЭМБРИОН" [4]. Я невольно обратил внимание на некую общность природы и внутренней структуры ФВ, ТП и виртуальных информационных полей и их структур, возникающих в НК. Возможно, такое сравнительное рассмотрение этих явлений поможет глубже понять природу ФВ и ТП и создать уже рабочие нетрадиционные торсионные технологии для их практического использования.

Следует обратить внимание на то, что НК "Эмбрион", его виртуальные информационные поля - это не фило-

софское, теоретическое и математическое построение, или абстрактная эвристическая модель, и не программная реализация нейрокомпьютерной модели на обычном персональном компьютере, а аппаратно изготовленные и испытанные на ряде актуальных задач опытные образцы электронных нейрокомпьютеров.

Итак, обратимся к семи уровням мироздания и понятию праматерии - этому "ничто", из чего все возникает, и все в нее возвращается [5]. Ее предполагаемая самоскомпенсированность по заряду, право - и левовращениям, массе покоя делает ее недоступной для мира физической реальности [6], мира физической материи.

Реальны ли ТП?

Я приведу определение "физической реальности", данное А. Эйнштейном, в качестве рабочего критерия определения наличия ТП и торсионных излучений По А. Эйнштейну [6] достаточным критерием физической реальности является существование элемента этой реальности, соответствующего некоторой физической величине, если можно достоверно предсказать значение этой величины без разрушения системы в процессе измерения и эксперимента.

Следующее понятие, которое нам понадобится - это понятие "самозарождение" или возникновение "ниоткуда", из ФВ виртуальных частиц, пар "частица-античастица" и их исчезновение в той же точке среды (ФВ), где они возникли. Этот флюктуационный процесс свидетельствует, правда косвенно, о существовании материальной среды, находящейся за порогом физической реальности (в "зазеркалье" как говорит Н. Бехтерева).

Итак, принимается, что существует неовещественная, не материализованная среда с массой покоя равной нулю или Физический Вакуум и физическая материя, вещество в виде частиц, плазмы, полей, атомов, молекул, твердых, жидких и газообразных тел, явлений, предметов нашего реального мира. Вакуум не нейтральная среда. Вакуум и Надвакуумный мир, в котором мы с Вами живем и существуем, можно представить в виде семи уровней реальности. Наибольший интерес, с точки зрения их фундаментальности и прямого отношения к моему сообщению представляют V, VI и VII уровни субфизики.

Это следующие уровни: "физический вакуум"(V), "первичный вакуум" (первичные торсионные поля)(VI) и "абсолютное ничто" (особое состояние материи)(VII). Начнем с высшего седьмого уровня.

1. УРОВЕНЬ "АБСОЛЮТНОЕ НИЧТО" (АН)

Он описывается двумя исходными фундаментальными тождествами:

закон сохранения энергии

$$0=0, \quad (1)$$

закон сохранения энергии-импульса и количества движения (или 4-х мерного импульса)

$$0=0. \tag{2}$$

Этот седьмой уровень есть безграничное упорядоченное многообразие пронумерованных точек с заданной геометрией [5].

Два другие тождества или структурные уравнения Картана, описывающие элементы седьмого уровня в момент его рождения или сотворения Творцом, имеют следующий вид:

$$\nabla \left[\begin{smallmatrix} e^a \\ k^{ij} \end{smallmatrix} \right] T^i_{[kj]} e^a_i = 0, \tag{3}$$

$$R^i_{jkm} = -2\nabla \left[\begin{smallmatrix} T^i_{[kj]} \\ k^{ijlm} \end{smallmatrix} \right] - 2T^i_c \left[\begin{smallmatrix} T^c_{[ijm]} \\ k^{ijlm} \end{smallmatrix} \right] = 0, \tag{4}$$

где T^i_{kj} - поле инерции или "внутренние" поля, R^i_{jkm} - тензор Римана или "внешние" поля переносящие энергию, e^a_j - базис четырехмерной ускоренной системы отсчета.

Тождества (3) и (4) описывают пустое пространство, не имеющее ни кривизны, ни кручения. Но это пространство, тем не менее, пронумеровано, т.е. имеет "первичное сознание", которое осознает "Абсолютное ничто" и делает его упорядоченным. Это активное начало или Бог.

Момент активизации или представления "Абсолютного Ничто" как безграничного многообразия с геометрией абсолютного параллелизма и есть точка сингулярности (особая точка) или момент рождения, создания или творения VII-го уровня.

Какие свойства проявляются на этом уровне?

Во-первых, это среда с программой, матрицей или планом возможного. В них заложены структура и свойства всех нижележащих уровней реальности. Для реализации возможного и превращения его в действительное (реальное), нужно, во-вторых, активное воздействие или воля и сознание. Планы и возможности на уровне VII существуют потенциально (или виртуально). Этот уровень реальности можно назвать по-разному, кому как это необходимо и приемлемо: Бог, Абсолют и др.

"Тонкая материя" этого уровня заполняет все бесконечное пространство, а на V-м уровне физического вакуума - это "грубая материя", которая заполняет всю Вселенную и, которую мы знаем и воспринимаем как все наше реальное пространство. Поле сознания (или уровень VII) не локализовано ни в одной из точек Вселенной. Оно есть всюду.

Итак, уровень VII - "Абсолютное Ничто" это нелокализованное сверхсознание.

Нами рассмотрено представление о седьмом уровне создателей теории ФВ и ТП [5]. А при чем же здесь нейрокompьютер?

Нейрокompьютер (НК) "Эмбрион" как некая физическая модель, содержит основные функциональные и структурные компоненты, такие же, как и живой мозг. Его блок-схема показана на рисунке.

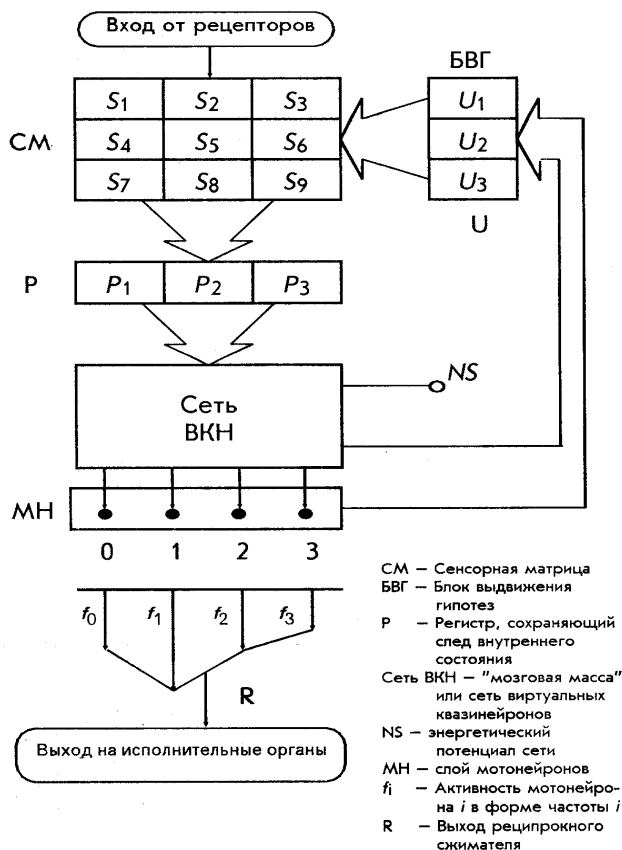


Рисунок 1 - Блок-схема НК "Эмбрион"

НК генерирует множество виртуальных квазинейронных, которые возникают в процессе возбуждения неравновесной структуры (поляризованного ФВ) первичного атома или первоатома Живой Вселенной [7].

n -атом - это некий полевой объект или вихрь в виде виртуального пространства-времени (ПВ), элементы которого могут быть названы фитонами Акимова.

Два связанных n -разрядных регистра P и S , содержащих n частиц или фитонов четырех видов. Два из них, c и g (операторы G_1 и G_4) - пассивные или неполяризованные частицы, а два вида, e и d (операторы G_2 и G_3) - активные или поляризованные частицы. Первоатом - это упорядоченное пронумерованное многообразие точек с геометрией тора (или кольца) и есть нейрокompьютерная модель VII-го уровня реальности или уровня "Абсолютного ничто".

Покажем возможную физическую интерпретацию тождеств вида $0=0$ (1) и (2) и нейрокompьютерную природу структурных уравнений Картана (3) и (4).

Исходное фундаментальное тождество (1) или закон сохранения энергии для НК "Эмбрион" интерпретируется как условие сохранения разрядности нейрокompьютера или размерности первоатома ($n = const$) и сохранения суммы вероятностей возможных состояний нейронов ($\sum P = 1$) при любых преобразованиях информации, при любых двоичных числах в регистрах P и S . Энергия ниоткуда не

появляется и не исчезает.

Закон сохранения энергии-импульса (2) - это, в НК как стохастическом автомате, закон сохранения суммы вероятностей переходов $(\sum P_{ij}) = 1$ в матрице возможного. О ней речь пойдет ниже.

Таблица 1 - Матрица возможного уровня

| Входное слово | Внутреннее состояние | Выходное слово | | | | |
|---------------|----------------------|----------------|---------|-------------|---------|-------------|
| | | Y_1^{t+1} | \dots | Y_j^{t+1} | \dots | Y_m^{t+1} |
| P_1^t | P_1^t | W_{11} | \dots | W_{1j} | \dots | W_{1m} |
| \vdots | | \vdots | | | | \vdots |
| S_1^t | P_i^t | W_{i1} | | W_{ij} | | W_{im} |
| \vdots | | \vdots | | | | \vdots |
| S_1^t | $P_m^t = 2^n$ | W_{m1} | \dots | W_{mj} | \dots | W_{mm} |

Уровень "Абсолютное ничто" - это среда с программой или матрицей возможного (смотри таблицу 1). В ней виртуально или потенциально заложены структура и свойства всех нижележащих уровней реальности. Это есть "Тонкая материя" уровня VII.

Для реализации возможного и превращения его в действительное имеется активное внутреннее начало - силовое воздействие от генератора активности NS или, как это активное начало называет Акимов, волевое воздействие, точнее - воля. Другим условием такого превращения-перехода служит пока нелокализованное поле сознания. В НК "Эмбрион" - это множество нереализованных гипотез восприятия $\{U\}$, находящихся в данный момент в Блоке Выдвижения Гипотез (БВГ) или подкорке мозга нейрокомпьютера. Нелокализованное поле сознания потенциально охватывает равномерно все пространство п-атома посредством генератора белого шума. Это общая флюктуация ФВ.

Начальный момент внутреннего, собственного ($U=0$) времени мозга нейрокомпьютера есть точка сингулярности. Это точки: 000, сингулярность 1, точка X_n , точка $Y_3 = 2100$ как моменты начала активизации матрицы возможного. Это моменты рождения и проявления сознания [9], [11].

Итак, уровень VII - "Абсолютное ничто" - это нелокализованное и не проявленное сознание, нереализованная гипотеза восприятия $\{U\}$ под воздействием воли NS. Локализация сознания - это попадание случайного волевого импульса $NS = 1$ из генератора белого шума (ГШ) в один из n разрядов регистра P. Это и будет проявление созна-

ния [8], [9].

Рассмотрим нейрокомпьютерную интерпретацию уравнения (3)

$$\nabla \begin{bmatrix} k^{e^a} \\ j \end{bmatrix} = -T^i \begin{bmatrix} k \\ j \end{bmatrix}^{e^a}$$

В нем геометрия ускоренной системы отсчета (слева) определяется "внутренним" полем инерции T_{kj}^i (справа).

В НК при невязке равной нулю ($J=0$) величина внутреннего поля инерции определяется расстоянием по Хеммингу между кодами состояния регистра $P_1 = (0, 0, 0, \dots, 0)$ и состояния $P_2 = (1, 1, 1, \dots, 1)$.

Уравнение (4) описывает характеристику невязки ($J>0$) - неуравновешенности между состоянием внешней S оболочки п-атома и состоянием его внутреннего ядра P.

Тензор Римана R_{jkm}^i в данном случае характеризует "внешние" силовые или энергетические поля. Внешнее поле может быть создано в НК путем замены двоичного кода j -й строки Сенсорной Матрицы (СМ) S_j на $j+1$ -ю строку S_{j+1} с новым, несовпадающим кодом. Внешние поля могут действовать не только на внешнюю S_j оболочку п-атома, но и на его ядро P путем замены его двоичного кода.

Итак, в точке сингулярности 1 при $U=0$ мы имеем в НК, в общем случае, созданный Творцом или Главным конструктором НК VII-й уровень физической реальности или "Абсолютное ничто" как напряженную виртуальную матрицу возможного, нереализованную волю ($NS=0$) и нелокализованное Сверхсознание в виде множества ($\{U\}$) гипотез восприятия, погруженные в Физический Вакуум.

Рассмотрим теперь шестой уровень (VI) - Первичный Вакуум или первичное торсионное поле.

2. УРОВЕНЬ VI - ПЕРВИЧНЫЕ ТОРСИОННЫЕ ПОЛЯ (ПТП)

Будем, в соответствии с [5], считать, что Сверхсознание или Сверхразум и воля уровня VII, рассмотренного в предыдущем разделе, реализуют матрицу возможного в виде первичного вакуума или в виде первичных полей кручения или торсионных полей (ПТП). Они представляют собой в НК набор информационных вихрей, вид которых можно представить в виде квазичастицы или "Живой Вселенной" [16]. Этих вихрей бесконечное множество, мощность которого зависит от n - числа разрядов P и S регистров НК. Это квантовые объекты или квантовые вихреобразные частицы информации и сами являются источниками излучений физических полей, в том числе, и электромагнитных полей в широком диапазоне спектров. Эти вихри заполняют все пространство Вселенной на VI-м уровне реальности.

Уравнения (3) и (4) уровня VII преобразуются, начиная с точки сингулярности при $U>0$, в уравнения (5) и (6), описывающие динамику первичных торсионных полей:

$$T_{jm} = 2/v \left(\nabla \left[\begin{matrix} r_{i|jm}^i \\ r_{i|jm}^s \end{matrix} \right] + T^i \right), \quad (5)$$

$$E = \int T_{jm} g_{jm} (-g)^{1/2} dV = 0. \quad (6)$$

Материя как вещество на этом VI-м уровне не существует в виде вещества. VI-й уровень в НК "Эмбрион" - это уровень виртуального вероятностного поля или уровень полевой материи.

Действительно, запустим НК - нашу квантовую систему, т.е. выполним условие: $NS \gg 1$; $U > 0$. Возникает вероятностное торсионное поле. Информационные вихри не имеют массы покоя. Это безмассовые (в классическом понимании этого термина) объекты физической реальности. В каждой точке бифуркации торсионного поля вихри взаимодействуют, рождаются и исчезают, что хорошо видно при испытании НК. Происходит не энергетический, а информационный обмен между ними и ФВ. Имеет место перетекание потоков вероятностей по информационным уровням-кодам между локальными аттракторами и концентрация торсионного поля в глобальном аттракторе. Это, например, точка X_k , нейроны типа (110) и (011).

На данном шестом уровне реальности нет обычного вещества, с его зарядом, массой, а имеются лишь квантовые взаимодействующие виртуальные вероятностные вихри. Этот уровень можно интерпретировать в виде голограммы, заполняющей всю нашу Вселенную.

Это значит, что структура невязки J (или волновой геном НК) развернута в пространстве-времени (ПВ) и ее информация распределена, "размазана" по всему объему ПВ в виде некоторой волновой функции Ψ возбужденного n -атома.

Процесс генерации первичного торсионного поля вероятно детерминирован и предопределен матрицей возможного. А значит, возможен и обратный процесс свертки или восстановления информации или исходного образа при движении внутреннего времени НК вспять (обращение времени $U \rightarrow -U$).

Математически процесс генерации ПТП может в НК "Эмбрион" быть описан уравнением Колмогорова-Чемпена

$$P_r^U = P_r^{(U-1)} |W_{ij}| \quad (7)$$

или возведением в U -ю степень матрицы возможных переходов $|W_{ij}|$.

Все точки поля связаны между собой через фазовый портрет голограммы. Это в высшей степени когерентное поле. Пространство нашей Живой Вселенной в НК обладает свойствами голограммы. В каждой ее точке имеется информация обо всей Вселенной, обо всем ПВ.

Я, однако, не могу согласиться с постулатом Акимова [5] о мгновенной скорости передачи информации в ПТП. Если матрица возможного мгновенно в момент $U = 0$, в точке сингулярности 1 предопределяет все вероятности переходов и состояния всех точек ПВ Вселенной, то при $U > 0$ имеет место течение собственного времени, а значит

и существует скорость информационных взаимодействий. Вопрос может лишь стоять в определении величины этой скорости.

В структуре первичного торсионного поля уровня VI одновременно содержатся сведения обо всех свойствах элементов следующего V-го уровня - уровня физического вакуума и последующих уровней.

Рассмотрим особенности нейрокомпьютерной модели уровня ФВ.

3. ФИЗИЧЕСКИЙ ВАКУУМ (ФВ)

V-й уровень физической реальности представляет особую форму материи, так называемой физической вакуум (ФВ).

К известным нам уже свойствам шестого уровня ПТП на данном уровне добавляются еще несколько характеристик. А именно, здесь определяются параметры будущих, рождающихся на этом уровне частиц (их заряд, масса, спин и др.). Частицы еще отсутствуют. Вещества, "грубой материи" еще нет.

Этот уровень - это материальная, но не вещественная среда. Это чисто информационная полевая структура, которая порождает новые частицы и античастицы, такие, например, как (0/0), (1/0), (0/1), (1/1) и которые могут превращаться друг в друга и аннигилировать. Более подробное введение в атомную квантовую информатику имеется в книге "Вселенная Хокинга и нейрокомпьютер" [7], где НК рассматривается как квантовая система.

Если создать условие, препятствующее аннигиляции виртуальных пар частиц, которое в НК достигаются подключением интегрирующего детектора, отбирающего частицы заданного качества или заданных свойств, то рождается вещество или у нас код, структура с конкретной величиной массы, с геометрией, зарядом, спином.

В НК это осуществляется путем регистрации формы гистограммы распределения вероятностей кодов нейронов и частот их импульсации.

Происходит процесс материализации или проявления физического вакуума. Рождается наш обычный вещественный мир. Набор возможных объектов материального мира определяется матрицей.

Мы в нашем мире получаем предопределенный набор возможных информационных и физических (вещественных) атомов с их свойствами, выраженными в виде Информационного периодического закона элементов, аналогичного Периодической системе Д.И. Менделеева.

Здесь имеет место нейрокомпьютерная реализация известного волнового генома Гаряева. Регистр P играет роль хромосомного аппарата, вокруг которого генерируется волновое поле - волновой геном НК и из протоплазмы ядра - окружающего вакуума извлекаются и синтезируются молекулярные, а у нас смысловые, последовательности и структуры [8].

Другими словами, все, что сотворено на седьмом уровне АН, предопределяет все виды сочетаний материи на всех последующих уровнях ее организации, вплоть до уровня живой материи [16].

Мы здесь с Вами оказываемся перед необъятными просторами нейрокомпьютерного моделирования механизмов

мов генетики [17], эволюции, явлений клонирования и квантовой телепортации вещества, моделирования явлений квантовой механики и микромира.

Из НК модели Вселенной, процесса проявления, материализации вещества из "Абсолютного ничто" до момента $U=n$ рождается ВСЕ возможное, предопределенное матрицей возможного, множество элементов, состояний и свойств первичного торсионного поля (ПТП).

Поэтому зарождение или появление "человека" гарантировано не позже, чем через время $U \leq n$ после Большого Взрыва (точки сингулярности 1), если Вселенная едина в Метасистеме. Но, как следует из нашей нейрокомпьютерной модели [7] и данных космологов, Вселенных бесчисленное множество и моменты, точки сингулярности, могут существовать всегда, в любом интервале наблюдения.

Я считаю, в отличие от [5], последовательность или время проявления конкретного элемента таблицы Периодической системы или закона строго определены матрицей возможного. Что раньше может появиться (материализоваться) в процессе эволюции, человек, вирус или бактерия, - этот вопрос требует более тщательного изучения и дополнительного анализа.

Установлена взаимосвязь всех уровней реальности с уровнем VII АН. А.Е. Акимов [5] называет поляризацией результат взаимодействия проявленной материи с ФВ. Это взаимодействие и есть изменение пространства-времени Вселенной. Эффект такого изменения проявляется в изменении невязки $|J|$ как не равновесия (поляризации) возбуждения информационного n -атома.

$$|J| = |P| - |S|. \quad (8)$$

Невязка порождает в свою очередь конкретный набор частиц и их характеристик.

Здесь мы с Вами подходим к возможности более глубокого понимания информационной природы гравитации и электромагнетизма, т.е. природы физического, нас окружающего пространства и его связей с ФВ и к созданию единой теории поля.

Итак, по моему убеждению, с целью уточнения терминологии, целесообразно субфизику представить как ФВ, состоящий из трех компонент-уровней: АН - "Абсолютное ничто", ПТП - первичные торсионные поля и V - вакуум (эфир).

$$\Phi V = АН + ПТП + V. \quad (9)$$

Это исключит путаницу в терминах между ФВ как едином поле и физическим вакуумом уровня V .

4. О КВАНТОВОЙ ТЕЛЕПОРТАЦИИ ВЕЩЕСТВА С ПОМОЩЬЮ НЕЙРОКОМПЬЮТЕРА "ЭМБРИОН"

Обратиться к данной проблеме меня побудило знакомство с публикацией Г. Кузнецова "Телепортация посредством телепатии" [10], идеи Я.И. Корчмарюка о "переселении" личности из старого в новый, молодой организм [11] и публикация фирмы IBM Quantum Teleportation [12]

в Интернете.

Эти работы имеют прямое отношение к рассматриваемой в настоящей работе модели ФВ и тесно связаны с ЭПР-парадоксом [6].

Действительно, если пространство-время Вселенной рождено и его свойства предопределены, включая процессы материализации и дематериализации (аннигиляции) на последующих уровнях физической реальности, матрицей возможного $|W_{ij}|$, волей (NS) и полем сознания $\{U\}$, т. е.

Сверхсознанием Творца (Богом, Абсолютом) и представляет собою торсионное поле в виде голограммы взаимосвязанных и взаимодействующих элементов единой системы, то любое изменение в какой-либо точке Вселенной (ПВ-материи) проявится, породит, материализует вначале информационный, а вслед за ним и материальный, вещественный процесс синтеза или разложения вещества с заданными свойствами в заданном месте.

