



# ВЕСТНИК

ЮЖНО-УРАЛЬСКОГО № 26 (159)  
ГОСУДАРСТВЕННОГО  
УНИВЕРСИТЕТА 2009

ISSN 1991-976X

СЕРИЯ

## «КОМПЬЮТЕРНЫЕ ТЕХНОЛОГИИ, УПРАВЛЕНИЕ, РАДИОЭЛЕКТРОНИКА»

Выпуск 10

Учредитель – ГОУ ВПО «Южно-Уральский государственный университет».

**Редакционная коллегия серии:**

д.т.н., профессор **Шестаков А.Л.**  
(*отв. редактор*);  
д.т.н., профессор **Казаринов Л.С.**  
(*зам. отв. редактора*);  
д.т.н., профессор **Щипицын А.Г.**  
(*зам. отв. редактора*);  
д.ф.-м.н., профессор, чл.-кор. РАН  
**Ушаков В.Н.**;  
д.т.н., профессор **Войтович Н.И.**;  
д.т.н., профессор **Карманов Ю.Т.**;  
д.т.н., профессор **Логиновский О.В.**;  
д.т.н., профессор **Лысов А.Н.**;  
д.ф.-м.н., профессор **Танана В.П.**;  
д.т.н., профессор **Тележкин В.Ф.**;  
д.ф.-м.н., профессор **Ухоботов В.И.**;  
к.т.н., доцент **Кафтанников И.Л.**;  
**Фокин Л.А.** (*техн. редактор*)

Серия основана в 2001 году.

Свидетельство о регистрации ПИ № ФС77-26455  
выдано 13 декабря 2006 г. Федеральной службой по  
надзору законодательства в сфере массовых комму-  
никаций и охране культурного наследия.

Журнал включен в Реферативный журнал и Базы  
данных ВИНТИ. Сведения о журнале ежегодно пуб-  
ликуются в международной справочной системе по  
периодическим и продолжающимся изданиям  
«Ulrich's Periodicals Directory».

Решением Высшей аттестационной комиссии  
Министерства образования и науки Российской Фе-  
дерации журнал включен в «Перечень ведущих ре-  
цензируемых научных журналов и изданий, в кото-  
рых должны быть опубликованы основные научные  
результаты диссертации на соискание ученой степени  
доктора и кандидата наук».

Подписной индекс 29008 в объединенном катало-  
ге «Пресса России».

Периодичность выхода – 2 номера в год.

## СОДЕРЖАНИЕ

ЛАВРОВ Н.Г., СТРАШНИН Е.Э., ШАЛИМОВ Л.Н. Применение концепции обратных задач динамики к проблеме управления угловым движением спускаемого аппарата .....	4
ШЕСТАКОВ А.Л., ЛАПИН А.П., ЛАПИНА Е.А. Алгоритмы выбора и обоснования моделей функций преобразования измерительных преобразователей давления .....	10
ГУДКОВ В.Ю. Математическая модель изображения отпечатка пальца на основе топологических векторов для линий .....	13
ГУДКОВ В.Ю. Оценка точности автоматического кодирования частных признаков изображений отпечатков пальцев .....	19
ЛЕБЕДЕВ А.А. Применение модельного прогнозирующего управления для регулирования активной мощности энергоблока тепловой электростанции .....	23
ВИШНЕВЕТСКАЯ Е.В., ГОРБУНОВ И.А. Использование нейросетевого моделирования для диагностики структуры эмоциональной компетентности ребенка по факторам, включающим физиологические параметры .....	26
ПАВЛОВСКАЯ О.О. Статические методы оценки надежности программного обеспечения .....	35
КАРМАНОВ Ю.Т., НИКОЛАЕВ А.Н. Непараметрическое обнаружение узкополосного радиосигнала с неизвестной несущей частотой на фоне гауссовских шумов.....	38
КАРМАНОВ Ю.Т., НЕПОМНЯЩИЙ Г.А. Способ защиты РЛС со сложным сигналом от имитирующей помехи .....	41
КАРМАНОВ М.Л. Протокол маршрутизации для ad-hoc сетей .....	47
РОЖКОВ С.А. Терминальная система WTPRO: архитектура, функциональность, схема применения .....	52
ЗБИЦКИЙ П.В. Модель метаморфного преобразования исполняемого кода .....	57
САПОЖНИКОВ А.А. Наблюдение аномалий в сети .....	62
ШУРЫГИН А.Н. Имитационное моделирование технического обслуживания основных производственных фондов .....	66
ГОЙТИНА Е.В., ШНАЙДЕР Д.А. Оперативное управление экономичностью водяных тепловых сетей на основе макро моделирования .....	71
ВОЛОВИЧ Г.И., ТОПОЛЬСКАЯ И.Г., ТОПОЛЬСКИЙ Д.В. Моделирование работы электронного измерителя механического момента электрических машин в системах автоматического управления .....	77
КЕЛЛЕР А.В. Алгоритм численного решения задачи Шоуолтера–Сидорова для систем леонтьевского типа .....	82
ПОПОВ А.Е. Применение методов интервального анализа при калибровке преобразователей давления .....	87