Если мы будем на Сенсорную Матрицу (СМ) НК вводить последовательные срезы информационного портрета личности (фантом передатчика), то в виртуальном торсионном поле-пространстве НК на приемном конце можно автоматически синтезировать вещество с характеристиками и свойствами, закодированными в виртуальном вероятностном поле НК - передатчика.

В настоящей работе нет возможности более подробно остановиться на этой очень важной и интересной проблеме, на модели квантовой передаче вещества на расстояние и я оставляю за собой право более детально осветить этот механизм информационной связи [13-17].

ПЕРЕЧЕНЬ ССЫЛОК

1. Сознание и физический мир: сборник статей /Под ред. А.Е.Акимова.- М.: Яхтсмен, 1995.
2. *Иеромах Анатолий (Берестов), Решетов В.* "Колдуны" в законе. М. Московский Патриархат. 1999.
3. *Шипов Г.И.* Теория физического вакуума. М. ИТ-Центр, 1993.
4. *Цыганков В.Д.* Нейрокомпьютер и его применение. М. Сол Систем, 1993.
5. *Акимов А.Е.* Облик физики и технологий в начале XXI века. М. Шарк, 1999.
6. *А.Эйнштейн.* Собрание сочинений т.4
7. *Цыганков В.Д.* Вселенная Хокинга и нейрокомпьютер. М. Синтег, 2000.
8. *Налимов В.В.* Спонтанность сознания. Вероятностная теория смыслов и смысловая архитектура личности. М. Прометей, 1989.
9. *Цыганков В.Д.* Нейрокомпьютер "Эмбрион" как генератор и преобразователь смыслов в вероятностной модели личности. // РИУ. № 1. 2000.
10. *Кузнецов Г.* Телепортация посредством телепатии //ж-л. Компьютера, №16, 1997.
11. *Корчмарюк Я.И.* "Сеттлеретика" - новая междисциплинарная наука о "переселении" личности. <http://www.user.cytyline.ru/-neur news/ nmatan/ settler>.
12. Quantum Teleportation. <http://www.research.ibm.com/ quantuminfo/ teleportation>.
13. *Цыганков В.Д.* Виртуальный нейрокомпьютер "Эмбрион" как квантовый вычислитель // Теория и общие вопросы обработки аналоговой информации. Труды Международной конференции. Ульяновск. 1999. стр. 42 - 51.
14. *Цыганков В.Д.* Нейрокомпьютерная квантовая Orch OR - модель сознания. // РИУ. № 2. 1999.
15. *Цыганков В.Д.* Виртуальный нейрокомпьютер "Эмбрион" как информационный лазер. // Там же.
16. *Цыганков В.Д.* Живая Вселенная. Нейрокомпьютер и мозг. М. МАИ. 1995, Нью-Йорк. (в печати).
17. *Цыганков В.Д.* О возможности применения нейрокомпьютера в генетике. Сознание и физическая реальность. М. МАИ. 1995.

ІІІ. УПРАВЛІННЯ

УДК 621.11-52

О ВЫБОРЕ ВЕСОВЫХ КОЭФФИЦИЕНТОВ ОПТИМИЗИРУЕМОГО ФУНКЦИОНАЛА В ТЕОРИИ АНАЛИТИЧЕСКОГО КОНСТРУИРОВАНИЯ ОПТИМАЛЬНЫХ РЕГУЛЯТОРОВ

И. Е. Александрова, Т. Е. Александрова

Предлагается методика выбора весовых коэффициентов аддитивного интегрального квадратичного критерия качества замкнутой системы автоматического управления при решении задачи аналитического конструирования оптимального регулятора.

Пропонується методика вибору вагових коефіцієнтів адитивного інтегрального квадратичного функціоналу якості замкненої системи автоматичного керування при рішенні задачі аналітичного конструювання оптимального регулятора.

The case of value technique of quality criterion for close-loop automatic control system for solving problem of optimal regulator analytic constriction is proposed.

ВВЕДЕНИЕ

Одной из важнейших и самых сложных проблем синтеза оптимальных регуляторов является формализация требований к динамическим характеристикам замкнутой системы и представление этих требований в виде требования минимума интегрального квадратичного функционала

$$I = \int_0^T \left[\sum_{i=1}^n \beta_i^2 \left(x_i^2(t) + \sum_{j=1}^m u_j^2(t) \right) \right] dt. \quad (1)$$

Такой функционал представляет собой меру отклонения возмущенного движения замкнутой системы от невозмущенного и может быть указано сколько угодно таких мер, каждая из которых определяет оптимальный в смысле минимума функционала (1) вектор управления $U(t)$. В данной статье рассматривается методика выбора значений весовых коэффициентов β_i ($i = \overline{1, n}$) функционала (1) при конструировании оптимального регулятора для линейного объекта

$$\dot{X}(t) = AX(t) + BU(t); U(t) \in G_u. \quad (2)$$

Предположение о линейности объекта управления не

ограничивает общности предлагаемой методики, которая может быть без изменений распространена на нелинейные объекты вида

$$\dot{X}(t) = \Phi[X(t)] + BU(t); U(t) \in G_u.$$

ПОСТАНОВКА ЗАДАЧИ

В работах [1-3] n -мерный вектор $\beta = \{\beta_i\} \in G_\beta$ рекомендуется определять из условия экстремума одного или нескольких дополнительных критериев качества (γ -критериев), однако при таком подходе к проблеме выбора основного критерия остается определенный произвол в выборе дополнительных критериев.

Проблеме выбора критериев оптимальности при конструировании оптимальных регуляторов посвящены также работы М.Е.Салуквадзе [4,5]. Предложенный в них метод выбора вектора $\beta \in G_\beta$ основан на идее синтеза парето-оптимального управления, обеспечивающего максимальную близость аддитивного функционала (1) к каждому из функционалов

$$I_i = \int_0^T \left(x_i^2(t) + \sum_{j=1}^m u_j^2(t) \right) dt. \quad (3)$$

Вместе с тем, метод М.Е.Салуквадзе не предполагает ранжировки критериев (3) по степени их важности.

В работах [6-8] рассмотрена методика выбора вектора $\beta \in G_\beta$, основанная на требовании быстрейшего убывания функции Ляпунова-Беллмана вдоль траектории замкнутой системы. Действительно, функцию Ляпунова-Беллмана можно трактовать как текущее расстояние между возмущенным состоянием замкнутой системы и невозмущенным. Динамические свойства системы тем выше, чем быстрее убывает это расстояние в переходном процессе. Вместе с тем, этот метод требует большого

объема вычислений в процессе его реализации, а также может приводить к весьма большим значениям коэффициентов усиления оптимального регулятора, соответствующих малым значениям весовых коэффициентов β_i ($i = \overline{1, n}$).

Потребуем, чтобы весовые коэффициенты функционала (1) удовлетворяли ограничению

$$\sum_{i=1}^n \beta_i = 1, \quad (4)$$

чтобы избежать тривиального решения $\beta_i = 0$ ($i = \overline{1, n}$) задачи выбора величин β_i ($i = \overline{1, n}$), а задачу аналитического конструирования оптимального регулятора (АКОР) сформулируем в виде: требуется отыскать вектор $U(t) \in G_u$, доставляющий на решениях системы (2) минимум функционалу (1) при одновременном отыскании вектора $\beta \in G_\beta$, с ограничением (4).

МЕТОД И АЛГОРИТМ РЕШЕНИЯ

С учетом обозначений (3) функционал (1) представим в следующем виде:

$$I = \sum_{i=1}^n \beta_i^2 I_i^*. \quad (5)$$

Через I_i^* обозначим минимальное значение функционала (3), получаемое в результате решения задачи АКОР для объекта (2) при минимизации только лишь функционала (3).

При фиксированных значениях β_i минимально возможное значение функционала (5) составляет

$$I = \sum_{i=1}^n \beta_i^2 I_i^*. \quad (6)$$

Отыщем минимум функционала (6) по $\beta \in G_\beta$ при условии (4). Рассматриваемая задача представляет собой задачу на условный экстремум. Для решения ее составим функцию Лагранжа

$$F(\beta) = \sum_{i=1}^n \beta_i^2 I_i^* + \lambda \left(1 - \sum_{i=1}^n \beta_i \right), \quad (7)$$

где λ - множитель Лагранжа. Для отыскания минимума функции (7) приравняем нулю производные по компонентам вектора β

$$\frac{\partial}{\partial \beta_i} F(\beta) = 2\beta_i I_i^* - \lambda = 0; \quad (i = \overline{1, n}). \quad (8)$$

Из соотношений (8) получаем

$$\beta_i = \frac{\lambda}{2I_i^*}, \quad (i = \overline{1, n}). \quad (9)$$

Вектор β с компонентами (9) доставляет минимум функции Лагранжа (7), поскольку гессиан

$$\frac{\partial^2}{\partial \beta \partial \beta} F(\beta) = 2 \begin{vmatrix} I_1^* & 0 & \dots & 0 \\ 0 & I_2^* & \dots & 0 \\ \dots & \dots & \dots & \dots \\ 0 & 0 & \dots & I_n^* \end{vmatrix}$$

в силу положительности I_i^* ($i = \overline{1, n}$), является положительно-определенной матрицей.

Подставляя соотношения (9) в формулу (4), получим выражение для множителя Лагранжа

$$\lambda = \frac{2}{n \sum_{i=1}^n \frac{1}{I_i^*}}. \quad (10)$$

Подставив (10) в (9), получим искомые значения компонент вектора β

$$\beta_i = \frac{1}{n \sum_{i=1}^n \frac{1}{I_i^*}}, \quad (i = \overline{1, n}). \quad (11)$$

ПРИМЕР

Для объекта, возмущенное движение которого описывается уравнениями [9]

$$\dot{x}_1(t) = x_2(t), \quad \dot{x}_2(t) = u(t), \quad (12)$$

требуется отыскать управление $u(t)$, доставляющее минимум интегральному квадратичному функционалу

$$I = \int_0^T [\beta_1^2 x_1^2(t) + \beta_2^2 x_2^2(t) + (\beta_1^2 + \beta_2^2) u^2(t)] dt, \quad (13)$$

где весовые коэффициенты β_1 и β_2 удовлетворяют условию

$$\beta_1 + \beta_2 = 1$$

и подлежат выбору.

В соответствии с вышеизложенным, функционал (13)

представим в виде суммы (5), причем

$$I_1 = \int_0^T [x_1^2(t) + u^2(t)] dt, \quad (14)$$

$$I_2 = \int_0^T [x_2^2(t) + u^2(t)] dt. \quad (15)$$

Минимальные значения функционалов (14) и (15) равны [10]:

$$I_1^* = \langle X(0), K_1 X(0) \rangle; I_2^* = \langle X(0), K_2 X(0) \rangle, \quad (16)$$

где $X(0)$ - вектор состояния объекта управления в момент $t = 0$, а квадратные симметрические матрицы K_1 и K_2 удовлетворяют матричным уравнениям

$$Q_1 + K_1 A + A^T K_1 - K_1 B B^T K_1 = 0; \quad (17)$$

$$Q_2 + K_2 A + A^T K_2 - K_2 B B^T K_2 = 0. \quad (18)$$

Матрицы, входящие в уравнения (17) и (18), равны

$$Q_1 = \begin{bmatrix} 1 & 0 \\ 0 & 0 \end{bmatrix}; Q_2 = \begin{bmatrix} 0 & 0 \\ 0 & 1 \end{bmatrix}; A = \begin{bmatrix} 0 & 1 \\ 0 & 0 \end{bmatrix}; B = \begin{bmatrix} 0 \\ 1 \end{bmatrix}.$$

Решения уравнений (17) и (18) записываются в виде

$$K_1 = \begin{bmatrix} \sqrt{2} & 1 \\ 1 & \sqrt{2} \end{bmatrix}; K_2 = \begin{bmatrix} 0 & 0 \\ 0 & 1 \end{bmatrix},$$

а значения функционалов (16) составляют

$$I_1^* = \sqrt{2}x_1^2(0) + 2x_1(0)x_2(0) + \sqrt{2}x_2^2(0), I_2^* = x_2^2(0). \quad (19)$$

В соответствии с принципом минимакса [7], в формулах (19) следует принять $x_1(0) = x_{1\max}$; $x_2(0) = x_{2\max}$.

Положим $x_{1\max} = x_{2\max} = 1$. В результате получаем

$$I_1^* = 2(1 + \sqrt{2}), I_2^* = 1.$$

Подставим полученные значения I_1^* и I_2^* в (11). В результате получим искомые значения весовых коэффициентов функционала (13)

$$\beta_1 = \frac{1}{3 + 2\sqrt{2}}; \beta_2 = \frac{2 + 2\sqrt{2}}{3 + 2\sqrt{2}}.$$

Управление $U(t)$, доставляющее минимум функционалу (13) на решениях системы (12), в соответствии с общей

теорией АКОР [9] записывается в виде

$$U(t) = -\frac{1}{\beta_1^2 + \beta_2^2} B^T K X(t), \quad (20)$$

где матрица K удовлетворяет матричному уравнению

$$Q + KA + A^T K - K B B^T K \cdot \frac{1}{\beta_1^2 + \beta_2^2} = 0, \quad (21)$$

а матрица Q равна

$$Q = \begin{bmatrix} \beta_1^2 & 0 \\ 0 & \beta_2^2 \end{bmatrix}.$$

Решение уравнения (21) приводит к следующей матрице K :

$$K = \begin{bmatrix} K_{11} & K_{12} \\ K_{12} & K_{22} \end{bmatrix}, \quad (22)$$

где $K_{11} = \beta_1 \sqrt{\beta_2^2 + 2\beta_1 \sqrt{\beta_1^2 + \beta_2^2}}$; $K_{12} = \beta_1 \sqrt{\beta_1^2 + \beta_2^2}$;

$$K_{22} = \sqrt{\beta_1^2 + \beta_2^2} \sqrt{\beta_2^2 + 2\beta_1 \sqrt{\beta_1^2 + \beta_2^2}}.$$

Подставляя матрицу (22) в соотношение (20), получаем формулу для оптимального управления

$$U(t) = -\frac{1}{\sqrt{\beta_1^2 + \beta_2^2}} [\beta_1 x_1(t) + \sqrt{\beta_2^2 + 2\beta_1 \sqrt{\beta_1^2 + \beta_2^2}} \cdot x_2(t)].$$

ВЫВОД

Предложена методика выбора весовых коэффициентов аддитивного функционала качества в задаче аналитического конструирования оптимальных регуляторов.

ПЕРЕЧЕНЬ ССЫЛОК

1. Летов А.М. Динамика полета и управление.-М.: Наука, 1969.-312с.
2. Янушевский Р.Т. Теория линейных оптимальных многосвязных систем управления. -М.: Наука, 1973.-464с.
3. Деруссо П., Рой Р., Клоуз Ч. Пространство состояний в теории управления. -М.: Наука, 1970.-484с.
4. Салуквадзе М.Е. Об оптимизации векторных функционалов // Автоматика и телемеханика. -1971.-№9.-С.5-15.
5. Салуквадзе М.Е. Задачи векторной оптимизации в теории управления. - Тбилиси, 1975.-201с.
6. Аврамов В.П., Александров Е.Е. О выборе показателя качества в теории аналитического конструирования оптимальных регуляторов //Изв. вузов. Приборостроение.- 1976.-№10.-С.32-36.
7. Александров Е. Е., Бех М.В. Автоматизированное проектирование динамических систем с помощью функций Ляпунова. - Харьков: Основа, 1993.-113с.
8. Александров Е.Е., Кузнецов Б.И., Тернюк Н.Э. и др. Оптимизация электромеханических систем с упругими элементами. -К.: ВИПОЛ, 1995.- 288с.
9. Ройтенберг Я.Н. Автоматическое управление. - М.: Наука, 1978. - 551с.
10. Александров Е. Е., Кузнецов Б. И., Богаенко И. Н. и др. Многоканальные системы оптимального управления.- К.: Техніка, 1995.-228с.

СОСТОЯНИЕ, ТЕНДЕНЦИИ И ПРОБЛЕМЫ В ОБЛАСТИ МЕТОДОВ УПРАВЛЕНИЯ АСИНХРОННЫМИ ДВИГАТЕЛЯМИ

В. Г. Бичай, Д. М. Пиза, Е. Е. Потапенко, Е. М. Потапенко

Сделан обзор доступной литературы в области современных методов управления асинхронным электроприводом, опубликованной за период с 1994г. по 2000г. включительно. Всего было изучено более двухсот публикаций, 102 наиболее интересные из которых включены в обзор. Это позволило определить состояние, тенденции и проблемы в области управления асинхронным электроприводом.

Зроблено огляд доступної літератури в галузі сучасних методів керування асинхронним електроприводом, яка опублікована за термін з 1994р. до 2000р. включно. Усього було вивчено більш двохсот публікацій, 102 найбільш цікавих з яких включено до огляду. Це дозволило з'ясувати стан, тенденції та проблеми в галузі керування асинхронним електроприводом.

The survey of the accessible literature in the field of the modern control methods of induction drives which was published during from 1994 to 2000 year was done. It was studied more than 200 publications, 102 the most interesting from which were included in the survey. This allowed to determine the state, the tendencies and the problems in the field of the modern control of induction drives.

В настоящее время более 60% всей вырабатываемой в мире электроэнергии потребляется электродвигателями [1,2]. Наибольшее распространение получили электроприводы (ЭП) с двигателями постоянного тока (ДПТ) и асинхронными двигателями с короткозамкнутым ротором (АД). Причем около половины вырабатываемой в мире электроэнергии потребляется АД [3]. Указанные соотношения справедливы и для Украины. ДПТ применяются для регулируемых ЭП, АД - в основном, для нерегулируемых ЭП (около 80% АД - нерегулируемые [1]). Это объясняется, с одной стороны, хорошими регулировочными характеристиками ДПТ и наличием коллектора, снижающего его надежность и не позволяющего эксплуатировать ДПТ в агрессивных и взрывоопасных средах. С другой стороны, простотой конструкции, дешевизной, высокой надежностью, но плохими регулировочными характеристиками АД.

Большинство технологических процессов требует регулирования. Это регулирование при нерегулируемом ЭП осуществляется путем управляемого изменения в самом технологическом процессе. В частности, в насосах, компрессорах, вентиляторах - путем открытия и закрытия заслонок. Показано [4-6], что путем отказа от регулирования заслонками и путем перехода к регулированию АД можно достигнуть экономии (50...68)% по сравнению с нерегулируемыми АД. В связи с таким количеством эксплуатируемых АД, которые требуют регулирования, весьма актуальным является вопрос повы-

шения экономичности регулируемых АД.

Применение регулируемых АД до начала 70-х годов сдерживалось отсутствием тиристорных с малым временем выключения. В начале 70-х годов появились специальные инверторные тиристоры с требуемыми характеристиками, что обусловило появление частотно-регулируемых ЭП с АД. Однако удельные массогабаритные показатели и стоимость ЭП с АД были весьма велики из-за преобразователей напряжения, а по показателям статических и динамических характеристик ЭП с АД существенно уступали ЭП с ДПТ [7]. Последнее десятилетие XX-го века ознаменовалось значительными успехами силовой электроники. Было освоено промышленное производство биполярных транзисторов с изолированным затвором (IGBT), силовых модулей на их основе и др. средств. Все это стимулировало разработку новых алгоритмов управления и средств их вычисления.

В 1971 г. Ф.Блашке [8-10] предложил новый принцип управления АД. В соответствии с этим принципом управления движения рассматриваются не в стационарной (кларковской) системе координат, а в системе координат Парка-Горева, связанной с потокосцеплением ротора. В этой вращающейся системе координат все основные переменные изменяются не по гармоническому закону, а как в ДПТ. Это позволило строить системы управления АД по тем же принципам, что и системы управления ДПТ, в частности, по подчиненному принципу. Этот принцип управления АД получил название Field Oriented Control (FOC) (управление по полю машины). В этом принципе управления все многофазные (чаще трехфазные) переменные (токи, напряжения, потокосцепление) формируют соответствующие векторы, поведение которых рассматривается или в векторном виде, или в проекциях на две оси вращающейся системы координат. Поэтому в русскоязычной литературе этот принцип управления получил название "Векторное управление" (ВУ). Векторное управление теоретически позволяло достигнуть в АД тех же регулировочных характеристик, что и в ДПТ. Однако для реализации векторного управления надо иметь информацию о всем векторе состояния АД. В 70-х годах это предлагалось достигать за счет установки внутри АД датчиков Холла, с помощью которых замерять потокосцепление в воздушном зазоре АД, и тахогенераторов или импульсных датчиков, посредством которых измеряется угловая скорость ротора. Однако установка этих датчиков требовала замены или

доработки огромного парка общепромышленных АД, которые не предназначались для регулирования и не были снабжены этими датчиками. Наличие тахогенератора и датчиков Холла существенно ухудшали эксплуатационные характеристики АД по сравнению с ДПТ.

К началу 80-х годов были достигнуты большие успехи в области современной теории управления, которая позволяет оценить неизвестные составляющие вектора состояния. Применение методов современной теории управления позволило, в первую очередь, без специальных датчиков оценить вектор потокосцепления. После этого векторное управление с измерением вектора потокосцепления с помощью датчиков получило название прямого векторного управления (ПВУ) (Direct FOC (DFOC)), а без применения датчиков потокосцепления - косвенное векторное управление (КВУ) (Indirect FOC (IFOC)) [11-13]. Однако в определении ПВУ (DFOC) и КВУ (IFOC) нет единообразия. Некоторые авторы [14,15] под ПВУ понимают управление с непосредственным использованием векторов потокосцепления, полученных с помощью наблюдателей, а под КВУ понимают управление без явного использования векторов потокосцепления. Поскольку системы с непосредственным измерением с помощью датчиков векторов потокосцепления не получили распространения и не имеют перспективы, то в дальнейшем будет использоваться терминология авторов работ [14,15].

Как в ПВУ, так и в КВУ отсутствие датчиков потокосцепления компенсируется вычислениями по дифференциальным уравнениям, описывающим работу АД, в которые входят неточно известные и переменные параметры, такие как сопротивления ротора и статора, приведенный момент инерции ротора, момент нагрузки. Собственные индуктивности ротора и статора, а также их взаимная индуктивность вследствие эффекта магнитного насыщения имеют сложную зависимость от тока намагничивания. В связи с этим системы ВУ без специальных мероприятий являются очень чувствительными к неопределенности этих параметров. К счастью, в настоящее время современная теория управления обладает мощным арсеналом методов обеспечения робастности (нечувствительности) систем управления к разбросу их параметров. Это адаптивные системы, системы с переменной структурой, робастные комбинированные системы с наблюдателями, нейронные системы, системы с размытой логикой и др. В обзоре будут упомянуты только работы с ПВУ и КВУ с мероприятиями по обеспечению робастности.

Для обеспечения питания АД в настоящее время используются инверторы напряжения с широтно-импульсной модуляцией (ШИМ), которые создают большие высокочастотные пульсации в напряжении. Пульсации не позволяют использовать непосредственные измерения напряжения при его цифровой обработке в контроллере привода. Поэтому, прежде чем подать измеренное напря-

жение в цифровую систему, его пропускают через аналоговый фильтр [19,37]. Для исключения датчиков напряжения и снятия проблемы высокочастотных помех в работах [19,20,21,78] вместо истинного напряжения, подаваемого на АД, используется его программное значение, вырабатываемое микропроцессором и подаваемое на преобразователь напряжения. Кроме того, информация о напряжении и токе статора может быть получена по напряжению и току звена постоянного тока инверторного преобразователя [22,47,56,80]. Таким образом, для управления АД можно использовать лишь датчик скорости и датчики токов статора. Более того, ниже будут рассмотрены системы управления АД, использующие только датчики тока статора.

Вопросам управления АД с использованием датчика скорости посвящено огромное количество работ. Поэтому остановимся лишь на системах управления с использованием датчиков скорости, в которых применяются специальные мероприятия по обеспечению робастности [7,12,17-44].

Недостатком КВУ является сильная зависимость его показателей качества от постоянной времени роторной цепи (T_r). Данное обстоятельство и неизвестность потокосцепления приводят к затягиванию переходных процессов. Это обусловило появление работ [12,19,21,26,30,31,38] по непосредственной оценке постоянной времени или сопротивления роторной цепи (R_r) без оценки потокосцепления ротора. В работе [30] в КВУ для оценки T_r использована процедура, дающая в качестве побочного продукта оценку потокосцепления ротора. При наличии оценки потокосцепления ротора представляется более рациональным использовать не КВУ, а ПВУ.

Адаптивной оценке потокосцепления ротора и ее использованию для ПВУ посвящено большое количество работ [7,14,21,24,25,27,34-36,39,41-44] (см. также биографию данных работ).

Основным источником погрешности ПВУ является неточность определения опорного вектора (вектора потокосцепления ротора) в статорной системе координат. Для устранения этой неточности широко предлагается использовать адаптивные наблюдатели типа MRAS (Model reference adaptive system - адаптивная система с настраиваемой моделью) [22,27,29,34-36,78], генетические алгоритмы [39], нейронные сети [40-42], размытую логику [40,43,44], наблюдатели переменной структуры [60,79]. Для устранения этой же неточности в работе [28] предложено использовать в качестве синхронной системы координат систему координат, связанную не с вектором потокосцепления ротора, а с оценкой вектора потокосцепления ротора с помощью наблюдателя. В работе [20] для этой же цели в качестве синхронной системы координат используется система координат, связанная с вектором статорного тока.

Существуют три физические основы для определения

скорости ротора: 1) использование тахогенераторов или инкрементных датчиков (последние проще и лучше приспособлены к цифровому управлению, 2) использование против э.д.с., 3) обработка показаний датчиков тока с целью получения частоты пазовых пульсаций. Применение датчиков скорости вращения ротора позволяет организовать высококачественное и алгоритмически сравнительно простое управление АД. Однако наличие датчиков скорости существенно ухудшает эксплуатационные характеристики электропривода. В связи с этим в последнее десятилетие уделялось большое внимание разработке алгоритмов векторного управления с использованием лишь датчиков тока и напряжения в сети питания АД или, вообще, с использованием лишь датчиков тока [45-60,71,80,100,102]. Эта задача намного сложнее по сравнению с управлением с использованием датчиков скорости. Сравнительно простой метод вычисления угловой скорости можно получить на основании [81] с использованием ЭДС АД. Вследствие недостаточной точности простых алгоритмов были разработаны более точные, но более сложные, методы. В них задача решается с помощью различных типов наблюдателей. В работах [45,54,60] используются наблюдатели на скользящих режимах. В работе [49] применен наблюдатель потокосцепления статора, с помощью которого можно вычислить скорость ротора. Широкое распространение для оценки потокосцепления и скорости ротора получили системы типа MRAS [46,48,50,56], где используются два векторных уравнения потокосцепления ротора, одно из которых является базовой, а второе - настраиваемой моделями. Близкой к MRAS является система, нашедшая применение в Японии [51]. По сути дела, к системам типа MRAS примыкают адаптивные наблюдатели потокосцепления и скорости [48,52,55], где базовой моделью является сам АД, а настраиваемой моделью является наблюдатель тока статора и потокосцепления ротора. В работе [55] для адаптации используется сравнение выходов двух одновременно настраиваемых наблюдателей. Хорошие результаты по оценке скорости были получены с помощью расширенного фильтра Калмана [48, 52]. Однако они связаны со сложными вычислениями. Все перечисленные методы в той или иной степени чувствительны к разбросу параметров электромагнитной цепи АД. В этом отношении представляет интерес определение скорости по роторным пазовым пульсациям тока статора [48], который, к сожалению, имеет свои недостатки. Предприняты попытки для построения систем управления АД без датчиков скорости с помощью нейронных сетей [46,52]. Все рассмотренные методы оперируют уравнениями в статорной системе координат. Исключением составляют работы [31,46,59], где восстановление вектора состояния осуществляется во вращающейся системе координат, связанной с потокосцеплением ротора.

Для обеспечения робастности систем управления АД и диагностики их работоспособности в публикациях полу-

чил распространение регрессионный рекуррентный метод наименьших квадратов [61-66], который позволяет в реальном времени оценивать параметры АД. Более того, в работе [66] получен алгоритм, который, помимо идентификации всех параметров АД, позволяет вычислить угловую скорость его ротора, используя только датчики тока и напряжения статора. Это открывает пути получения высокоточного адаптивного управления АД с помощью измерения только клеммных переменных (напряжений и токов).

Недостатком рассмотренных выше алгоритмов векторного управления являются: 1) большой объем вычислений при прямом и обратном взаимном преобразовании неподвижной и вращающейся систем координат, 2) наличие запаздывания в формировании электромагнитного момента. Для устранения этих недостатков были предложены [67,68] алгоритмы управления, получившие название Direct Torque Control (DTC) (непосредственное управление моментом и потокосцеплением). Эти идеи в последнее время получили бурное развитие [7,11,69-77,100]. Алгоритмы DTC отличаются простотой (нет преобразования координат и регулирования токовых контуров), робастностью по отношению к неопределенности параметров, высоким быстродействием. Недостатком является наличие пульсаций в электромагнитном моменте и потокосцеплении, что снижает точность регулирования, повышает электропотребление и увеличивает акустический шум АД. Это ставит новую проблему в разработке алгоритмов управления.

Как было указано в начале статьи, существенным качеством электродвигателя является его экономичность. Вопросу повышения экономичности АД при скалярном управлении уделено очень большое внимание [16,81-88]. Оптимизацию энергопотребления можно проводить путем формирования фазовой траектории и реализующего ее управления, при перемещении по которой между двумя точками минимизируется потребляемая электроэнергия [85-89]. Такое управление называется оптимальным в большом. При таком управлении электроприводом минимизируется потребляемая за определенный промежуток времени электроэнергия. Другим путем оптимизации является минимизация потребляемой мощности или какой-либо ее составляющей в каждой точке заданной траектории. Такое управление называется локально оптимальным [81-84].

Оптимальные алгоритмы управления, полученные для скалярного управления, трудно перевести на язык векторного управления. Это объясняется тем, что оба принципа управления оперируют разными переменными, сложным образом связанными между собой. Кроме того, для упрощения алгоритмов векторного управления в подавляющем большинстве случаев намагничивающую составляющую статорного тока делают постоянной [93], что устраняет необходимую для оптимизации степень свободы и не позволяет осуществлять локальную опти-

мизацию. Поэтому локальной оптимизации по мощности векторного управления посвящено мало работ [27,90-96,101,102]. В работе [90] сформулированы условия, при которых можно осуществить оптимизацию. (Не со всеми условия можно согласиться). В [91] предложен экстремальный поисковый метод путем непосредственного замера потребляемой мощности, что приводит к ухудшению точности управления. Работы [92,93] посвящены максимизации электромагнитного момента с учетом ограничений по току и напряжению, которые, однако, не дают аналитического решения. В статье [27] получено простое аналитическое выражение программного изменения потокосцепления ротора в зависимости от электромагнитного момента. В работах [94-96] синтезированы аналитические программные значения намагничивающей и моментной составляющих тока статора, минимизирующие различные сочетания видов потерь мощности. В отличие от [27], полученные выражения справедливы не только для стационарных, а и для переходных режимов. Кроме того, они, в отличие от [27], робастны по отношению к неточности знания статора индуктивностей и сопротивления ротора.

Проведенный анализ показывает, что предложено большое количество разнообразных алгоритмов управления, обладающих различными качествами. Выбор того или иного алгоритма управления должен осуществляться исходя из следующих требований:

- 1) диапазон регулирования скорости,
- 2) точность обработки программных значений скорости,
- 3) работоспособность при малых скоростях,
- 4) время переходного процесса при подаче ступенчатого задающего воздействия,
- 5) робастность (нечувствительность) по отношению к неопределенностям параметров и внешних воздействий,
- 6) минимизация электропотребления как в переходных, так и в установившихся режимах при различных нагрузках,
- 7) минимальное влияние на точность наброса и сброса нагрузки, как при больших, так и при малых скоростях,
- 8) быстрота готовности к работе после включения двигателя,
- 9) минимальное количество датчиков,
- 10) минимальность объема вычислений,
- 11) способность восстанавливать свою работоспособность после сбоев в системе или прекращения подачи питания без повторного перезапуска системы после останова ротора.

Сравнение различных алгоритмов управления очень трудоемко. Однако по п. 10 можно дать рекомендации. Поскольку неизвестен весь вектор состояния системы и значения ряда его параметров, то алгоритмы управления содержат наблюдатели, представляющие собой системы дифференциальных (в микропроцессоре - разностных) уравнений. Некоторые звенья регулятора также описываются дифференциальными уравнениями. Известно,

что объем вычислений пропорционален кубу порядка решаемой системы дифференциальных уравнений. Поэтому надо стремиться к уменьшению порядка системы дифференциальных уравнений регулятора и наблюдателя. Кроме того, при векторном управлении осуществляется неоднократное преобразование систем координат (неподвижной в подвижную и наоборот по различным переменным). Эти преобразования используют синусные и косинусные функции, вследствие чего требуют большого объема вычислений. Поэтому надо стремиться уменьшить количество преобразований координат.

При выборе алгоритмов управления следует принимать во внимание, что робастные системы (системы с переменной структурой, комбинированные системы с наблюдателями) имеют большее быстродействие, чем адаптивные системы.

Сделанный обзор литературы позволяет наметить следующие направления исследований:

- 1) учет влияния ограничений по току и напряжению, а также насыщения магнитной цепи на работоспособность систем управления,
- 2) получение простых алгоритмов управления, минимизирующих электропотребление с учетом ограничений по току, напряжению и при наличии насыщения магнитной цепи,
- 3) обеспечение робастности косвенных методов определения скорости ротора (повышение их точности) по отношению к разбросу параметров системы,
- 4) уменьшение нижней границы диапазона скоростей ротора АД, при которых система без датчика скорости сохраняет работоспособность,
- 5) разработка и исследование дискретных алгоритмов управления,
- 6) разработка и исследование методов идентификации параметров АД для диагностики их состояния и выявления неисправностей,
- 7) разработка и исследование алгоритмов управления в системах с непосредственным управлением моментом (DTC), минимизирующих пульсации электромагнитного момента и токов статора,
- 8) аналитическое доказательство робастной устойчивости систем управления, гарантирующее их работоспособность в заданных условиях эксплуатации,
- 9) сравнение известных алгоритмов управления и определение предпочтительных областей применения каждой из них.

Реализация последнего пункта не только трудоемка из-за большого количества предложенных систем управления АД, но и требует широкого кругозора исполнителей в области современных систем и методов управления, таких как системы с переменной структурой, комбинированные системы с наблюдателями неопределенности, адаптивные системы, нейронные сети, размытая логика, методы идентификации и оптимального управления, теория фильтрации.

Следует отметить, что классические методы проектирования систем управления электроприводом для рассмотренных в обзоре систем не применимы. Все это выдвигает требование пересмотра учебного плана для подготовки магистров по специальности 8.092203 - электромеханические системы автоматизации и электропривод с целью углубленного изучения современной теории управления (1 семестр) и современных методов управления электроприводом (1 семестр).

Авторы благодарны Пересаде С.М. за сотрудничество в "добывании" использованной в обзоре литературы.

ПЕРЕЧЕНЬ ССЫЛОК

1. Козаченко В. Ф. Основные тенденции развития встроенных систем управления двигателями и требования к микроконтроллерам. // Chip News. - 1999. - №1(34). - С.2-9.
2. Браславский И. Я. О возможностях энергосбережения при использовании регулируемых регулируемых асинхронных электроприводов. // Электротехника. - 1998. - №8. - С.2-6.
3. Хашимов А. А. Энергосберегающие системы автоматизированного электропривода переменного тока // Электротехника. - 1995. - №11. - С.34-39.
4. Браславский И. Я., Ишматов З. Ш., Барац Е. И. Принципы построения микропроцессорной системы управления частотно-регулируемым асинхронным электроприводом насоса. // Электротехника. - 1998. - №8. - С.6-10.
5. Никитин В.М., Поздеев А.Д., Ковалев Ф.И., Шестоперов Г.Н. Энергосберегающие электроприводы. // Электротехника. - 1996. - №4. - С.52-55.
6. Попович Н. Г., Печник Н. В. Электромеханические системы автоматизации и задача энергосбережения. // Вісник ХДПУ. Збірка наукових праць. Тематичний випуск 113. - Харків, ХДПУ, 2000. - С.297-300.
7. Дацковский Л.Х., Роговой В.И., Абрамов Б.И., Моцохейн Б.И., Жижин С.П. Современное состояние и тенденции в частотно-регулируемом электроприводе (Краткий аналитический обзор). // Электротехника. - 1996. - №10. - С.18-28.
8. Blaschke F. Das Prinzip der Feldorientierung die Grundlage für die TRANSVECTOR - Regelung von Asynchronmaschinen. // Siemens-Zeitschrift. - 1971. - 45. - P.757.
9. Blaschke F. The principle of field orientation applied to the new transvector closed-loop control system for rotating field machines. // Siemens - Rev. - 1972. - 39. - P.217-220.
10. Рудаков В. В., Столяров И. М., Дартау В. А. Асинхронные электроприводы с векторным управлением. - Л.: Энергоатомиздат, 1987. - 136с.
11. Jezernik K. Robust Direct Torque and Flux Vector Control of Induction motor // In Proc. IECON '98. Germany, Sept. - 1998. - V.2. - P.667-672.
12. Valdeñebro L. R., Edson B. Fuzzy Optimization for Rotor Time Constant Identification of an Indirect Vector Controlled Induction Motor Drive // In Proc. IECON '99. Slovenia, Bled, Sept. - 1999. - P. 504-509.
13. Kleinhans C. E., Diana G., Harley R. G. and al. Analysing a CSI - Fed Field Oriented Controlled Induction Motor Using a New Simulation Package CASED // In Proc. IECON '94, Sept. - 1994. - V.1. - P.192-197.
14. Damiano A., Gatto G., Marongiu I., Pisano A. Synthesis and Digital Implementation of Reduced Order Rotor Flux Observer for IM Drive // In Proc. IECON '99. Slovenia, Sept. - 1999. - P.729-734.
15. Пересада С. М. Обобщенная теория косвенного векторного управления асинхронным электродвигателем. Часть 1. Проблема векторного управления в асинхронном электроприводе: краткий обзор и формулировка проблемы // Техническая электродинамика. - 1999. - №3. - С.27-31.
16. Шрейнер Р. Т., Поляков В. А. Адаптивная система векторного управления асинхронным электроприводом с ориентацией поля ротора // Электротехника. - 1998. - №2. - С.23-29.
17. Clerc G., Chouiter D., Aurioi P., Retif J. M. H_{∞} Controllers Design for Robust Field Oriented Asynchronous Machines in Distributed Environment // Preprints of Conference on Control of Industrial Systems. - Belfort, France, May, 1997. - V3/3. - P.114-119.
18. Millet C., Leroux D., Li Y., Feuvrie B., Bergmann C. Identification for a Field-Oriented Control Design of an Asynchronous Machine // Preprints of Conference on Control of Industrial Systems. Belfort, France, May, 1997. - V.1/3. - P.718-723.
19. Minh Ta-Cao, Hoahg Le-Huy. Rotor Resistance Estimation Using Fuzzy Logic for High Performance Induction Motor Drives // In Proc. IECON '98. Germany, Sept., 1998. - 1998. - P.303-308.
20. Krzeminski Z. Application of Observer System to Nonlinearly Controlled Induction Motor Fed by Voltage Source Inverter // In Proc. IECON '94. Italy, Bologna. Sept., 1994. - 1994. - V.1 - P.311-315.
21. Gyu-Sik Kim, Jae-Yoon Kim, Ju-Yeop Choi, Ick Choy. Recursive Rotor Resistance Adaptation Algorithm for Induction Motor Control // In Proc. IECON '98. Germany, Sept., 1998. - 1998. - P.883-887.
22. Kozo Ide, Zhi-Guo Bai, Zi-Jiang Yang and Teruo Tsuji. Torque Control of Induction Machine by Vector Approximation with Parameter Adaptation Based on MRAS // In Proc. IECON '94. Italy, Bologna. Sept., 1994. - 1994. - V.1 - P.281-286.
23. Пересада С. М. Робастное управление в следящем асинхронном электроприводе // Проблемы автоматизированного электропривода. Теория и практика: Вестник ХГПУ. Специальный выпуск. - Харьков: ХГПУ, 1998. - С.117-120.
24. R. Marino, S. Peresada, P. Tomei. On-Line Stator and Rotor Resistance Estimation for Induction Motors // IEEE Trans. Control Syst. Technol. - 2000. - V.8, №3. - P.570-579.
25. R. Marino, S. Peresada, P. Tomei. Global adaptive output feedback control of induction motors with uncertain rotor resistance // IEEE Trans. Automatic Control. - 1999. - V.44, №5. - P.967-983.
26. Kanokvate Tungpimolrut, Fang-Zheng Peng and Tadashi Fukao. A Robust Rotor Time Constant Estimation Method for Vector Control of Induction Motor Under Any Operating Conditions // In Proc. IECON '94. Italy, Bologna. Sept. 1994. - 1994. - V.1. - P.275-280.
27. Tetsuo Yamada, Kouki Matsuse, Masahiko Tsukakoshi and Lippei Huang. A Determining Method of Flux Level for Deadbeat Flux Level Controlled Direct-Field-Oriented Induction Motors Using Adaptive Flux Observer // In the same Proc. - P.287-292.
28. H. K. Khalil and E. G. Strangas. Robust Speed Control of Induction Motors Using Position and Current Measurements // IEEE Trans. on Automatic Control. - 1996. - V.41, №8. - P.1216-1220.
29. Wen-Jieh Wang, Chun-Chieh Wang. A New Composite Adaptive Speed Controller for Induction Motor Based on Feedback Linearization // IEEE Trans. on Energy Conversion. - 1998. - V.13, №1. - P.1-6.
30. L. R. Valdeñebro, Edson Bim. Fuzzy Optimization for Rotor Time Constant Identification of an Indirect Vector-Controlled Induction Motor Drive // In Proc. IECON '99. Slovenia, Bled. Sept. 1999. - 1999. - P. 504-509.
31. K. Mueller, Efficient TR Estimation in Field Coordinates for Induction Motors // In the same Proc. - P.735-741.
32. S. I. Selem Jr., A. do Prado. Discrete Time State Observer and Torque Estimator of Induction Motors with Parameter Perturbation Rejection // Preprints of NOLCOS '98. The Netherlands, Enschede. July, 1998. - 1998. - V.3/3. - P.720-725.
33. A. S. Bazanella and R. Reginatto. Robustness Margins for Indirect Field-Oriented Control of Induction Motors // IEEE Trans. Automat. Contr. - 2000. - V. 45, №6. - P.1226-1231.
34. R. Marino, S. Peresada, P. Tomei. On-Line Rotor Resistance Estimation for Induction Motors // In Proc. IECON '94. - 1994. - V.3. - P. 2137-2142.
35. C. Attaianesi, A. Damiano, G. Gatto, I. Marongiu, A. Perfetto. Induction Motor Drive Parameters Identification // IEEE Trans. on Power Electronics. - 1998. - V.13, №6. - P.1112-1121.
36. Аттаианесе Ч., Дамиано А., Марони И., Перфетто А. Управление асинхронным двигателем с адаптацией с изменяющейся электромагнитной постоянной времени ротора // Электротехника. - 1996. - №7. - С.29-31.
37. L. F. A. Pereira, J. F. Haffner, E. M. Hemerly, H. A. Grundling. A Simulation Framework for Flux Estimation and Vector Control of Induction Machines // In Proc. IECON '98. - 1998. - P.1587-1591.
38. A. Ba-razzouk, A. Cheriti and V. Rajagopalan. Real Time Implementation of a Rotor Time-Constant Online Estimation Scheme // In Proc. IECON '99. - 1999. - PE-15.
39. G. Franceschini, A. Piazzi, C. Tassoni. A Genetic Algorithm Approach to Design Flux Observer for Induction Servo Motors // In Proc. IECON '94. - 1994. - V.3. - P.2132-2136.
40. T. C. Chen, C. Y. Liaw. Design of a Neural Fuzzy Controller for Induction Motor Speed Control // In Proc. IECON '94. - 1994. - V.1. - P.611-616.
41. P. Marino, M. Milano and F. Vasca. Linear Quadratic State Feedback and Robust Neural Network Estimator for Field-Oriented-Controlled Induction Motors // IEEE Trans. on Indust. Electron. - 1999. - V.46, №1. - P.150-161.
42. Lei Wang, Guo-xing Zhou, Yun-shi Xiao, Qi-di Wu. Hopfield Neural Network Based Identification and Control of Induction Motor Drive System // In Proc. of the 14th IFAC World Congr., 1999. - 1999. - Paper № I - 3b - 03 - 4.
43. C. J. Lopez-Toribio, R. J. Patton, S. Daley. Supervisory Takagi-Sugeno Fuzzy Fault-Tolerant Control of a Rail Traction System /