## CONTENTS

LAVROV N.G., STRASHININ E.E., SHALIMOV L.N. Back dynamic concept applied to re-entry vehicle attitude control problem .....	4
SHESTAKOV A.L., LAPIN A.P., LAPINA E.A. Choice algorithms and model substantiation of measuring pressure converters conversion functions .....	10
GUDKOV V.U. Mathematical model of fingerprint based on topology vectors for lines .....	13
GUDKOV V.U. Accuracy rating of automatic minutiae recognition of fingerprints .....	19
LEBEDEV A.A. Model predictive control of the thermal power plant generating unit output .....	23
VISHNEVETSKAYA E.V., GORBUNOV I.A. Use of neural modeling structure for diagnostics of emotional competence of a child with factors including physiological parameters .....	26
PAVLOVSKAYA O.O. Static methods of assessment of software .....	35
KARMANOV Y.T., NIKOLAEV A.N. Nonparametric narrowband detection of the signal with unknown carrier frequency with a background gaussian noise .....	38
KARMANOV Y.T., NEPOMNYASHIY G.A. One way to protect the radar of a complex signals from simulating interference .....	41
KARMANOV M.L. Routing in ad-hoc networks .....	47
ROZKHOV S.A. Terminal system WTPRO: architecture, functionality, application circuit ...	52
ZBITSKIY P.V. Executable code metamorphic transformation model .....	57
SAPOZHNIKOV A.A. Network anomalies observation .....	62
SHURYGIN A.N. Simulation of maintenance of fixed assets .....	66
GOYTINA E.V., SHNAYDER D.A. Operational management of efficient water heating networks based on macromodelling .....	71
VOLOVICH G.I., TOPOLSKAYA I.I., TOPOLSKIY D.V. Simulation of the electronic measuring device of the mechanical moment of electric drives in automatic control systems .....	77
Keller A.V. Algorithm of numerical solution of Shoulter–Sidorov problem for Leontiev-type systems .....	82
POPOV A.E. Application of interval analysis in calibration of pressure transducers .....	87

## ПРИМЕНЕНИЕ КОНЦЕПЦИИ ОБРАТНЫХ ЗАДАЧ ДИНАМИКИ К ПРОБЛЕМЕ УПРАВЛЕНИЯ УГЛОВЫМ ДВИЖЕНИЕМ СПУСКАЕМОГО АППАРАТА

*Н.Г. Лавров, Е.Э. Страшинин, Л.Н. Шалимов*

## BACK DYNAMIC CONCEPT APPLIED TO RE-ENTRY VEHICLE ATTITUDE CONTROL PROBLEM

*N.G. Lavrov, E.E. Strashinin, L.N. Shalimov*

В работе рассматривается подход к синтезу систем управления (СУ) на основе концепции обратных задач динамики (ОЗД). Описывается ставший уже классическим подход, основанный на ОЗД. Показаны особенности применения концепции ОЗД к синтезу СУ спускаемых аппаратов (СА). Предлагается модифицировать классический подход ОЗД для реализации СУ маневрирующего СА. Приведены результаты моделирования, подтверждающие эффективность модификации.

*Ключевые слова: система управления, спускаемый аппарат, обратные задачи динамики, система стабилизации, эталонные модели, адаптация, моделирование.*

The control system synthesis approach based on back dynamic structures is considered. Classical back dynamic structure is described. Re-entry vehicle control system features are shown. Back dynamic structure classical approach is suggested to be modified. Modeling results are shown.