- / In Proc. of the 14th IFAC World Congr, 1999. - 1999. - Paper № P - 7e - 03 - 4.
44. Xi Fu, Ying Li, Xihua Li. The Self-Adaptive SMVS-F Control System for Induction Motor // In Proc. of the 14th IFAC World Congr., 1999. - Paper № I - 3b - 01 - 04.
45. Изосимов Д. Б. Многосвязный нелинейный идентификатор состояния асинхронного двигателя на скользящих режимах // Проблемы управления многосвязными системами. - М.: Наука, 1983. - С.133-139.
46. D. Schröder, C. Schaffner and U. Lenz. Neural-Net Based Observes for Sensorless Drives // In Proc. IECON '94. - 1994. - V.3. - P.1599-1610.
47. Ho-Sun Yoo, In-Joong Ha, Sang-Hoon Lee, Sung-Jung Yoon, Yo Ko. A Polar Coordinate-Oriented Method of Identifying Rotor Flux and Speed of Induction Motors without Rotational Transducers // In Proc. IECON '94. - 1994. - V.3. - P.2116-2121.
48. C. Ilas, A. Bettini, L. Ferraris, G. Griva, F. Profumo. Comparison of Different Schemes without Shaft Sensors for Field Oriented Control Drives // In Proc. IECON '94. - 1994. - V.3. - P.1579-1588.
49. M. Marchesoni, P. Segarich, E. Soressi. A Simple Approach to Flux and Speed Observation in Induction Motor Drives // In Proc. IECON '94. - 1994. - V.1. - P.305-310.
50. Tung-Hai Chin. Approaches for Vector Control of Induction Motor without Speed Sensor // In Proc. IECON '94. - 1994. - V.3. - P.1616-1620.
51. Ichiro Miyeshita, Akio Imayanayida, Takashi Koga. Recent Industrial Application of Speed Sensorless Vector Control in Japan // In Proc. IECON '94. - 1994. - V.3. - P.1573-1578.
52. Hisao Kubota and Kouki Matsuse. Speed Sensorless Field Oriented Control of Induction Machines using Flux Observer // In Proc. IECON '94. - 1994. - V.3. - P.1611-1615.
53. Сакае Ямамура. Спирально-векторная теория электрических машин переменного тока // Электротехника. - 1996. - №10. - С.7-15.
54. R. Castro-Linares, Jm. Alvarez-Gallegos, V. Vasques-Lopez. A Reduced Order Nonlinear Controller-Observer Scheme for the Control of Induction Motors // Preprints of the 4th IFAC Symp. NOLCOS '98. - 1998. - V.2. - P.303-308.
55. Young Ahn Kwon and Dae Won Jin. A Novel MRAS Based Speed Sensorless Control of Induction Motor // In Proc. IECON '99. - 1999. - PE - 15.
56. J. L. Mora, J. N. Tombs, R. Pachon, A. Torralba, M. Barranco and L. G. Franquelo. ASIC-Based Tachometer Without Mechanical Transducer for Induction Machines // In Proc. IECON '99. - 1999. - SP - 2.
57. Cyril Spiteri Staines, Greg M. Asher and Keith J. Bradley. A Periodic Burst Injection Method for Deriving Rotor Position in Saturated Cage-Salient Induction Motors without a Shaft Encoder // IEEE Trans. on Industry Applications. - 1999. - V.35, №4 - P. 851-858.
58. Lazhar Ben-Brahim, Susumu Tadakuma and Alper Akdag. Speed Control of Induction Motor without Rotational Transducers // IEEE Trans. on Industry Applications. - 1999. - V.35, №4 - P. 844-848.
59. Калашников В. И., Булахов В. В. Глубокорегулируемые системы векторного управления // Проблемы автоматизированного электропривода. Теория и практика. Вестник ХГПУ. Вып. 113. - Харьков, ХДПУ, 2000. - С.85-86.
60. Краснова С. А., Уткин В. А. Каскадный синтез асимптотических наблюдателей состояния асинхронного бездатчикового электропривода // Труды Международной конференции "Идентификация систем и задачи управления" (SICPRO '2000). - М, ИПУ. - 2000.
61. A. Dell' Aquila, V. Giliberti, F. S. Lovecchio, L. Salvatore. Real-Time Estimation of Induction Motor Parameters by LSE // In Proc. IECON '94. - 1994. - V.3. - P.2127-2131.
62. Millet C., Leroux D., Li Y., Feuvrie B., Bergmann C. Identification for a Field-Oriented Control Design of an Asynchronous Machine // Preprints of IFIB/IFAC/IMACS Conf. on Control of Industrial Systems. - 1997. - V.1/3. - P.713-723.
63. Pedro J. R. de Oliveira, Zelia M. A. Peixoto. Parameter Estimation of an Induction Machine Using a Continuous Time Model // In Proc. IECON '98. - 1998. - P.292-296.
64. F. Barrero, J. Rerez, R. Millan and L. G. Franquelo. Self-Commissioning for Voltage-Referenced Voltage-Fed Vector Controlled Induction Motor Drives // In Proc. IECON '99. - 1999. - SP - 2.
65. Steven R. Shaw and Stiven B. Leeb. Identification of Induction Motor Parameters from Transient Stator Current Measurements // IEEE Trans. on Industrial Electronics. - 1999. - V.46, №1. - P.139-149.
66. Потапенко Е. Е., Потапенко Е. М. К вопросу идентификации параметров и восстановления вектора состояния асинхронного двигателя // Вісник ХДПУ. Збірка наукових праць. Тематичний випуск 61. - Харків, ХДПУ. - С.82-84.
67. I. Takahashi and T. Noguchi. A Novel Quick Response and High Efficiency Control Strategy of an Induction Motor // IEEE Trans. Industry Application. - 1986. - V.IA - 22, №5. - P. 820-827.
68. M. Depenbrock. Direct Self Control (DSC) of Inverter-Fed Induction Machine // IEEE Trans. Power Electronics. - 1988. - V.3, №4. - P.420-429.
69. A. M. Walczyna. Problems of Application of Direct Flux and Torque Control Methods to High Power VSI-Fed Drives Operating at Low Speed // In Proc. IECON '94. - 1994. - V.1. - P.293-298.
70. D. Casadei, G. Grandi, G. Serra, A. Tani. Effects of Flux and Torque Hysteresis Band Amplitude in Direct Torque Control of Induction Machines // In Proc. IECON '94. - 1994. - V.1. - P.299-304.
71. K. Jezernik, M. Rodic, D. Drevensek. Speed Sensorless Torque Tracking Control of Induction Motor // In Proc. IECON '99. - Report № 00223.
72. Hoang Le-Huy. Behavioral Modeling and Simulation of a Direct-Torque-Control Induction Motor Drive Using PSPICE // In Proc. IECON '99. - 1999. - PE - 18.
73. V. Perelmutter. A Simplified Modeling of Induction Motor Drives with Direct Torque Control // In Proc. IECON '99. - 1999. - P.486-491.
74. H. Ludvigsen, R. Ortega, P. Albertos and O. Egeland. On Hybrid Control of Nonlinear Systems under Slow Sampling: Application to Induction Machines // In Proc. NOLCOS '9. - 1999. - V.2. - P.309-314.
75. Jun-Koo Kang and Seung-Ki Sul. New Direct Torque Control of Induction Motor for Minimum Torque Ripple and Constant Switching Frequency // IEEE Trans. Industry Applications. - 1999. - V.35, №5. - P.1076-1082.
76. Toshihiko Noguchi, Masaki Yamamoto, Seiji Kondo and Isao Takahashi. Enlarging Switching Frequency in Direct Torque-Controlled Inverter by Means of Dithering // IEEE Trans. on Industry Applications. - 1999. - V.35, №6. - P.1358-1366.
77. Ciro Attaianesse, Vito Nardi, Aldo Perfetto and Giuseppe Tomasso. Vectorial Control Torque: A Novel Approach to Torque and Flux Control of Induction Motor Drives // IEEE Trans. on Industry Applications. - 1999. - V.35, №6. - P.1399-1405.
78. Yang Qiu, Lipei Huang, Bing Bai, Zhiyun Xue. A Simple Online Estimation Method of Rotor Resistance for Induction Motor Using MRAC // In Proc. IECON '99. - 1999. - SP - 2.
79. K. Jezernik, V. Volcanjk. VSC Robust Control of an IM Servo-drive // In Proc. IECON '94. - 1994. - V.1. - P.627-632.
80. Z. Janda, M. Jankovic, J. Bebic, S. Vukosavich, V. Vuckovich. The Realization of a Novel Speed-Sensorless Induction Motor Drive // In Proc. IECON '94. - 1994. - V.3. - P.1621-1626.
81. Seung Ki Sul and Min ho Park. A Novel Technique for Optimal Efficiency Control of a Current-Source Inverter-Fed Induction Motor // IEEE Trans. on Power Electronics. - 1988. - V.3, №2. - P.192-198.
82. Сандлер А. С., Сарбатов П. С. Автоматическое частичное управление асинхронными двигателями. - М.: Энергия, 1974. - 328 с.
83. Булагов А. А. Частотное управление асинхронными двигателями. - М.: Энергоиздат, 1982. - 216 с.
84. A. Kusko and D. Galler. Control Means for Minimization of Losses in AC and DC Motor Drives // IEEE Trans. on Industry Applications. - 1983. - V.19, №4. - P.561-570.
85. Петров Ю. П. Оптимальное управление электрическим приводом. - М., - Л.: Госэнергоатомиздат, 1961. - 187с.
86. Петров Ю. П. Оптимальное управление электрическим приводом с учетом ограничений по нагреву. - Л.: Энергия, 1971. - 144 с.
87. Абдулаев Н. Д., Петров Ю. П. Теория и методы проектирования оптимальных регуляторов. - Л.: Энергоатомиздат, 1985. - 240 с.
88. Петров Ю. П. Синтез оптимальных систем управления при неполноте известных возмущающих сил. - Л.: Издательство Ленингр. ун-та, 1987. - 292 с.
89. C. Canudas de Wit and J. Ramirez. Optimal Torque Control for Current-Fed Induction Motors // IEEE Trans. on Automatic Control. - 1999. - V.44, №5. - P.1084-1089.
90. Попович Н.Г., Пересада С.М., Ковбаса С.Н., Король С.В. Энергетически эффективные алгоритмы управления асинхронными двигателями электромеханических систем // Вісник ХДПУ. Збірка наукових праць. Тематичний випуск 113. - Харків, ХДПУ, 2000. - С.25-29.
91. D. S. Kirschen, D. W. Novotny, T. A. Lipo. Optimal Efficiency Control of an Induction Motor Drive // IEEE Trans. on Energy Conversion. - 1987. - V.2, №1. - P. 70-75.
92. Sang-Hoon Kim, Seung-Ki Sul. Voltage Control Strategy for Maximum Torque Operation of an Induction Machine in the Field Weakening Region // In Proc. IECON '94. - 1994. - V.1. - P.599-604.
93. R. T. Novotnak, J. Chiasson and M. Bodson. High-Performance Motion Control of an Induction Motor with Magnetic Saturation // IEEE Trans. on Control Systems Technology. - 1999. - V.7, №3. - P.315-327.
94. Потапенко Е. Е., Потапенко Е. М. Энергосберегающее робастное управление асинхронными приводами // Механіка

- та машинобудування. - 1999. - №1. - С.106-111.
95. *Потапенко Е.Е., Потапенко Е.М.* Синтез экстремального робастного управления асинхронным приводом // Технічна електродинаміка. Тематичний випуск. - 2000. - Ч.6. - С.34-37.
96. *Бичай В. Г., Потапенко Е. Е., Потапенко Е. М.* Робастное экстремальное управление асинхронным приводом // Автоматика - 2000. Міжнародна конференція з автоматичного управління: Праці у 7-ми томах. - Т.5. - Львів: ДНДІІІ, 2000. - С.27-31.
97. *Грузов В. Л., Красильников А. Н., Машкин А. В.* Анализ и оптимизация алгоритмов управления в частотно регулируемых электроприводах с инверторами напряжения // Электротехника. - 2000. - №4. - С.15-20.
98. *Панкратов В. В., Нос О. В.* Оптимизация алгоритмов векторного управления асинхронным электроприводом на ос-

- нове методов непрерывной иерархии // Электричество. - 2000. - №6. - С.48-53.
99. *Изосимов Д. Б.* Новые подходы к синтезу цифрового управления в электроприводах переменного тока // Приводная техника. - 1997. - №4. - С.8-14; №5. - С.14-19.
100. *M. Rodic, K. Jezernik.* Torque Trajectory Controller for Induction Motor // In Proc. IECON'99. - P.641-646.
101. *Toshihiko Noguchi and Daichi Hiraishi.* Core Loss Compensation of Direct Field-Oriented Induction Motor Incorporating Robust Parameter Identification // In Proc. IECON'99. - Report № 00045.
102. *Sung-Don Wee, Myoung-Ho Shin, Dong-Seok Hyun.* Stator Flux-Oriented Vector Control of Induction Motors Considering Iron-Loss // In Proc. IECON'99. - Report № 00226.

УДК 517.938

ФОРМАЛИЗОВАННАЯ ИНТЕГРАЦИЯ ИНФОРМАЦИОННОЙ И ИСПОЛНИТЕЛЬНОЙ ПОДСИСТЕМ МНОГОСВЯЗНЫХ ОБЪЕКТОВ

В. Н. Буков, В. В. Косьянчук, В. Н. Рябченко

В статье излагается механизм формализованной интеграции систем наблюдения и управления. Показано, что если целью синтеза является обеспечение желаемого качества процессов замкнутой системы, то необходимо отказаться от традиционного принципа разделения и синтезировать совместно наблюдатель и регулятор.

Here it is considered the mechanism of the formalized integration of observation systems and controllers. It is shown that if the purpose of design being the achievement of desirable quality of processes of a loop system, it is necessary to refuse a conventional principle of separation and to design an observation system and a controller simultaneously.

ВВЕДЕНИЕ

Одним из перспективных направлений совершенствования сложных систем является интеграция. Под интеграцией понимается способ организации взаимодействия подсистем, обеспечивающий рациональное использование всех располагаемых ресурсов для достижения конечной цели [1]. Интеграция систем отражает в наиболее концентрированном виде (на фоне располагаемых ресурсов) такое фундаментальное свойство систем, как связность. При этом речь идет о связности элементов и подсистем во всех аспектах их существования и проявления. Именно это свойство определяет известные виды интеграции. Например: комплексирование, аппаратная интеграция, функциональная интеграция, технологическая интеграция, симбиотическая интеграция (симбиоз) [2,3]. Однако ни один из авторов не рассматривает формализованную (в виде формул) взаимосвязь различных по функциональному назначению систем. Этот подход будем называть формализованной интеграцией подсистем. Он основан на взаимосвязи решений задач подсистем в интересах достижения конечной цели, а неединственность алгебраических решений позволяет выбирать наиболее предпочтительный вариант в инте-

ресах удовлетворения неформализованных требований. Целью данной статьи является обоснование и разработка формализованной интеграции информационной и исполнительной подсистем многосвязного объекта.

ПОСТАНОВКА ЗАДАЧИ

Рассмотрим задачу формализованной интеграции информационных и исполнительных подсистем единого комплекса как задачу совместного синтеза линейного наблюдателя и регулятора в новой постановке.

В общем случае эту задачу можно представить схемой, показанной на рис. 1.

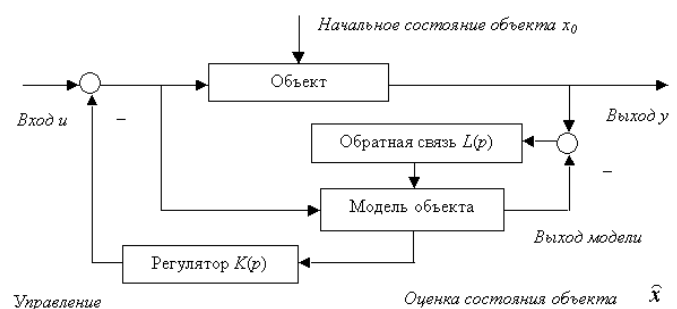


Рисунок 1

Управление объектом осуществляется по оцениваемому наблюдателем вектору состояния. Известно [4], что если объект управления представляется в пространстве состояния в виде

$$\begin{aligned} \dot{x} &= Ax + B\varepsilon, \quad y = Cx, \quad x(t_0) = x_0, \\ x &\in \mathbb{R}^n, \quad \varepsilon \in \mathbb{R}^s, \quad y \in \mathbb{R}^m, \end{aligned} \quad (1)$$

для которого введены следующие векторы: x - вектор состояния, ε - входной вектор, y - выходной вектор, x_0 - начальное состояние и A, B, C - действительные матрицы¹, то

1. Существует такая обратная связь $L(p)$, что все корни характеристического полинома λ_i могут быть размещены желаемым образом:

$$\exists L(p) \rightarrow \forall \lambda_i: \det(\lambda_i I_n - A + LC) = 0. \quad (2)$$

2. Характеристический полином наблюдателя и регулятора может быть представлен произведением характеристического полинома наблюдателя и характеристического полинома регулятора

$$\Delta p = (pI_n - A + BK)(pI_n - A + LC). \quad (3)$$

Однако анализ и синтез системы управления только по полюсам недостаточен для практики! Во-первых, при модальном управлении полюсы, системы всегда рассматриваются в общей совокупности, без уточнения канала "вход - выход", к которому каждый из корней относится, а во-вторых, на характер переходных процессов в линейных системах влияют не только корни знаменателя (полюсы), но и корни числителя (нули), а они для многосвязных систем традиционно не учитываются.

Поэтому предлагается осуществлять синтез как с учетом полюсов, так и нулей определяемых желаемыми матричными передаточными функциями (МПФ).

Поведение выходного сигнала такой системы, как и любой линейной системы, описывается матричным оператором

$$y = E_y^u(p)u + E_y^0(p)x_0 + E_y^{\hat{0}}(p)\hat{x}_0, \quad (4)$$

с матричными передаточными функциями (подчеркнем: в замкнутом контуре оценивания - управления) к выходу системы y от управления u на ее входе $E_y^u(p)$, начального состояния x_0 объекта $E_y^0(p)$ и начального состояния \hat{x}_0 наблюдателя $E_y^{\hat{0}}(p)$. Здесь p - оператор дифференцирования по времени и оператор преобразования Лапласа. Эти матричные передаточные функции исчерпывающим образом определяют динамические свойства рассматриваемой системы в случае ее минимальной реализации.

Начальное состояние наблюдателя можно считать нулевым, поэтому будем осуществлять синтез системы управления по двум желаемым МПФ $E_y^u(p)$ и $E_y^0(p)$

$$E_y^u(p) = C(pI_n - A + BK)^{-1}B, \quad (5)$$

$$E_y^0(p) = C(pI_n - A + BK(pI_n - A + BK + LC)^{-1}LC)^{-1}. \quad (6)$$

Если первое выражение достаточно широко известно, то вторая МПФ в литературе не встречается. Суть синтеза сводится к разрешению этих уравнений относительно K и L . В силу того, что матрицы C и B не квадратные, уравнения (5) и (6) известными методами неразрешимы. Именно поэтому разработчики методов модального управления использовали для синтеза только характеристический полином. Анализ матричных передаточных функций показывает, что

1. Матричные передаточные функции от входного воздействия u к состоянию x и выходу y объекта зависят от выбора только матрицы связей $K(p)$ регулятора.

2. Матричные передаточные функции от начального условия объекта x_0 к состоянию x и выходу y объекта

зависят одновременно от выбора матрицы связей $K(p)$ регулятора и матрицы обратных связей $L(p)$ наблюдателя. Следовательно, если под синтезом подразумевать одновременное обеспечение желаемых матричных передаточных функций $E_y^u(p)$ и $E_y^0(p)$ путем выбора матриц $K(p)$ и $L(p)$, то в общем случае разделение задач регулирования и наблюдения невозможно. Напомним, что классическая теорема разделения Калмана в детерминированных системах сформулирована и имеет силу для задачи обеспечения желаемых значений только корней характеристического полинома замкнутого контура "объект - наблюдатель состояния - регулятор".

Сформулируем задачу совместного синтеза линейного наблюдателя и регулятора как задачу управления выходом объекта.

Пусть дана система, включающая, линейный объект (1); линейный наблюдатель состояния

$$\begin{aligned} \dot{\hat{x}} &= A\hat{x} + L(y - \hat{y}) + B\varepsilon, \quad \hat{y} = C\hat{x}, \quad \hat{x}(t_0) = \hat{x}_0, \\ \hat{x} &\in \mathfrak{R}^n, \end{aligned} \quad (7)$$

где \hat{x} - оценка состояния объекта, \hat{x}_0 - начальное значение этой оценки, L - матрица обратных связей наблюдателя; линейный регулятор

$$\varepsilon = u - K\hat{x}, \quad u \in \mathfrak{R}^s. \quad (8)$$

где K - матрица обратных связей регулятора, u - вектор управления на входе системы.

Заданы желаемые матричные передаточные функции: по управлению $E_y^u(p)$ и по начальному состоянию $E_y^0(p)$.

1. Мы пока не уточняем, удовлетворяются ли критерии наблюдаемости и управляемости.

Необходимо найти дробно-рациональные полиномиальные матрицы $K(p)$ и $L(p)$.

АЛГОРИТМ РЕШЕНИЯ

Решение задачи основано на использовании так называемой технологии вложения систем [5-7]. В результате выполнения процедур этой технологии применительно к сформулированной задаче получен следующий алгоритм. Не раскрывая механизма получения конечных выражений, приведем и прокомментируем результат.

Предварительно по известным формулам определяются МПФ объекта управления

$$F_y^u(p) = C(pI_n - A)^{-1}B, \tag{9}$$

$$F_x^u(p) = (pI_n - A)^{-1}B, \tag{10}$$

$$F_y^0(p) = C(pI_n - A)^{-1}. \tag{11}$$

Алгоритм решения состоит из четырех шагов:

1. Решить относительно $K(p)$ матричное линейное уравнение

$$E_y^u(p)K(p)F_x^u(p) = F_y^u(p) - E_y^0(p). \tag{12}$$

В результате определяется искомая матрица связей регулятора $K(p)$.

2. Вычислить для идеального наблюдения МПФ замкнутого контура

$$\Phi_y^0(p) = C(pI_n - A + BK)^{-1}. \tag{13}$$

3. Вычислить вспомогательную (интеграционную) матрицу

$$\pi_{\hat{x}}(p) = \Phi_y^0(p) - E_y^0(p). \tag{14}$$

4. Решить относительно $L(p)$ матричное линейное уравнение

$$\pi_{\hat{x}}(p)L(p)F_y^0(p) = E_y^0(p) - F_y^0(p). \tag{15}$$

В результате определяется искомая матрица связей наблюдателя $L(p)$.

Из этого алгоритма следует, что:

1. Матрица обратных связей регулятора не зависит от наблюдателя;

2. Матрица обратных связей наблюдателя зависит от матрицы обратных связей регулятора. Это еще раз подтверждает сформулированное положение о необходимости совместного синтеза наблюдателя и регулятора, названного формализованной интеграцией. Здесь впервые

в виде конструкций, определяющих и использующих матрицу $\pi_{\hat{x}}$, формализована связь задач синтеза наблюдателя и регулятора, часто считающихся независимыми. Выбор этой матрицы определяет весь диапазон возможных вариаций взаимосвязи указанных систем. Итак, обеспечить желаемую реакцию замкнутой системы "объект - наблюдатель состояния - регулятор" на ненулевые начальные условия можно как изменением коэффициентов регулятора, так и наблюдателя.

3. Недоопределенность уравнений (12) и (15), имеющая место в большинстве практических случаев, позволяет построить множество эквивалентных альтернативных решений.

4. Основные области применения разработанного подхода связаны, во-первых, с восстановлением координат состояния, недоступных непосредственному измерению, а во-вторых, с комплексированием систем наблюдения, включая задачи коррекции инерциальных систем.

ПРИМЕР

Рассмотрим решение задачи совместного синтеза наблюдателя и регулятора на примере построения системы позиционной коррекции навигационной системы летательного аппарата, показанной на рис. 2.

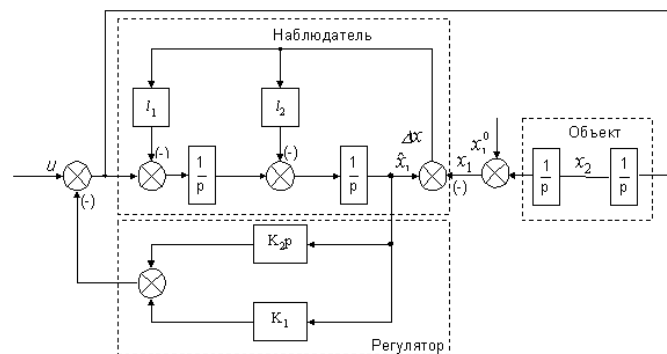


Рисунок 2

Упрощенная модель бокового движения объекта при полете вдоль заданного маршрута над плоской землей представляется последовательным соединением двух интеграторов. Управление движением осуществляется по сигналам, снимаемым с инерциальной системы (наблюдатель), в частности, по координате x_1 . В режиме позиционной коррекции сигналы, соответствующие координатам объекта x_1 и наблюдателя \hat{x}_1 , сравниваются. Начальные условия объекта ненулевые и равны x_1^0 . Необходимо определить матрицу L наблюдающего устройства, которая обеспечит списание ошибки по

координате x_1 желаемым образом.

Пусть объект представлен в виде (1), (7), (8) с соответствующими матрицами

$$A = \begin{bmatrix} 0 & 1 \\ 0 & 0 \end{bmatrix}, B = \begin{bmatrix} 0 \\ 1 \end{bmatrix}, C = \begin{bmatrix} 1 & 0 \end{bmatrix}. \quad (16)$$

Пусть матричные передаточные функции регулятора определены и имеют значения $k_1(p) = 1$, $k_2(p) = 5$.

В соответствии с (11) определим передаточную функцию объекта

$$F_y^0(p) = \begin{bmatrix} \frac{1}{p} & \frac{1}{p^2} \end{bmatrix}. \quad (17)$$

Желаемые матричные передаточные функции выбраны в виде

$$E_x^0(p) = \begin{bmatrix} E_{11}^0 & \frac{1}{p}E_{11}^0 \\ E_{21}^0 & \frac{1}{p}(1 + E_{21}^0) \end{bmatrix}, \quad (18)$$

где

$$E_{11}^0(p) = \frac{p^3 + 7p^2 + 17p}{p^4 + 7p^3 + 17p^2 + 32p + 6},$$

$$E_{21}^0(p) = \frac{-(2p^2 + 24p + 8)}{p^4 + 7p^3 + 17p^2 + 32p + 6}. \quad (19)$$

Свободной составляющей выхода системы при идеальном наблюдении соответствует матричная передаточная функция

$$\Phi_x^0(p) = \frac{1}{p^2 + k_2(p)p + k_1(p)} \begin{bmatrix} p + k_2(p) & 1 \\ -k_1(p) & p \end{bmatrix}. \quad (20)$$

В итоге матрица $L(p)$ имеет вид числовой матрицы

$$L(p) = \begin{bmatrix} 2 \\ 6 \end{bmatrix}. \quad (21)$$

Реакция системы на начальные условия x_1^0 представлена на рис. 3.

Видно, что при $L(p) = 0$ реакция затянута. При

$L(p) = \begin{bmatrix} 2 \\ 6 \end{bmatrix}$ реакция объекта соответствует желаемой.

ВЫВОДЫ

Таким образом, в работе получены следующие результаты.

1. Обоснована необходимость совместного синтеза информационной и исполнительной подсистем многосвязных объектов. Показано, что для обеспечения желаемого качества процессов функционирования многосвязных систем в интегрированном комплексе необходимо отказаться от принципа разделения.
2. Разработан формализованный подход интеграции процедуры синтеза многосвязных законов функционирования информационных и исполнительных подсистем многосвязного объекта. Предлагаемый подход позволяет обеспечить как размещение полюсов и нулей замкнутой системы "объект - наблюдатель состояния - регулятор" на комплексной плоскости, так и "закрепление" их за конкретными входами и выходами.
3. Особо хотелось бы отметить, что понятие управляемости и наблюдаемости не используются. Вместо этих широко известных критериев используются условия разрешимости системы матричных линейных уравнений и казуальность (физическая реализуемость) полученных решений.

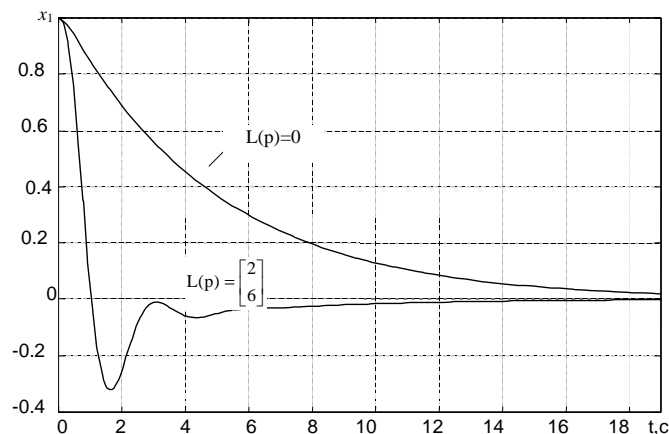


Рисунок 3

ПЕРЕЧЕНЬ ССЫЛОК

1. Красовский А.А., Ермилов А.С. Боевое применение и эффективность пилотажно-навигационных комплексов. - М.: ВВИА им. Н.Е. Жуковского, - 1989. - 240с.
2. Алиев Р.А. Методы интеграции в системах производств. - М.: Энергоиздат, - 1989. - 190с.
3. Рубинштейн М.И. Оптимальная группировка взаимосвязанных объектов. М.: Наука, - 1989. - 355с.
4. Zhou K., Doyle J.C. Essentials of robust control. Upper Saddle River, New Jersey, Prentice Hall, - 1998. - 280p.
5. Буков В.Н., Косьянчук В.В. Вложение систем. Линейное наблюдение // Автоматика и телемеханика. - 2001. - № 2. - С. 3-20.
6. Буков В.Н., Рябченко В.Н. Вложение систем. Произвольные образы // Автоматика и телемеханика. - 2000. - № 12. - С.15-32.
7. Буков В.Н., Рябченко В.Н. Вложение систем. Линейное управление // Автоматика и телемеханика - 2001. - № 1. - С.3-23.

ОПТИМИЗАЦИЯ ПАРАМЕТРОВ ЦИФРОВЫХ НЕЧЕТКИХ РЕГУЛЯТОРОВ

В. И. Гостев, И. П. Лесовой, А. Е. Чуприн

Изложены важные для проектирования цифровых нечетких регуляторов вопросы оптимизации их основных параметров (диапазонов изменения лингвистических переменных, формы и параметров функций принадлежности лингвистических величин) путем минимизации квадратичного критерия качества с целью получения оптимальных переходных процессов в системе автоматического управления с нечетким регулятором.

Викладені важливі для проектування цифрових нечітких регуляторів питання оптимізації їх основних параметрів (діапазонів зміни лінгвістичних змінних, форми і параметрів функцій приналежності лінгвістичних величин) шляхом мінімізації квадратичного критерію якості з метою здобуття оптимальних перехідних процесів в системі автоматичного управління з нечітким регулятором.

The problems of optimization of leading particulars of digital fuzzy-controllers (ranges of variation linguistic variables, shape and parameters of functions of an inhering of linguistic magnitudes) are explained important for their projection by minimization of square-law criterion of quality with the purpose of deriving optimum transient processes in a system of automatic control with an fuzzy-controller.

Применение нечетких регуляторов (регуляторов, работающих на базе нечеткой логики) для управления различными (в частности, нестационарными и нелинейными) объектами показывает их высокую эффективность и в ряде случаев существенные преимущества перед линейными цифровыми регуляторами [1-7]. Основными параметрами цифровых нечетких регуляторов, при которых производится их синтез и расчет, являются, во-первых, количество и форма функций принадлежности $\mu^T(u)$ лингвистических величин и, во-вторых, диапазоны изменения входных и выходной лингвистических переменных *ошибка*, *первая производная ошибки*, *вторая производная ошибки*, *управляющее воздействие на объект*, т.е.

$$[\theta_{\min}, \theta_{\max}], [\dot{\theta}_{\min}, \dot{\theta}_{\max}], [\ddot{\theta}_{\min}, \ddot{\theta}_{\max}], [m_{\min}, m_{\max}].$$

Количество функций принадлежности ФП (число термов, описывающих входные и выходную переменные) обычно ограничивают, по возможности, наименьшим числом. ФП аппроксимируют обычно треугольниками. Для настройки треугольных ФП на экспертные данные можно пользоваться операцией возведения в степень: $[\mu^T(u)]^k$, где показатель степени определяет изменение формы ФП. Так, если на универсальном множестве $U = [0, 1]$ заданы два нечетких подмножества, треугольные функции принадлежности которых для каждой лингвистической величины (например, термов *положительный-1*, *отрицательный-2*) определяются в виде (см. рис.1).

Если $\mu^1(u) = 1 - u$; $\mu^2(u) = u$, $u \in [0, 1]$, то для настройки можно использовать ФП вида

$$\mu^1(u) = (1 - u)^c; \mu^2(u) = u^c, u \in [0, 1], \quad (1)$$

и варьировать коэффициентом (показателем степени) c .

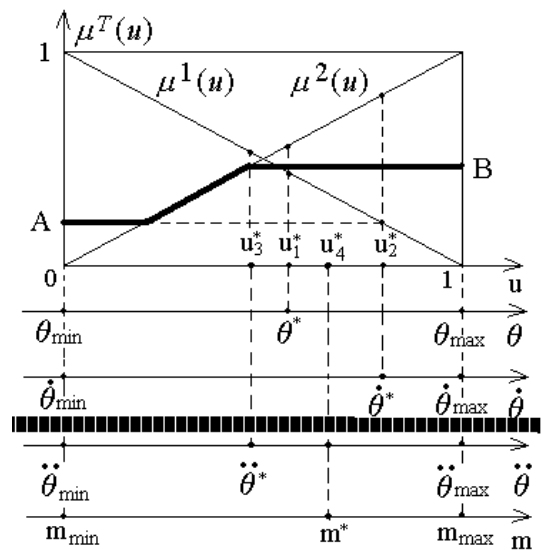


Рисунок 1

Если на универсальном множестве $U = [0, 1]$ заданы два нечетких подмножества, функции принадлежности которых для каждой лингвистической величины записываются в виде $\mu^T(u) = 1 / \left[1 + \left(\frac{u-b}{c} \right)^2 \right]$, $u \in [0, 1]$, и определяются как

$$\mu^1(u) = 1 / \left[1 + (u/c)^2 \right], \mu^2(u) = 1 / \left[1 + \left(\frac{u-1}{c} \right)^2 \right], \quad u \in [0, 1], \quad (2)$$

то для настройки можно также варьировать коэффициентом c .

На универсальном множестве $U = [0, 1]$ два нечетких подмножества можно описать также функциями принадлежности

$$\mu^1(u) = e^{-cu}, \mu^2(u) = e^{-c(1-u)}, u \in [0, 1], \quad (3)$$

и для настройки варьировать коэффициентом c .

Для каждой лингвистической переменной можно использовать свои функции принадлежности. Таким образом, существует достаточно большое число вариантов задания функций принадлежности при оптимизации параметров нечеткого регулятора.

Результирующую ФП получают обычным способом, а расчет абсциссы центра тяжести $s_c = S(u_c, \mu_c)$ участка площади, охватываемой результирующей ФП $\mu(u)$ в пределах изменения переменной u от $u = U_1$ до $u = U_2$, удобно определять, используя численное интегрирование по методу трапеций, по формуле

$$u_c = \frac{\frac{U_1 \mu_0}{2} + \sum_{i=1}^{M-1} u_i \mu_i + \frac{U_2 \mu_M}{2}}{\frac{\mu_0}{2} + \sum_{i=1}^{M-1} \mu_i + \frac{\mu_M}{2}}, \quad (4)$$

где $(U_2 - U_1)/M = u_0$ - шаг дискретизации, M - число дискрет на интервале $U_2 - U_1$, $i = 1, 2, 3, \dots, M - 1$.

При определении результирующей ФП необходимо определять абсциссы точек пересечения ФП нечетких подмножеств (например, термов *положительный-1*, *отрицательный-2*) с горизонтальными прямыми. Наиболее просто это выполнить для треугольных ФП. Для ФП вида

$$\mu^1(u) = (1 - u)^c; \mu^2(u) = u^c, u \in [0, 1]$$

абсциссы точек пересечения определяются как

$$u^* = 1 - \sqrt[c]{\mu^1(u^*)} \text{ и } u^* = \sqrt[c]{\mu^2(u^*)}. \quad (5)$$

Для ФП вида

$$\mu^1(u) = 1/[1 + (u/c)^2], \mu^2(u) = 1/[1 + ((u-1)/c)^2], \\ u \in [0, 1]$$

абсциссы точек пересечения определяются как

$$u^* = c \times \sqrt{\frac{1}{\mu^1(u^*)} - 1} \text{ и } u^* = 1 + c \times \sqrt{\frac{1}{\mu^2(u^*)} - 1}. \quad (6)$$

Для ФП вида

$$\mu^1(u) = e^{-cu}, \mu^2(u) = e^{-c(1-u)}, u \in [0, 1]$$

абсциссы точек пересечения определяются как

$$u^* = -\frac{1}{c} \ln \mu^1(u^*) \text{ и } u^* = 1 + \frac{1}{c} \ln \mu^2(u^*). \quad (7)$$

При оптимизации параметров цифровых нечетких регуляторов в системах автоматического управления необходимо задавать критерий качества и функции воздействий (управляющее и/или возмущающие воздействия) на систему. Наиболее часто при оптимизации используют один из квадратичных критериев качества, например,

$$J = \frac{1}{L} \sum_{v=0}^{L-1} \theta_v^2 \Rightarrow \min, \quad (8)$$

где ошибка системы θ_v вычисляется с шагом моделирования h_0 , а число L определяет интервал наблюдения. Оптимальные параметры соответствуют минимальному значению критерия качества, а минимизация критерия качества автоматически приводит к оптимизации переходных процессов в системе управления. Можно использовать различные алгоритмы условной и безусловной оптимизации.

Рассмотрим систему управления (см. рис.2) с цифровым нечетким регулятором и нестационарным объектом управления, "замороженная" передаточная функция которого определяется формулой

$$G(s) = \frac{\alpha_2(s)}{m(s)} = \frac{\alpha}{s(s^2 + bs + a)}.$$

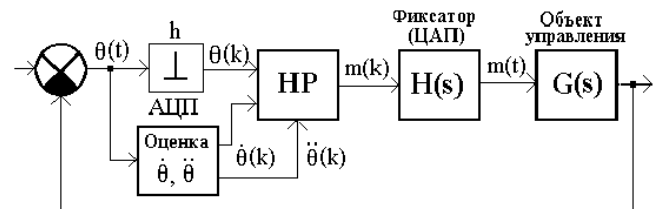


Рисунок 2

Синтез нечеткого регулятора для такого объекта выполнен в работе [5].

При заданных в работе [5] законах изменения параметров передаточной функции объекта управления и входного воздействия и треугольных ФП (см. рис.1) выбранные без оптимизации (настройка "вручную") диапазоны изменения входных и выходного параметров нечеткого регулятора (диапазоны изменения переменных $\theta, \hat{\theta}, \ddot{\theta}, m$) определяются следующим образом:

$$[\theta_{\min}, \theta_{\max}] = [-1,02; 1,02], [\hat{\theta}_{\min}, \hat{\theta}_{\max}] = [-4, 4], \\ [\ddot{\theta}_{\min}, \ddot{\theta}_{\max}] = [-24, 24] \text{ и } [m_{\min}, m_{\max}] = [-1, 1].$$

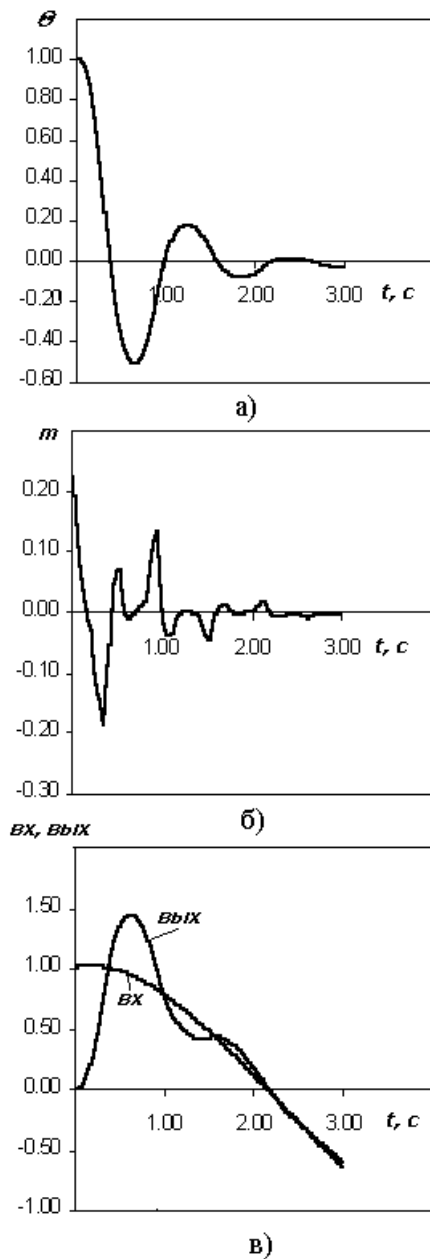


Рисунок 3

Шаг квантования в цифровом регуляторе $h=0,01с$, шаг моделирования $h_0=0,0005с$. Интервал наблюдения равен 3 с.

Квадратичный критерий качества имеет показатель $J = 0,021$.

Проведем оптимизацию этих параметров по указанному выше критерию при заданных в работе [5] воздействиях на входе системы и законах изменения параметров передаточной функции объекта, используя метод оптимизации Хука-Дживса.

После оптимизации получаем следующие диапазоны изменения входных и выходного параметров (диапазоны изменения переменных $\theta, \dot{\theta}, \ddot{\theta}, m$):

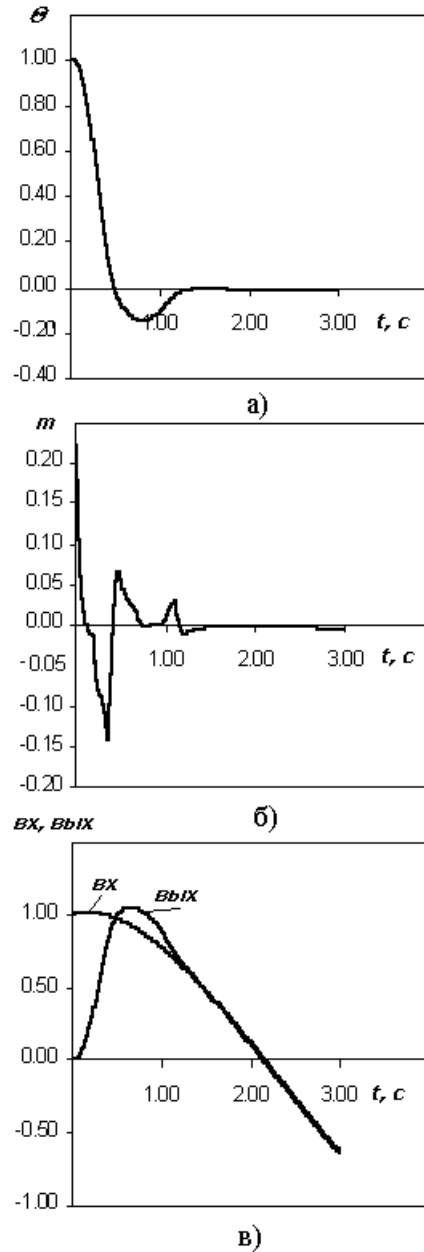


Рисунок 4

$$[\theta_{\min}, \theta_{\max}] = [-1,02; 1,02],$$

$$[\dot{\theta}_{\min}, \dot{\theta}_{\max}] = [-2,75; 2,75],$$

$$[\ddot{\theta}_{\min}, \ddot{\theta}_{\max}] = [-16,52; 16,52] \text{ и}$$

$$[m_{\min}, m_{\max}] = [-1, 1].$$

Квадратичный критерий качества имеет показатель $J = 0,0168$.

Ниже представлены результаты исследования точности обработки системой автоматического управления с нечетким регулятором заданного закона изменения входного воздействия: на рис.3 - регулятор с неоптимизируемыми параметрами; на рис.4 - регулятор с оптималь-

ными параметрами $[\dot{\theta}_{\min}, \dot{\theta}_{\max}]$ и $[\ddot{\theta}_{\min}, \ddot{\theta}_{\max}]$. Сравнивая результаты, заключаем, что оптимизация параметров приводит к значительному улучшению качества системы управления, которое характеризуется величиной текущей ошибки и численно определяется показателем J .

Можно использовать различные рассмотренные выше формы функций принадлежности (см. формулы (1)-(3)) и, варьируя одновременно коэффициентом c и диапазонами изменения входных и выходного параметров (диапазонами изменения переменных $\theta, \dot{\theta}, \ddot{\theta}, m$), находить минимальное значение показателя J . Проведенные расчеты методом оптимизации Хука-Дживса с использованием формул (4)-(8) дают следующие результаты. Наибольшее из минимальных значений показателя J получается при использовании ФП, определяемых по формуле (2), при следующих параметрах цифрового нечеткого регулятора: $c = 10^{-4}$;

$$[\theta_{\min}, \theta_{\max}] = [-1,02; 1,02],$$

$$[\dot{\theta}_{\min}, \dot{\theta}_{\max}] = [-2,14, 2,14],$$

$$[\ddot{\theta}_{\min}, \ddot{\theta}_{\max}] = [-22, 23; 22, 23] \text{ и}$$

$$[m_{\min}, m_{\max}] = [-1, 1].$$

Квадратичный критерий качества имеет показатель $J = 0,0197$.

Наименьшее из минимальных значений показателя J при использовании экспоненциальных ФП (см. формулы (3) и (7)) при следующих параметрах цифрового регулятор: $c = 72,5$;

$$[\theta_{\min}, \theta_{\max}] = [-1,02; 1,02],$$

$$[\dot{\theta}_{\min}, \dot{\theta}_{\max}] = [-3,75; 3,75],$$

$$[\ddot{\theta}_{\min}, \ddot{\theta}_{\max}] = [-22,98; 22,98] \text{ и}$$

$$[m_{\min}, m_{\max}] = [-1, 1].$$

Квадратичный критерий качества имеет показатель $J = 0,0136$.

Таким образом, для рассмотренной системы автоматического управления экспоненциальные функции принадлежности являются наилучшими при выбранном критерии качества. С целью выбора оптимальных параметров нечетких регуляторов для конкретных объектов управления следует производить расчеты для различных ФП и, используя оптимизационные программы, выбирать ФП, при которых вычисляемый по формуле (8) показатель качества J является минимальным.

ПЕРЕЧЕНЬ ССЫЛОК

1. Архангельский В.И., Богаенко И.Н., Грабовский Г.Г., Рюшин Н.А. Системы функций-управления.- К.: Техника, 1997.- 208 с.
2. Архангельский В.И., Богаенко И.Н., Грабовский Г.Г., Рюшин Н.А. Досвід розвитку і застосування систем функцій-управління // Автоматизація виробничих процесів.- 1997.- №2(5).- С.1-10.
3. Ротштейн А.П. Интеллектуальные технологии идентификации.- Винница: "УНІВЕРСУМ - Вінниця", 1999.- 320 с.
4. Гостев В.И., Чуприн А.Е. Система управления частотой вращения ротора газотурбинного двигателя на базе нечеткой логики // Электротехника и электроэнергетика.- 2000.- N1.- С.5-9.
5. Гостев В.И., Чуприн А.Е., Лесовой И.П. Синтез цифрового регулятора системы управления нестационарным объектом на базе нечеткой логики // Механіка та машинобудування.- 2000.- №1.- С.128-133.
6. Гостев В.И., Баранов А.А., Чуприн А.Е., Худолий Д.А. Синтез цифрового нечеткого регулятора системы управления нестационарным объектом // Праці міжнародної конференції з автоматичного управління "Автоматика-2000": Львів, 11-15 вересня 2000 р.- Т.2 - Львів: Державний НДІ інформаційної інфраструктури, 2000.- С.63-67.
7. Гостев В.И., Лесовой И.П., Чуприн А.Е. Синтез цифрового нечеткого регулятора системы управления объектом с нелинейностью типа "люфт" // Автоматизація виробничих процесів.- 2000.- N1(10).- С.113-116.

УДК 658.562

ИНСТРУМЕНТАЛЬНЫЕ СРЕДСТВА ВСЕОБЩЕГО УПРАВЛЕНИЯ КАЧЕСТВОМ: КОНТРОЛЬНЫЕ КАРТЫ

Ю. Н. Дорошенко, В. И. Дубровин

Рассмотрены проблемы в области контрольных карт. Приведена классификация контрольных карт и особенности их применения. Доказана эффективность многомерного контроля качества и приведен пример построения карты Хотеллинга.

Розглянуто проблеми в галузі контрольних карт. Наведена класифікація контрольних карт та особливості їх викори-

стання. Доведена ефективність багатомірнього контролю якості і наведений приклад побудови карти Хотеллінга.

The problems in areas of control charts are considered. The classification of control charts and features their applications is adduced. The efficiency of multivariate control is proven and the example of construction of Hotelling charts is adduced.

ВВЕДЕНИЕ

Одной из наиболее актуальных задач современного производства является выпуск конкурентоспособной продукции. Качественная, а значит конкурентоспособная, продукция обеспечивает конкурентоспособность предприятия и оказывает положительное воздействие на экономику. Поэтому повышение качества продукции является не только проблемой отдельных предприятий, а оказывает существенное влияние на уровень жизни всего населения [1]. Таким образом, на первый план выходят задачи создания и внедрения эффективного инструментария управления качеством. Статистические методы управления качеством [2], в частности, контрольные карты [3,4,5], представляют собой действенное средство разработки новой технологии и контроля качества производственных процессов.

1. КЛАССИФИКАЦИЯ КОНТРОЛЬНЫХ КАРТ

Согласно стандарту [6,7], контрольная карта - карта, на которой для наглядности отображения состояния логистического процесса отмечают значения соответствующей выборочной характеристики смежных выборок во временной последовательности. Иными словами, контрольная карта - разновидность графика, который отличается наличием контрольных границ, обозначающих допустимый диапазон разброса характеристик в обычных условиях течения процесса. Контрольная карта предназначена для оценки степени статистической устойчивости процесса.

В зависимости от различных признаков, существует несколько классификаций контрольных карт [3,8,9].

1. Классификация на основе статистических оценок:

- контрольные карты для регулирования по количественным признакам;
- контрольные карты для регулирования по качественным признакам.

2. Классификация в зависимости от сферы применения:

- контрольные карты для регулирования технологических процессов;
- контрольные карты для анализа технологических процессов.

3. Классификация по количеству представляемых процессов:

- краткие контрольные карты;
- многопоточные групповые карты.

В таблицах 1 и 2 представлены типы контрольных карт для регулирования по количественным и качественным признакам.

Преимущества контрольных карт по качественному признаку:

- позволяют быстро получить общее представление о различных аспектах качества анализируемого изделия, то есть, на основании различных критериев качества можно сразу принять или забраковать продукцию;

Таблица 1 - Типы контрольных карт для регулирования по количественным признакам

| Тип контрольной карты | Наименование (сопряженный вид распределения) | Назначение и условия применения |
|------------------------|----------------------------------------------------------------------------|-------------------------------------------------------------------------------|
| \bar{x} -карта | Измеряемое значение (распределение средних арифметических) | В случае, когда сведения о процессе поступают через большие интервалы времени |
| $(\bar{x} - R)$ -карта | Среднее значение и размах (распределение средних арифметических и размаха) | Для анализа таких показателей качества, как длина, масса, диаметр, время |

Таблица 2 - Типы контрольных карт для регулирования по качественным признакам

| Тип контрольной карты | Наименование (сопряженный вид распределения) | Назначение и условия применения |
|-----------------------|------------------------------------------------------|-------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|
| p-карта | Доля дефектов (биномиальное распределение) | Для доли дефектных изделий и определения интенсивности выпуска продукции |
| np-карта | Число дефектных изделий (биномиальное распределение) | В случае, когда контролируемым параметром является число дефектных изделий при постоянном объеме выборки n |
| c-карта | Число дефектов (распределение Пуассона) | В случае, когда контролируемым параметром является число дефектов, обнаруживаемых среди постоянных объемов продукции |
| u-карта | Число дефектов на единицу (распределение Пуассона) | В случае, когда контролируемым параметром является число дефектов, однако обращаться с выборкой как с определенным объемом продукции невозможно (когда площадь, длина, масса непостоянны) |

- позволяют обойтись без применения дорогих точных приборов и требующих значительных затрат времени измерительных процедур. Кроме того, этот тип контрольных карт более понятен менеджерам, которые не разбираются в тонкостях методов контроля качества.

Таким образом, с помощью таких карт можно более убедительно продемонстрировать руководству наличие проблем с качеством изделий.

Преимуществом контрольных карт по количественному признаку является большая чувствительность по сравнению с контрольными картами по качественному признаку. Благодаря этому, контрольные карты по количественному признаку могут указать на существование проблемы ухудшения качества, прежде чем в потоке продукции появятся бракованные изделия, выделяемые с помощью контрольной карты по качественному признаку.

Контрольные карты применяются не только для регулирования процесса, а также для его анализа и совершенствования, так как являются источником полезной информации о процессе благодаря тому, что положение границ регулирования позволяет дать вероятностную оценку положения средних значений, т.е. уровня настройки технологического процесса и рассеивания значений.

Краткая контрольная карта (контрольная карта для кратких производственных серий) представляет собой график наблюдаемых значений характеристик качества (значений непрерывной переменной или альтернативного признака) для нескольких частей процесса, причем все значения контролируемой характеристики наносятся на одну и ту же карту. Разработка кратких контрольных карт стала следствием необходимости адаптации контрольных карт к тем ситуациям, когда требуется выполнить несколько десятков измерений контролируемой характеристики процесса, прежде чем вычислить контрольные пределы. Часто данное требование выполняется с трудом на тех стадиях производственного процесса, в ходе которых изготавливается ограниченное (малое) число деталей, которые необходимо подвергнуть измерениям.

Групповая контрольная карта дает возможность нанести данные для нескольких потоков наблюдаемых значений непрерывной переменной или альтернативного признака (характеристик качества) на одну и ту же карту. Это упрощает интерпретацию карты при одновременном управлении большим числом процессов или их характеристик. Здесь термином "потоки процесса" могут обозначаться данные, полученные для различных станков, сборочных линий, операторов и так далее. Все эти данные могут быть нанесены на одну контрольную карту: для каждой из выборок с измерениями контролируемой характеристики на карту наносится две точки, в результате чего на графике образуются две линии. Верхняя из них представляет собой график наиболее высоких средних значений каждой выборки для всех нанесенных на карту потоков переменных или альтернативных признаков, а нижняя - подобный график наименьших средних значений каждой выборки. Следовательно, с помощью групповой карты, можно быстро

определить, не началась ли разладка процесса в одном или нескольких потоках процесса или контролируемых характеристиках, не переходя к проверке всех измерений подряд.

2. КОНТРОЛЬНЫЕ КАРТЫ В АВТОМАТИЗИРОВАННЫХ СИСТЕМАХ И ПАКЕТАХ

Ранее одним из показателей, определяющих целесообразность той или иной процедуры управления качеством, была сложность расчетов и, соответственно, затрат на них. Однако, с появлением недорогих персональных компьютеров, все большую популярность приобретают процедуры, требующие проведения большего объема вычислений. В пакетах Statistica [10,11,12], Statgraphics Plus [13] имеются инструменты для построения и анализа различных видов контрольных карт. В частности, в модуле Контроль качества пакета Statistica предлагается широкий набор инструментов для исследования характеристик процесса, в том числе установление значений плановых спецификаций, задание параметров критериев серий, вычисление негауссовских контрольных пределов, индексов пригодности и операционных характеристик. Все таблицы результатов и графики, как и в других модулях системы, можно редактировать, сохранять и автоматически помещать в файл отчета. Автоматизированные опции и средства для быстрого вызова упрощают рутинные операции, а практически все многочисленные графические опции могут быть изменены в любое время (сохранены как системные установки по умолчанию или как шаблоны многократного использования).

Такого рода инструменты включены также в автоматизированные системы управления предприятия типа SAP R/3, BAAN IV. В частности, в систему R/3 встроены модуль "Управление качеством" (QM - Quality Management), включающий процедуры для построения контрольных карт:

- контрольная карта для среднего значения с включением допуска (приемочная карта);
- контрольная карта для среднего значения без включения допуска (карта Шухарта);
- контрольная карта для среднеквадратического отклонения (карта Шухарта).

3. ПРОБЛЕМАТИКА КОНТРОЛЬНЫХ КАРТ

В последнее время специалистами обсуждается ряд вопросов, связанных с контрольными картами.

1. Дискуссии относительно эволюции в теории контрольных карт [14]. Сейчас большое внимание уделяется вопросам, связанным с теорией управления качеством, концепцией TQM, идеями сертификации, в то время как теории контрольных карт не уделяется достаточного внимания. Со времен Шухарта контрольные карты дей-

ствительно эволюционировали недостаточно.

2. Дискуссии относительно терминов в области контрольных карт.

- Термин "статистически управляемый процесс". С одной стороны, процесс считается статистически управляемым тогда и только тогда, когда распределения всех его измеримых характеристик постоянны во времени (все средние и дисперсии всех измеримых характеристик постоянны во времени) [15]. С другой стороны, процесс является управляемым, когда на контрольной карте нет точек, сигнализирующих о наличии специальных причин вариации [14].

- Термин "чувствительности". Согласно теории проверки статистических гипотез Неймана и Пирсона, один критерий является более чувствительным по отношению к другому, если первый критерий при всех возможных альтернативных значениях параметра дает большую вероятность обнаружения ложности проверяемой гипотезы, чем второй. В [15] предложен следующий способ определения индекса чувствительности правил выявления наличия специальных причин вариации на контрольной карте: индекс чувствительности некоторого правила, пригодность которого для интерпретации показаний контрольных карт требуется определить, - это отношение вероятности правила 1 (точка на контрольной карте лежит за контрольными пределами) к вероятности того правила, для которого индекс чувствительности определяется. Однако, как указывается в [14], это определение не учитывает число выборок, необходимых для срабатывания правил (каждое правило требует различное число точек).

3. Дискуссии вокруг правил интерпретации контрольных карт и эффективности использования вероятностных правил. Использование вероятностных пределов на контрольной карте дает увеличение ее чувствительности, однако требует знания вида и параметров закона распределения рассматриваемых характеристик, что на практике не всегда известно. Использование доверительных интервалов, кратных стандартному отклонению, проще, однако дает меньшую чувствительность карты к изменению параметра процесса. Чувствительность дает преимущество в тех случаях, когда контроль связан с большими затратами, но, когда затраты на контроль невелики, это преимущество утрачивается.

Таким образом, теория контрольных карт, несмотря на значительный возраст, не вполне устоялась.

Несмотря на то, что контрольные карты являются достаточно эффективным инструментом управления качеством, при их использовании возникает ряд проблем [14,16]:

1. Классические контрольные карты в ряде случаев не позволяют быстро обнаружить разладку процесса, тогда как среднее число выборок, предшествующее ложному сигналу о разладке, для них невелико. Для ускорения обнаружения разладки существует три пути:

- Увеличение объема выборки n и использование классических контрольных карт. Недостатком этого способа является то, что увеличение объема выборок не всегда возможно и целесообразно по ряду причин. При наличии большого числа результатов наблюдений для контрольных карт размахов, а также средних арифметических и размахов, оказывается удобным формировать их в выборки определенного размера. Периоды отбора таких выборок могут быть определены на основе обработки результатов достаточно простых производственных экспериментов.

- Использование более эффективных модификаций контрольных карт.

В этом случае могут использоваться карты кумулятивных сумм, для которых тенденция роста наносимых на карту точек позволяет быстрее обнаружить разладку.

- Использование контрольных карт с предупредительными границами также позволяет ускорить обнаружение разладки процесса в случае малых и медленно образующихся смещений контролируемого параметра. Для таких карт устанавливается пара предупредительных границ, размещенных между границами регулирования, и правило, согласно которому принимается решение о корректировке процесса. Это правило заключается в том, что всегда можно указать такое количество последовательных точек, вероятность попадания которых в область между внутренней и внешней границей будет настолько малой, что можно утверждать о разладке процесса.

2. Большинство контрольных карт разработаны для характеристик, распределение которых близко к нормальному. Однако измеряемые показатели качества часто подчинены распределениям, отличным от нормального. Повышение вероятности выработки правильного решения в случае ненормального распределения контролируемой характеристики может быть достигнуто за счет составления специальных таблиц для констант контрольных карт. Также разработаны специальные модификации контрольных карт.

- X-карты для данных с негауссовским распределением. При малых объемах выборок и сильной асимметрии распределения показателей качества, построенные по таким данным стандартные контрольные пределы приводят как к получению большого числа ложных сигналов тревоги (т.е. росту вероятности альфа-ошибки), так и к увеличению числа случаев, когда при фактически произошедшей разладке процесс продолжает считаться контролируемым (росту вероятности бета-ошибки). В негауссовских x-картах расчет контрольных пределов ведется на основе так называемых кривых Джонсона, с помощью которых аппроксимируется асимметрия и эксцесс большой группы негауссовских распределений. Негауссовские x-карты рекомендуется применять в том случае, когда распределение выборочных средних обладает явной асимметрией или является негауссовским.

- Контрольные карты для логнормального распределения (распределение показателя качества асимметрично). Путем трансформирования показателя качества (путем логарифмирования его значений) удается преобразовать асимметричное распределение в симметричное. Если логарифмы числовых значений показателя качества образуют нормальное распределение, то эффективно использование данного вида контрольных карт.

3. В случае существования нескольких показателей качества, которые коррелированы между собой, использование классических контрольных карт часто приводит к ошибочным результатам. В этом случае следует применять многомерные контрольные карты.

4. МНОГОМЕРНЫЙ КОНТРОЛЬ КАЧЕСТВА

Одномерные контрольные карты достаточно эффективны в случае наличия независимых показателей качества продукции (X_1, X_2, \dots, X_i), когда проводится построение отдельных контрольных карт по каждому из показателей и управляемость процесса определяется суммированием результатов анализа по этим картам. Однако для прослеживания центральной тенденции определенного числа измеримых показателей качества, особенно при наличии корреляции этих показателей между собой, эффективно использование многомерного контроля [16]. В этом случае интегральный показатель качества не может быть определен достаточно точно, а построение отдельных контрольных карт по каждому из показателей приводит к искажению результатов.

Многомерная контрольная карта представляет собой карту оценки развития процесса с учетом реакции двух и более характеристик, объединенных в качестве одной статистики для мгновенной выборки. Если переменные или характеристики рассматриваются как взаимозависимые, то, как правило, составляется χ^2 статистика. Если переменные или характеристики коррелированы, то используется T^2 статистика.

Разработка многомерных контрольных карт основана на T^2 контрольных картах Хотеллинга. Q-карты, разработанные Гхаром и Торгерсоном, предназначены для представления основной тенденции нескольких показателей качества на одной карте. Отслеживание стабильности выбранной переменной с помощью многомерной контрольной карты в подходящей доверительной области базируется на том факте, что квадратическая форма многомерного нормального распределения подчинена хиквадрат. Частным случаем Q-карты являются двухпараметрические контрольные карты, предназначенные для определения наличия одной причины разладки данного параметра.

Модель экономического обоснования T^2 контрольных карт Хотеллинга для контроля двух или более переменных, разработанная Монтгомери и Клаттом, осно-

вана на предположении, что:

- состояние процесса (налаженное или разлаженное) обнаруживается точно во время извлечения частной выборки;
- функция затрат для случаев, когда действуют очень большие объемы выборок и сложные правила контроля, носит экспоненциальный характер.

Предположим, что производится контроль качества некоторого процесса. Через определенные интервалы времени производится измерение состава в некотором продукте трех ингредиентов: Пр1, Пр2 и Пр3. Данные представлены в таблице 3.

Таблица 3 - Исходные данные

| № | Пр1 | Пр2 | Пр3 | № | Пр1 | Пр2 | Пр3 |
|---|-----|------|------|----|------|------|------|
| 1 | 6.5 | 34.5 | 52.9 | 7 | 7.1 | 34.6 | 47.6 |
| | 6.7 | 33.9 | 50.8 | | 7.1 | 39.2 | 53.7 |
| | 8 | 38.1 | 47.7 | | 7.8 | 27.7 | 65.4 |
| | 9.3 | 37.1 | 46.4 | | 8.6 | 41.5 | 44.1 |
| | 7.2 | 33.8 | 49.2 | | 7.1 | 39.2 | 49.3 |
| 2 | 7.9 | 38.8 | 56.3 | 8 | 6.7 | 29.1 | 48.1 |
| | 8.9 | 38.4 | 48.4 | | 6.5 | 35.9 | 50.1 |
| | 7.9 | 31.9 | 54.9 | | 9.2 | 33.6 | 49 |
| | 8.7 | 37.2 | 48.8 | | 6.4 | 43.5 | 57.8 |
| | 7.3 | 33.8 | 47.8 | | 6.7 | 34.6 | 50.8 |
| 3 | 7.8 | 36.3 | 52 | 9 | 8.7 | 33.3 | 44.8 |
| | 6.9 | 38.5 | 50.5 | | 6.8 | 36.8 | 43.3 |
| | 6.5 | 31.7 | 52 | | 7 | 35.3 | 51.3 |
| | 8.6 | 35.3 | 56.4 | | 7.5 | 34.8 | 47.3 |
| | 6.6 | 30.6 | 57.5 | | 8.9 | 28.1 | 48.4 |
| 4 | 8.8 | 37.9 | 57 | 10 | 8.8 | 27.9 | 53.5 |
| | 7.9 | 34.7 | 51.7 | | 6 | 31.4 | 48.4 |
| | 6.5 | 40.6 | 47 | | 8.2 | 27.8 | 47.1 |
| | 8 | 33.8 | 48.3 | | 6.6 | 31.8 | 48.7 |
| | 7.8 | 39 | 47.6 | | 6.9 | 31.2 | 47.1 |
| 5 | 8.7 | 37.5 | 53.8 | 11 | 7.7 | 40.1 | 63 |
| | 7.2 | 35.6 | 48.7 | | 7.9 | 28.9 | 45.6 |
| | 8.6 | 35 | 47.4 | | 8.4 | 31.6 | 56.4 |
| | 8.9 | 31.3 | 54.1 | | 7.7 | 48.1 | 54.1 |
| | 7.7 | 34.4 | 47.5 | | 9.2 | 39.4 | 48.5 |
| 6 | 7.1 | 32.8 | 56.9 | 12 | 9.1 | 40.2 | 50.1 |
| | 6.6 | 43.9 | 48.8 | | 10.4 | 37.8 | 48.2 |
| | 7 | 42.8 | 49.6 | | 7 | 32.1 | 49.9 |
| | 8.9 | 33.5 | 48.2 | | 8.1 | 35.5 | 48.5 |
| | 7.3 | 36.6 | 52.7 | | 8.4 | 34.2 | 60.4 |

Для карты Хотеллинга исходные коррелируемые признаки преобразуются в новые переменные, которые являются независимыми, и затем создается единая статистика, основанная на новых переменных. Верхний контрольный предел определяется по известному объему выборки и числу переменных для соответствующих вероятностей альфа-ошибки (вероятности ошибочно посчитать процесс вышедшим из-под контроля, когда на самом деле это не так, а появление выброса обусловлено случайной причиной). Для контрольной карты данного типа может быть задан только верхний контрольный предел, потому что статистика T^2 принимает только положительные значения.

Таблица 4 - Расчет статистики T^2 Хотеллинга

| № выб. | T^2 для выборки | T^2 для отдельных значений | № выб. | T^2 для выборки | T^2 для отдельных значений |
|--------|-------------------|------------------------------|--------|-------------------|------------------------------|
| 1 | 0.7680 | 1.87754 | 7 | 0.90558 | 1.05720 |
| | | 1.30476 | | | 1.67298 |
| | | 0.94449 | | | 13.6812 |
| | | 3.67873 | | | 4.97858 |
| | | 0.62196 | | | 1.35200 |
| 2 | 1.1745 | 2.21974 | 8 | 2.46576 | 3.87782 |
| | | 2.21140 | | | 1.71958 |
| | | 1.50745 | | | 2.62489 |
| | | 1.3585 | | | 7.92932 |
| | | 0.86078 | | | 1.21166 |
| 3 | 3.07600 | 0.11777 | 9 | 5.22633 | 3.02043 |
| | | 1.28956 | | | 4.02854 |
| | | 2.47738 | | | 0.59699 |
| | | 2.43421 | | | 0.73977 |
| | | 4.72408 | | | 4.68721 |
| 4 | 1.03217 | 3.62852 | 10 | 12.50413 | 4.66387 |
| | | 0.09248 | | | 4.60225 |
| | | 3.91066 | | | 4.17473 |
| | | 0.54388 | | | 2.44150 |
| | | 1.22033 | | | 2.56326 |
| 5 | 1.37798 | 1.75451 | 11 | 5.34974 | 8.77655 |
| | | 0.56934 | | | 3.83199 |
| | | 1.36778 | | | 2.82646 |
| | | 2.94168 | | | 9.54578 |
| | | 0.63617 | | | 3.42280 |
| 6 | 2.47685 | 2.56014 | 12 | 4.94481 | 3.31315 |
| | | 5.56714 | | | 8.25254 |
| | | 3.66520 | | | 1.27490 |
| | | 1.97252 | | | 0.40705 |
| | | 0.44742 | | | 5.21877 |

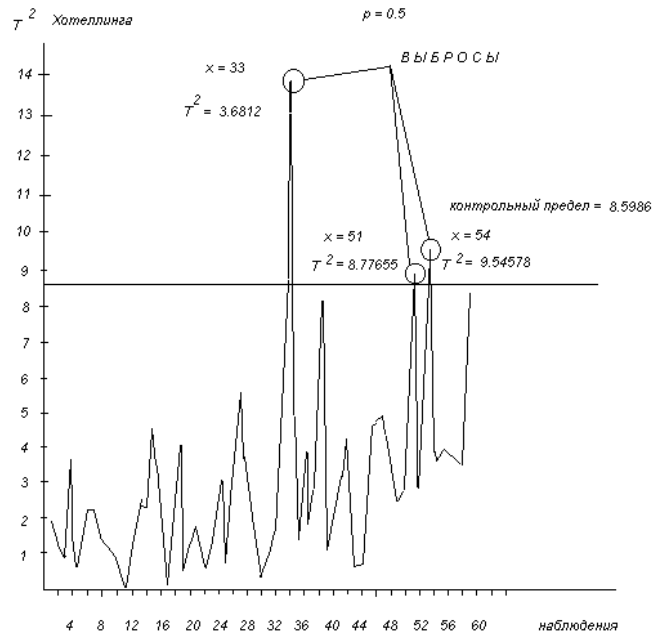


Рисунок 1 - Контрольная карта T^2 Хотеллинга для отдельных наблюдений

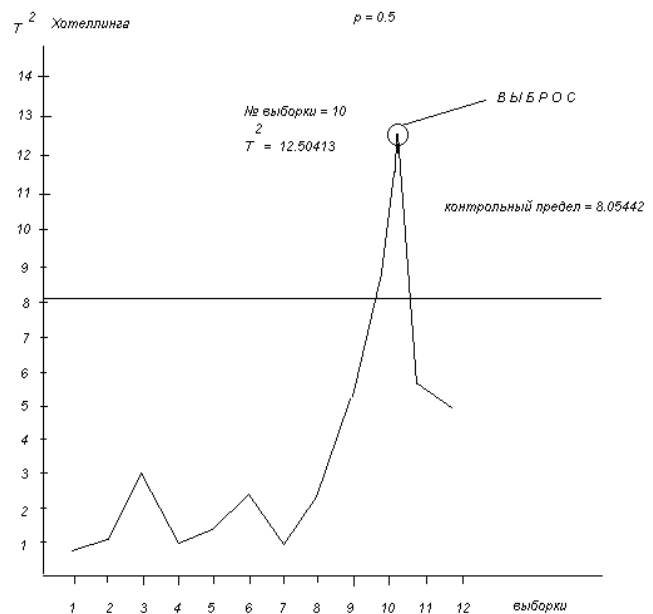


Рисунок 2 - Контрольная карта T^2 Хотеллинга для группированных данных

Из таблицы 4 и рис. 1-2 видно, что показания контрольной карты для выборок и для отдельных наблюдений не всегда совпадают. На основании анализа этих контрольных карт можно определить номер выборки или отдельного наблюдения, для которых имеется выброс

(одна или несколько составляющих вышли за пределы плановых спецификаций). Возможно, выброс обусловлен плохим качеством продукции, полученной от поставщика.

В [17] представлены результаты анализа эффективности многомерных контрольных карт по сравнению с одномерными. Этот анализ показал, что в большинстве случаев использование многомерной контрольной карты оказывается более эффективно, чем использование карт по отдельным показателям. Сравнение одномерных и многомерных контрольных карт проводилось по моменту обнаружения сдвига процесса на примере карт Шухарта для значений X_1 , X_2 и Хотеллинга для T^2 .

Факторы, свидетельствующие о необходимости использования многомерных контрольных карт:

- существует нескольких показателей качества объекта, коррелированных между собой;

- построение достаточно точной линейной регрессионной модели для интегрального показателя качества затруднено.

В настоящее время многомерные контрольные карты на практике используются крайне редко, поэтому существует необходимость развития теории многомерных контрольных карт, руководств по их практическому использованию, стандартов, а также включения инструментов для построения и анализа многомерных контрольных карт в компьютерные статистические пакеты.

ВЫВОДЫ

Контрольные карты, являющиеся одним из инструментов контроля качества, при их правильном применении в сочетании с другими инструментами контроля, позволяют производить эффективный анализ собранных сведений о показателях качества процесса. Это позволяет получать ценную информацию для решения возникающих проблем. Наибольшие сложности в теории контрольных карт и практике их использования вызывают следующие моменты:

- анализ на стадии подготовки технологического процесса и выбор типа контрольной карты, использование которой обеспечит минимальную вероятность принятия неправильного решения в сочетании с небольшими затратами;

- развитие теории многомерных контрольных карт и практики их применения;

- стандартизация терминов, методов построения контрольных карт и особенностей их использования.

Указанные проблемы целесообразно решать посредством:

- разработки автоматизированных систем, включающих процедуры выбора видов контрольных карт в зависимости от характера анализируемых показателей, построения контрольных карт, их анализа и выработки рекомендаций по управляющим воздействиям;

- теоретической разработки методов многомерного контроля качества и оценки его целесообразности в конкретных условиях;

- проведения обобщенного анализа информации и систематизации существующих видов контрольных карт с целью разработки стандартов.

ПЕРЕЧЕНЬ ССЫЛОК

1. *Огвоздин В.Ю.* Управление качеством. Основы теории и практики. - М.: Дело и сервис, 1999. - 160 с.
2. Семь инструментов качества в японской экономике. - М.: Издательство стандартов, 1990. - 88 с.
3. *Мердок Дж.* Контрольные карты. - М.: Финансы и статистика, 1986. - 251 с.
4. *Джонсон Н., Лион Ф.* Статистика и планирование эксперимента в технике и науке. - М.: Мир, Т.1, 1980. - 610 с.
5. Статистические методы контроля качества продукции / *Нюлер Л.* и др. - М.: Издательство стандартов, 1989. - 96 с.
6. ДСТУ 3514-97. Статистичні методи контролю та регулювання. Терміни та визначення. - К.: Держстандарт України, 1997. - 27 с.
7. ИСО 3534-2. Статистика - Словарь и условные обозначения - Часть 2: Статистическое управление качеством.
8. Статистические методы повышения качества / Под ред. *Х.Кумэ.* - М.: Финансы и статистика, 1990. - 304 с.
9. *Саката Сиро.* Практическое руководство по управлению качеством. - М.: Машиностроение, 1980. - 215 с.
10. *Боровиков В.П., Боровиков И.П.* STATISTICA - Статистический анализ и обработка данных в среде Windows. - М.: Информационно-издательский дом "Филинь", 1998. - 608 с.
11. *Боровиков В.П.* Популярное введение в программу STATISTICA - М.: Компьютер Пресс, 1998. - 267 с.
12. *Тюрин Ю.Н., Макаров А.А.* Статистический анализ данных на компьютере / Под ред. В.Э.Фигурнова. - М.: Инфра - М., 1998. - 528 с.
13. Руководство пользователя пакета STATGRAPHICS.
14. *Шпер В.А.* Еще раз о контрольных картах и вокруг них // Надежность и контроль качества, 1998, №10, - с.3-13.
15. *Hoyer R.W., Ellis W.C.* A Graphical Exploration of SPC. Part 1: SPC's definitions and procedures. - Quality Progress. - 1996. - Vol. 29. - № 5. - pp.65-73.
16. *Бендерский А.М., Богатырев А.А., Баумгартен Л.В.* Стандартизация статистических методов управления качеством. - М.: Издательство стандартов, 1983. - 152 с.
17. *Клячкин В.Н.* Оценка эффективности многомерного контроля качества технологического процесса // Вестник УлГУ. Информационные технологии, 1999, №2, С.59-62.

ОНЛАЙНОВЫЕ МЕТОДЫ УПРАВЛЕНИЯ КАЧЕСТВОМ: ГИБРИДНАЯ ДИАГНОСТИКА НА ОСНОВЕ НЕЙРОННЫХ СЕТЕЙ

В. И. Дубровин, С. А. Субботин

Предложена методика решения задач диагностики на основе интегрированной гибридной распознающей системы. Разработаны алгоритмы, реализующие отдельные этапы предложенной методики. Проведены эксперименты по практическому применению предложенной методики.

Запропоновано методику розв'язку задач діагностики на основі інтегрованої гібридної розпізнаючої системи. Розроблено алгоритми, що реалізують окремі етапи запропонованої методики. Проведено експерименти по практичному застосуванню запропонованої методики.

The technique of diagnostics problem solving on the basis of the integrated hybrid recognizing system is offered. The algorithms realizing separate stages of an offered technique are developed. The experiments on practical application of an offered technique are conducted.

ВВЕДЕНИЕ

С развитием сложных промышленных систем возникает все большая потребность в разработке и внедрении онлайн-методов управления качеством процессов и изделий, в частности, диагностики в режиме реального времени. Среди множества разработанных методов диагностики методы распознавания образов обеспечивают наиболее удобный подход к моделированию сложных систем, особенно в тех случаях, когда другие методы не могут использоваться или их использование сопряжено с большими затратами.

Напомним основные задачи распознающей системы: система должна исследовать процесс на основе наборов значений признаков; система должна функционировать по принципу "черного ящика", т.е. она должна некоторым образом моделировать взаимосвязь между входом и выходом; построенная модель должна удовлетворять заданной целевой функции.

Распознавание образов может быть формально определено как отображение из пространства наблюдений в пространство решений и, следовательно, представляет собой задачу классификации.

Цель настоящей работы состоит в том, чтобы создать методологию построения адаптивных робастных распознающих систем реального времени, допускающих контролируемое и неконтролируемое обучение.

В качестве основы при построении подобных систем предлагается использовать нейронные сети (НС), характеризующиеся высокой степенью адаптивности, универсальности и разнообразием архитектур и моделей [1-4].

1. АРХИТЕКТУРА ГИБРИДНЫХ СИСТЕМ ДИАГНОСТИКИ

При построении систем распознавания образов, к которым относятся и диагностические системы, необходимо решить три главных задачи:

- отбор информативных признаков (традиционно - оффлайн-методы);
- моделирование объекта (процесса);
- принятие решения - выдача определенного управляющего воздействия.

Традиционно большинство систем распознавания для решения конкретной задачи используют какой-либо один метод классификации. Однако ни один метод классификации не является универсальным и абсолютно надежным, и разные методы, как правило, основываются на разной философии. Вследствие этого, методы, обеспечивающие одинаковую или близкую надежность классификации для всей обучающей и/или контрольной выборки в целом, для отдельных экземпляров могут давать разные результаты (это особенно характерно для экземпляров, находящихся на стыках границ классов).

Естественно, что подобные системы, построенные на основе одного классификатора, не будут удовлетворять высоким требованиям надежности, предъявляемым во многих областях промышленности. Поэтому крайне важно создавать гибридные системы классификации, интегрирующие различные модели и методы классификации.

Но и этого не достаточно, поскольку построение эффективной модели объекта возможно только при наличии оптимального набора признаков, причем, для каждой модели, в общем случае, оптимальные наборы будут различными и одного только предварительного задания набора признаков на основе оффлайн-методов [5] (предварительный физический анализ, планирование и т.п.) будет недостаточно. Поэтому необходимо многократное итерационное повторение комбинации "отбор признаков - построение модели - отбор признаков".

Однако и этого не достаточно, поскольку, как правило, классификаторы содержат достаточное количество параметров, которые необходимо настроить вручную (для НС - это количество слоев, количество нейронов в слое, структура связей, параметры алгоритма обучения). Для автоматизации этого процесса необходима разработка процедур развития (роста) модели.

Получив систему, состоящую из классификаторов, имеющих оптимальное для данного классификатора

количество входов, необходимо оптимизировать их структуру (для НС - удалить (уменьшить) избыточные малоинформативные и усилить информативные связи при помощи алгоритма контрастирования).

Рассмотрев основные этапы диагностики, можно прийти к выводу, что их раздельная реализация неэффективна и для решения большинства прикладных задач целесообразно использовать некую гибридную интегрированную систему, включающую реализации различных блоков для решения рассмотренных задач. Структура такой системы может быть аналогична приведенной на рис. 1.

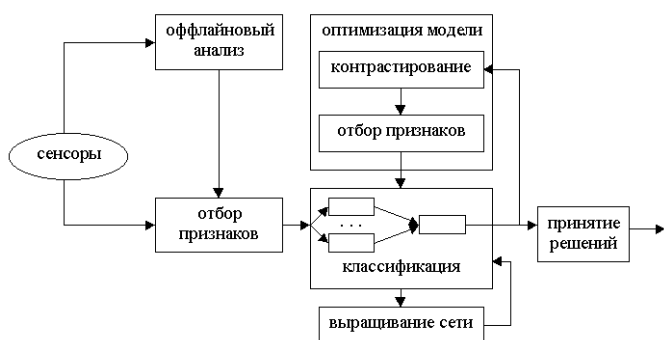


Рисунок 1 - Обобщенная схема гибридной диагностики

Как видно из рис. 1, процесс гибридной диагностики должен включать следующие этапы.

1. Предварительный отбор признаков на основе оффлайн методов и статистического анализа.
2. Построение модели объекта.
 - 2.1 Создание и обучение многоклассификаторной модели объекта.
 - 2.2 Модификация параметров отдельных классификаторов для улучшения надежности классификации (выращивание НС).
 - 2.3 Оптимизация нейросетевых моделей (отбор признаков и контрастирование).
 - 2.4 Создание и обучение блока-интегратора.
3. Создание и обучение блока принятия решений.

2. АЛГОРИТМЫ ГИБРИДНОЙ ДИАГНОСТИКИ

2.1. Алгоритм функционирования гибридной системы диагностики

Шаг 1. Задать набор признаков, характеризующих моделируемый объект и обучающую выборку.

Шаг 2. Произвести предварительный отбор информативных признаков.

Шаг 3. Построить и оптимизировать модель объекта.

Шаг 4. Проверить адекватность модели.

Шаг 5. Если модель адекватна, то перейти на шаг 6, в

противном случае - изменить параметры модели и перейти на шаг 3 или, если количество итераций шагов 3-4 превышает установленный предел, перейти на шаг 7, выдав сообщение о невозможности построения адекватной модели.

Шаг 6. Построить блок принятия решений.

Шаг 7. Останов.

2.2. Предварительный отбор информативных признаков

В настоящее время признано [1], что распознавание сложных образов наиболее целесообразно проводить на основе их относительного описания (описания в пространстве признаков). Выбор информативной системы признаков - наиболее важная задача теории распознавания. Однако удовлетворительного решения, определяющего порядок автоматического выделения признаков посредством переработки информации, получаемой на уровне абсолютного описания объекта, пока не найдено. Поэтому основным решением остается автоматизированный выбор наиболее информативных признаков из некоторого исходного множества (ансамбля) свойств, задаваемого эвристически.

При наличии достаточно большого объема диагностических данных можно производить их обработку с целью количественной оценки информативности совокупности имеющихся начальных признаков.

Оценку информативности начальных признаков, связанных с параметрами - критериями качества изделия, можно производить с помощью аппарата линейного регрессионно-корреляционного анализа [5] путем подсчета коэффициентов парной и / или множественной корреляции между экстремальными значениями параметра-критерия в процессе испытания (или значениями параметра-критерия в конце испытаний) и начальными признаками изделий.

Информативность признаков оценивается путем анализа полученных коэффициентов парной корреляции. Из начальных признаков выбираются те, которые имеют значимые коэффициенты парной корреляции с параметром-критерием. Затем из этих признаков отбираются те, которые имеют наименьшие коэффициенты парной корреляции между собой. По совокупности отобранных признаков подсчитываются коэффициенты линейной регрессии и коэффициент множественной корреляции данной совокупности начальных признаков с параметром-критерием. Коэффициент множественной корреляции является мерой информативности данной совокупности.

При наличии нормального распределения начальных признаков в каждом классе их частная информативность оценивается по формуле:

$$J_j = \frac{(\bar{x}_j^A - \bar{x}_j^B)^2}{\sigma^2}, \text{ если } \sigma = \sigma_{x_j^A} = \sigma_{x_j^B};$$

$$J_j = 0,5(\sigma_{x_j^A} - \sigma_{x_j^B})^2 \left(\frac{1}{(\sigma_{x_j^B})^2} - \frac{1}{(\sigma_{x_j^A})^2} \right) +$$

$$+ 0,5 \left(\frac{1}{(\sigma_{x_j^B})^2} - \frac{1}{(\sigma_{x_j^A})^2} \right) (\bar{x}_j^A - \bar{x}_j^B)^2, \text{ если } \sigma_{x_j^A} \neq \sigma_{x_j^B},$$

где J_j - частная мера информативности j -го признака;

\bar{x}_j - среднее значение j -го признака;

σ_{x_j} - среднеквадратическое отклонение j -го признака.

Индексы А и Б подразумевают вычисление соответствующей величины только для экземпляров соответствующего класса.

После нахождения оценок информативностей признаков вычисляется их среднее арифметическое J_{cp} , которое далее используется как порог при принятии решений о значимости признаков. Признак x_j относят к значимым, если $J_j > CJ_{cp}$, где C - некоторая заданная константа для регуляции значения порога принятия решения о значимости признаков, и к незначимым (малозначимым) - в противном случае.

2.3. Алгоритм построения и оптимизации модели объекта

Шаг 0. Задать набор классификаторов: определить используемые модели, начальные параметры моделей и методы обучения.

Шаг 1. Обучить все классификаторы на основе обучающей выборки.

Шаг 2. Для всех экземпляров обучающей выборки последовательно подать на входы классификаторов значения признаков для текущего экземпляра x^s , $s = 1, \dots, S$, где S - количество экземпляров в обучающей выборке, и определить K_k^s - класс s -го экземпляра для k -го классификатора, $k = 1, \dots, G$, где G - количество классификаторов.

Шаг 3. Набор результатов классификации $K_{вх} = \{K_k^s\}$ для всех экземпляров и классификаторов и набор фактических значений классов $K = \{K^s\}$ для всех экземпляров подать для обучения на входы и выход блока-интегратора, соответственно. Если в качестве блока-интегратора используется НС Кохонена [1], то подается только набор $K_{вх}$.

Шаг 4. Обучить интегратор моделированию зависимости между классами на выходах классификаторов и фактическим классом.

Шаг 5. Вычислить I_k - вклад k -го классификатора в формирование общего результата, $k = 1, 2, \dots, G$. Если в качестве блока-интегратора используется НС Кохоне-

на, то I_k могут быть определены как информативности входов НС на основе алгоритма отбора информативных признаков (см. раздел 2.5).

$$\sum_{k=1}^G I_k$$

Шаг 6. Для классификаторов, у которых $I_k > \frac{k=1}{G}$,

выполнить процедуру оптимизации структуры (контрастировать связи и/или оценить информативность и осуществить отбор признаков).

$$\sum_{k=1}^G I_k$$

Для классификаторов, у которых $I_k \leq \frac{k=1}{G}$, если

возможно, выполнить процедуру выращивания, иначе исключить классификатор из системы.

Шаг 7. Если значимость классификатора в процессе P повторений шагов 1-6 меньше определенного заданного числа, то исключить классификатор из системы, иначе - перейти на шаг 1.

Выход из цикла, определяемого шагами 1-7, осуществляется, если будет выполняться хотя бы одно из условий:

- 1) количество классификаторов $G \leq 2$;
- 2) достигнуты пределы увеличения / сокращения структуры для всех классификаторов (дальнейшее применение алгоритмов контрастирования и выращивания нецелесообразно);
- 3) требуемая точность классификации достигнута;
- 4) требуемая точность классификации не достигнута в течение заданного времени.

2.4. Алгоритм контрастирования сети

Шаг 0. Сохранить текущий набор значений весов и порогов сети.

Шаг 1. Задать или вычислить по определенному правилу граничное значение значимости весов \bar{w} . В качестве такого значения, например, можно использовать среднее арифметическое значение модулей весов и порогов сети.

Шаг 2. Все веса $|w_j^{(\mu, i)}| < \bar{w}$, где $w_j^{(\mu, i)}$ - вес j -го входа i -го нейрона μ -го слоя НС, уменьшить на основе правила $w_j^{(\mu, i)} = R^-(w_j^{(\mu, i)})$, где μ, i, j удовлетворяют заранее определенным правилам и ограничениям.

Шаг 3. Все веса $|w_j^{(\mu, i)}| \geq \bar{w}$ увеличить на основе правила $w_j^{(\mu, i)} = R^+(w_j^{(\mu, i)})$, где μ, i, j удовлетворяют заранее определенным правилам и ограничениям.

Шаг 4. Для всех экземпляров обучающей выборки x^s , $s = 1, \dots, S$ вычислить значение суммарной ошибки сети Еггог.

Шаг 5. Если $Error < MaxError$, где $MaxError$ - некоторое заданное максимальное граничное значение ошибки сети, то перейти на шаг 0, в противном случае - перейти на шаг 6.

Шаг 6. Восстановить предыдущие значения весов и порогов сети.

Шаг 7. Останов.

В качестве правил $R(w)$ предлагается использовать:

$$R^-(w) = \begin{cases} w - \alpha, & \text{если } w > \alpha, w \geq \alpha; \\ w + \alpha, & \text{если } w < \alpha, w \leq -\alpha; \\ 0, & \text{если } (w = 0) \text{ или } (w > 0 \text{ и } w < \alpha) \\ & \text{или } (w < 0 \text{ и } w > -\alpha), \end{cases}$$

$$R^+(w) = \begin{cases} w + \alpha, & \text{если } w > 0; \\ w - \alpha, & \text{если } w < 0; \\ 0, & \text{если } w = 0 \end{cases}$$

или

$$R^-(w) = \begin{cases} w\alpha, & 0 < \alpha < 1, \text{ если } w > \xi; \\ 0, & \text{если } w \leq \xi, \end{cases}$$

$$R^+(w) = w/\alpha, \quad 0 < \alpha < 1,$$

или

$$R^-(w) = w(1 - P(w)), \quad R^+(w) = w/P(w),$$

где α, ξ - заранее заданные константы,

$$P(w) = \begin{cases} -\frac{2}{\pi} \arctg \frac{1}{w}, & \text{если } w < 0; \\ \frac{2}{\pi} \arctg \frac{1}{w}, & \text{если } w > 0; \\ 1, & \text{если } w = 0. \end{cases}$$

При задании этих констант следует руководствоваться следующими соображениями:

1) слишком большие значения констант могут существенно ухудшить результат и затруднить процесс обучения;

2) слишком малые значения констант могут привести к существенному увеличению времени работы алгоритма контрастирования;

3) чем меньше значения констант, тем меньше будет дополнительная ошибка классификации, возникающая после выполнения алгоритма контрастирования.

В отличие от большинства других алгоритмов контрастирования [6,7], предполагающих удаление малоинформативных связей, данный алгоритм осуществляет последовательное увеличение больших весов и уменьшение

малых, что позволяет более гибко осуществлять процесс контрастирования.

2.5. Алгоритм отбора информативных признаков

После того, как НС обучена моделированию зависимости, на ее основе можно оценить информативность используемых признаков и затем отобрать наиболее информативные признаки для дальнейшего моделирования.

Для НС, имеющей один выход, то есть содержащей на последнем слое только один нейрон, степень влияния i -го признака x_i , значения которого поступают на входы нейронов первого слоя НС, на выходной сигнал у также будет, в основном, определяться значениями весов и признаков (входов) НС.

Количественная оценка значимости i -го признака будет определяться как сумма значимостей i -ых входов нейронов первого слоя НС. Для оценки значимости входов НС необходимо оценить не только их значимость относительно нейрона, но и значимость выхода нейрона как входа нейронов следующего слоя и так далее до нейронов последнего слоя. Очевидно, что реализовать подобную процедуру проще в обратном порядке, двигаясь от последнего слоя к первому, поскольку вычисление значимости входов нейрона последнего слоя будет производиться аналогично однослойному перцептрон, а значимости входов нейронов предыдущего слоя можно будет определить как произведение частных значимостей входов относительно выходов и значимостей выходов, как входов нейронов следующего слоя.

Для реализации такого подхода будем использовать процедуру, подобную применяемой в широко известном алгоритме обратного распространения ошибки (Error Backpropagation Method) [3,4], используемом при обучении НС. Алгоритм обратной оценки значимости признаков на основе НС имеет следующий вид.

Шаг 1. Установить счетчик слоев $q = M$.

Шаг 2. Для всех нейронов q -го слоя НС определить частные значимости их входов относительно их выходов (эти значимости будем называть частными, поскольку они не учитывают информативность выхода данного нейрона как входа нейронов следующего слоя):

$$z^*(x_i^{(q,j)*}) = \frac{|w_i^{(q,j)}|}{N_q \sum_{d=1} |w_d^{(q,j)}|}.$$

Шаг 3. Для всех нейронов q -го слоя определяют значимости их входов относительно выходов с учетом значимости выхода нейрона для нейронов следующего слоя:

$$z(x_i^{(q,j)*}) = z^*(x_i^{(q,j)*}) \sum_{p=1}^{N_{q+1}} z(x_j^{(q+1,p)*}).$$

Для нейронов M -го слоя $\sum_{p=1}^{N_{q+1}} z(x_j^{(q+1,p)*})$ принимается равной 1.

Шаг 4. Если $q > 1$, то установить $q = q - 1$ и перейти на шаг 2, в противном случае - перейти на шаг 5.

Шаг 5. Для всех признаков x_i $i = 1, 2, \dots, N$ определить оценки их значимости:

$$z(x_i^*) = \sum_{j=1}^{N_1} z(x_i^{(1,j)*}).$$

Оценки $z(x_i^*)$ в общем случае будут содержать определенную ошибку, однако, можно предположить, что на практике они позволят правильно определять качественную оценку значимости признаков, чего может быть вполне достаточно для многих приложений.

2.6. Алгоритм выращивания сети

Шаг 0. Инициализация: $\mu = 1$.

Шаг 1. Если количество нейронов в μ -ом слое N_μ меньше определенной величины β , то увеличить количество нейронов в μ -ом слое в соответствии с правилом $N_\mu = N^+(N_\mu)$, где $N^+(N_\mu)$ может определяться одним из выражений:

$$N^+(N_\mu) = \begin{cases} N_\mu + \gamma, & \text{если } (N_\mu + \gamma) < \beta; \\ \beta, & \text{если } (N_\mu + \gamma) \geq \beta, \end{cases}$$

$$N^+(N_\mu) = \begin{cases} \text{round}(N_\mu \gamma), \gamma > 1, & \text{если } \text{round}(N_\mu \gamma) < \beta; \\ \beta, & \text{если } \text{round}(N_\mu \gamma) \geq \beta, \end{cases}$$

где β и γ - заранее заданные константы, и перейти на шаг 4, в противном случае - перейти на шаг 2.

Шаг 2. Если $\mu < M$, то положить: $\mu = \mu + 1$ и перейти на шаг 4, в противном случае - на шаг 3.

Шаг 3. Если $\mu = M$ и $M < M_{\max}$, где M_{\max} - максимально допустимое количество слоев НС, то положить: $M = M + 1$, $N_M = 1$ и перейти на шаг 2, в противном случае - на шаг 4.

Шаг 4. Останов.

При задании констант β и γ следует учитывать следующие соображения:

- 1) при увеличении значения γ сеть будет расти быстрее;
- 2) значения β и γ должны быть больше нуля;
- 3) константа β должна быть целым числом;
- 4) если константа γ не является целым числом, то значение $N^+(N_\mu)$ должно округляться.
- 5) с увеличением значения β снижаются возможности роста сети.

2.7. Построение блока принятия решений

Основная задача блока принятия решений - обработка результата классификации для данного экземпляра и выдача соответствующих управляющих сигналов, определенных для данной ситуации.

Блок принятия решений можно строить различными способами:

- 1) на основе жестко заданных правил "Если (условие), то (действие)".
- 2) на основе нечеткой логики.
- 3) на основе нейросетевых моделей.

Последний вариант является более предпочтительным в тех случаях, когда нет заранее формализованных правил принятия решений, а есть только некоторое множество кортежей (пар данных), состоящих из описания ситуации и решения, которое целесообразно принять в такой ситуации.

3. Эксперименты и результаты

Для экспериментальной проверки предложенных моделей и методов на их основе проводилось решение прикладных задач технической диагностики для реальных изделий. Фрагменты результатов для отдельных этапов представлены в таблицах 1-3.

Как видно из табл. 1, с увеличением значения константы C сокращается количество признаков, с уменьшением - возрастает. Информативность отобранных признаков тем больше, чем больше значение C .

Из табл. 2 следует, что интеграция различных классификаторов с помощью НС Кохонена позволяет обеспечить на практике довольно высокий уровень надежности классификации.

Таблица 1 - Фрагмент результатов экспериментов по предварительному отбору признаков

| | | | | | | | | | |
|-------------------------------------------|----|-----|------|----|-----|------|----|-----|------|
| количество признаков в начальном наборе | 55 | 55 | 55 | 12 | 12 | 12 | 18 | 18 | 18 |
| константа значимости C | 1 | 1.5 | 0.75 | 1 | 1.5 | 0.75 | 1 | 1.5 | 0.75 |
| количество признаков в сокращенном наборе | 29 | 9 | 39 | 5 | 4 | 6 | 6 | 6 | 6 |

Таблица 2 - Фрагмент результатов экспериментов по построению и обучению гибридной системы классификаторов

| классификатор | вероятность правильной классификации в разных опытах (в %) | | | | | | | |
|---------------------------------------------------------|------------------------------------------------------------|------|------|-----|-----|-----|------|-----|
| | 100 | 100 | 85 | 97 | 98 | 100 | 82 | 97 |
| двухслойный перцептрон [2] | 100 | 100 | 85 | 97 | 98 | 100 | 82 | 97 |
| радиально-базисная сеть [3] | 100 | 100 | 100 | 100 | 98 | 100 | 100 | 93 |
| сеть LVQ [4] | 100 | 93,7 | 93,7 | 95 | 96 | 85 | 93,7 | 95 |
| интегрированная классификация на основе НС Кохонена [1] | 100 | 100 | 100 | 100 | 100 | 100 | 100 | 100 |

Таблица 3 - Фрагмент результатов экспериментов по отбору информативных признаков

| | | | |
|-------------------------------------------|----|----|-----|
| количество признаков в начальном наборе | 16 | 55 | 100 |
| количество признаков в сокращенном наборе | 11 | 15 | 20 |

Как видно из табл. 3, предлагаемый алгоритм оценки информативности и отбора признаков позволяет сокращать входное пространство классификаторов и тем самым снижает требования к ресурсам ЭВМ, увеличивает скорость работы классификаторов и позволяет уменьшить избыточность модели.

Предложенная методология построения диагностических систем на практике позволяет обеспечивать достаточно надежную классификацию для многих приложений. Наличие методов, позволяющих наращивать и сокращать структуру основных блоков системы в зависимости от сложности задачи, позволяет системе гибко адаптироваться к изменению задачи и разделению классов. Предложенная методология представляется достаточно перспективной для создания гибридных многоклассификаторных систем диагностики и может быть рекомендована для использования при решении задач распознавания, когда к надежности распознавания предъявляются высокие требования.

ПЕРЕЧЕНЬ ССЫЛОК

1. Дубровин В.И., Субботин С.А. Построение адаптивных систем классификации на основе нейронных сетей с латеральным торможением // *Радіоелектроніка. Інформатика. Управління*, 1999, № 2, С. 110-114
2. Дубровин В.И., Субботин С.А. Обобщенный градиентный алгоритм обучения многослойных нейронных сетей // *Електротехніка та електроенергетика*, 2000, № 1, С. 17-22
3. Дубровин В.И., Субботин С.А., Согорин А.А. Радиально-базисные нейронные сети в задачах технической диагностики // *Интернет, освіта, наука, друга міжнародна конференція ІОН-2000 / Збірник матеріалів конференції*. - Вінниця: Універсум - Вінниця, 2000. - С. 303-306.
4. *Neural Network Toolbox User Guide / Beale M., Demuth H.* - Natick: Mathworks, 1997. - 700 p.
5. *Справочник по прикладной статистике. В 2-х тт.: т. 1 / Под ред. Э. Ллойда, У. Ледермана, Ю.Н. Тюрина.* - М.: Финансы и статистика, 1989. - 510 с.
6. Горбань А.Н., Россиев Д.А. *Нейронные сети на персональном компьютере.* - Новосибирск: Наука. Сибирская издательская фирма РАН, 1996. - 276 с.
7. Ежов А.А., Шумский С.А. *Нейрокомпьютинг и его применения в экономике и бизнесе.* - М., 1998. - 222 с. (электронный документ <http://www.canopus.lpi.msk.su>).

АВТОРЫ НОМЕРА

Abdellatif Bourjij,

Dr., Nancy research centre for Automatic Control, Henri Poincare University, France

Mohamed Boutayeb,

Dr., Nancy research centre for Automatic Control, Henri Poincare University, France

Peana Hamburg,

Dr., Institut Arbeit und Technik Wissenschaftszentrum, Nordrein Westfalen, Gelsenkirchen, Germany

Stefan Balanica,

Dr., Computer Science Department, "Gh. Asachi" University of Iasi, Romania

Александрова Ирина Евгеньевна,

кандидат технических наук, доцент Национального технического университета "Харьковский политехнический институт"

Александрова Татьяна Евгеньевна,

аспирантка Национального технического университета "Харьковский политехнический институт"

Аникеев Максим Владимирович,

магистр, программист Таганрогского учебно-научного центра "Системы информационной безопасности"

Бабенко Людмила Климентьевна,

доктор технических наук, профессор Таганрогского государственного радиотехнического университета

Белоус Татьяна Ивановна,

аспирантка Национального технического университета Украины "Киевский политехнический институт"

Бичай Владимир Геннадиевич,

кандидат технических наук, старший научный сотрудник Запорожского государственного технического университета

Буков Валентин Николаевич,

доктор технических наук, заведующий кафедрой Военного Авиационного технического университета, г.Москва

Вершина Александр Иванович,

кандидат технических наук, доцент Запорожского государственного технического университета

Волгин Леонид Иванович,

доктор технических наук, профессор Ульяновского государственного технического университета

Габриэльян Дмитрий Давидович,

доктор технических наук, начальник научно-исследовательской лаборатории Ростовского военного института ракетных войск

Гостев Владимир Иванович,

доктор технических наук, начальник кафедры Киевского военного института управления и связи

Грядун Виктор Иванович,

старший преподаватель Запорожского государственного технического университета

Дорогов Александр Юрьевич,

кандидат технических наук, докторант Санкт-Петербургского государственного электротехнического университета

Дорошенко Юлия Николаевна,

аспирантка Запорожского государственного технического университета

Дубровин Валерий Иванович,

кандидат технических наук, докторант Запорожского государственного технического университета

Дулесов Александр Сергеевич,

кандидат технических наук, докторант, доцент Хакасского государственного университета им. Н.Ф.Катанова, г.Абакан

Жернаков Сергей Владимирович,

кандидат технических наук, доцент Уфимского государственного авиационного технического университета

Жураковский Юрий Павлович,

доктор технических наук, профессор Национального технического университета Украины "Киевский политехнический институт"

Калиниченко Владимир Николаевич,

аспирант института земледелия УАН, г.Киев

Карпуков Леонид Матвеевич,

кандидат технических наук, докторант Запорожского государственного технического университета

-
- Кику Андрей Григорьевич,*
кандидат технических наук, доцент Национального технического университета Украины "Киевский политехнический институт"
- Клячкин Владимир Николаевич*
кандидат технических наук, доцент Ульяновского государственного технического университета
- Корбецкий Орест Романович,*
кандидат технических наук, ассистент Государственного университета "Львівська політехніка"
- Косьянчук Владислав Викторович,*
кандидат технических наук, докторант Военного Авиационного технического университета, г.Москва
- Кузнецов Геннадий Васильевич,*
кандидат физико-математических наук, доцент Тульского государственного педагогического университета им. Л.Н. Толстого
- Легалов Александр Иванович*
кандидат технических наук, докторант Красноярского государственного технического университета
- Лесовой Иван Павлович,*
кандидат технических наук, директор Центра первичных сетей, г.Одесса
- Литвинов Евгений Николаевич,*
кандидат технических наук, доцент Национального технического университета Украины "Киевский политехнический институт"
- Лихоузов Алексей Сергеевич,*
аспирант Национального технического университета Украины "Киевский политехнический институт"
- Лихоузова Татьяна Анатольевна,*
аспирантка Национального технического университета Украины "Киевский политехнический институт"
- Лопин Вячеслав Николаевич,*
кандидат технических наук, доцент Курского государственного медицинского университета
- Макаревич Олег Борисович,*
доктор технических наук, заведующий кафедрой Таганрогского государственного радиотехнического университета
- Маркин Алексей Григорьевич,*
старший преподаватель Запорожского государственного технического университета
- Масуд М.Д. Мехеди,*
аспирант Харьковского государственного технического университета радиоэлектроники
- Мейстер Юрий Лазарович,*
начальник отдела казенного конструкторского бюро "Искра"
- Новицкий Николай Николаевич,*
доктор технических наук, заведующий лабораторией Института систем энергетики им. Л. А. Мелентьева (ИСЭМ) Сибирского отделения Российской академии наук, г.Иркутск
- Новотарский Михаил Анатольевич,*
кандидат технических наук, научный сотрудник Института математики НАН Украины
- Печенежский Дмитрий Сергеевич,*
аспирант Национального технического университета Украины "Киевский политехнический институт"
- Пица Дмитрий Макарович,*
доктор технических наук, проректор Запорожского государственного технического университета
- Пинчук Вячеслав Павлович,*
кандидат физико-математических наук, доцент Запорожского государственного технического университета
- Плаксин Сергей Викторович,*
кандидат физико-математических наук, старший научный сотрудник Института транспортных систем и технологий НАНУ "Трансмаг", г.Днепропетровск
- Потапенко Евгений Евгениевич,*
студент Запорожского государственного технического университета
- Потапенко Евгений Михайлович,*
доктор технических наук, профессор Запорожского государственного технического университета
- Привалов Владимир Николаевич,*
кандидат физико-математических наук, заведующий отделом Института транспортных систем и технологий НАНУ "Трансмаг", г.Днепропетровск

Прокофьев Вадим Павлович,

доктор технических наук, профессор, президент Международной общественной организации "Ассоциация ученых, специалистов и деятелей науки - "Наука"

Рябенко Владимир Николаевич,

кандидат технических наук, доцент Военного Авиационного технического университета, г.Москва

Сергеев Вячеслав Андреевич,

кандидат технических наук, проректор по международным связям Ульяновского государственного технического университета

Соколовский Иван Иванович,

кандидат физико-математических наук, ведущий научный сотрудник Института транспортных систем и технологий НАНУ "Трансмаг", г.Днепропетровск

Субботин Сергей Александрович,

преподаватель Запорожского государственного технического университета

Сысенко Ирина Юрьевна,

аспирантка Харьковского государственного технического университета радиоэлектроники

Теслюк Василий Николаевич,

кандидат технических наук, доцент Государственного университета "Львівська політехніка"

Токарев Вячеслав Вадимович,

научный сотрудник лаборатории Института систем энергетики им. Л. А. Мелентьева (ИСЭМ) Сибирского отделения Российской академии наук, г.Иркутск

Томашевский Валентин Николаевич,

доктор технических наук, профессор Национального технического университета Украины "Киевский политехнический институт"

Флора Валентин Данилович,

кандидат технических наук, доцент Запорожского государственного технического университета

Хак Х.М. Джахирул,

аспирант Харьковского государственного технического университета радиоэлектроники

Хаханов Владимир Иванович,

доктор технических наук, профессор Харьковского государственного технического университета радиоэлектроники

Цыганков Владимир Дмитриевич,

кандидат технических наук, начальник отдела ЦНИИТИ, г.Москва

Чиняев Юрий Петрович,

кандидат технических наук, начальник отдела Национального научно-исследовательского центра оборонных технологий и военной безопасности Украины

Чурин Андрей Евгеньевич,

адъюнкт Киевского военного института управления и связи

Шалагинова Зоя Ивановна,

кандидат технических наук, старший научный сотрудник лаборатории Института систем энергетики им. Л. А. Мелентьева (ИСЭМ) Сибирского отделения Российской академии наук, г.Иркутск

Шаповалов Виталий Павлович,

доктор физико-математических наук, профессор Запорожского государственного технического университета

Шацкий Виталий Валентинович,

кандидат технических наук, старший научный сотрудник Ростовского военного института ракетных войск

Шацкий Николай Витальевич,

кандидат технических наук, младший научный сотрудник Ростовского военного института ракетных войск

Шило Галина Николаевна,

аспирантка Запорожского государственного технического университета

Яшин Алексей Афанасьевич,

доктор технических наук, профессор, заместитель директора Государственного НИИ новых медицинских технологий, г.Тула