*Keywords: control system, re-entry vehicle, back dynamic structures, stabilization system, etalon model, adaptation, simulation.*

### Введение

Развитие науки и техники постоянно требует разработки и создания все новых и новых летательных аппаратов (ЛА) различных типов с широким набором целевых задач. Выполнение целевых задач предъявляет к ЛА и бортовой системе управления требования (зачастую очень жесткие) по реализации заданных характеристик управляемого движения. В ракетно-космической технике (РКТ) можно выделить три класса летательных аппаратов, исходя из динамики полета, это: космические аппараты (КА), спускаемые аппараты различного назначения (СА) и ракеты-носители (РН). Аппараты, принадлежащие к разным классам, сильно отличаются друг от друга по целому ряду параметров и в первую очередь по целям и решаемым задачам, откуда и вытекают все осталь-

ные различия. Тем не менее для всех упомянутых летательных аппаратов, независимо от принадлежности к конкретному классу, существует общий подход к математическому описанию их движения, который основан на широко известных уравнениях динамики:

$$\ddot{x} = f_x(t, x, \dot{x}, u_x),$$

$$\ddot{\varphi} = f_\varphi(t, \varphi, \dot{\varphi}, u_\varphi).$$

Применительно к СА, совершающему маневрирующий спуск в плотных слоях атмосферы Земли, полученная система является существенно нелинейной и нестационарной, что серьезно затрудняет синтез алгоритмов СУ классическими методами. Исследование альтернативных методов синтеза алгоритмов СУ дало результат, позволяющий говорить о том, что с многих позиций наиболее

---

**Лавров Никита Георгиевич** – начальник сектора ФГУП «НПО автоматики им. Н.А. Семихатова»; ngl@npoa.ru

**Страшинин Евгений Эрастович** – ведущий науч. сотрудник ФГУП «НПО автоматики им. Н.А. Семихатова»; ngl@npoa.ru

**Шалимов Леонид Николаевич** – генеральный директор ФГУП «НПО автоматики им. Н.А. Семихатова»; ngl@npoa.ru

---

**Lavrov Nikita Georgievich** – head of sector of Science & Production Association of Automatics; ngl@npoa.ru

**Strashinin Eugeny Erastovich** – lead scientific staff of Science & Production Association of Automatics; ngl@npoa.ru

**Shalimov Leonid Nikolaevich** – CEO of Science & Production Association of Automatics; ngl@npoa.ru

эффективным является подход к синтезу, основанный на концепции обратных задач динамики. Анализ работ Петрова Б.Н., Крутько П.Д., Попова Е.П. в области применения ОЗД к синтезу алгоритмов СУ [1–6] показал, что результирующие алгоритмы управления обладают нетрадиционной структурой, придающей им следующие важные свойства:

- слабую чувствительность к параметрическим и координатным возмущениям;
- естественную адаптивность;
- практически единую структуру как для линейных, так и для нелинейных систем;
- квазиоптимальность.

### **1. Классический подход к обратным задачам динамики**

Исторически обратные задачи динамики возникли в рамках теоретической механики и применяются для вычисления сил действующих на объект по известной траектории его движения. Применительно к задаче синтеза алгоритмов управления в соответствии с классической концепцией ОЗД можно выделить следующие основные этапы:

- формирование эталонных траекторий движения (эталонных моделей);
- формирование управляющих ускорений и сил;
- формирование уставок на отклонение органов управления.

Эталонная модель, которая представляет собой реализацию эталонной фазовой траектории, может быть задана в явной форме, в виде функции времени:

$$y(t) = c_1 e^{\lambda_1 t} + c_2 e^{\lambda_2 t},$$

или в неявной, когда задается дифференциальное уравнение, решением которого является функция  $y(t)$ :

$$g(t, y, \dot{y}, \dots) = 0.$$

Обычно эталонную модель, реализующую эталонную модель движения механической системы, удобно задать в виде линейного дифференциального уравнения требуемого порядка (например, второго) с постоянными коэффициентами:

$$T_{\xi}^2 \ddot{y} + 2T_{\xi} \xi_{\xi} \dot{y} + y = v,$$

где  $v$  – уставка;  $T_{\xi}, \xi_{\xi}$  – параметры, характеризующие динамику эталонной модели, определяются разработчиком СУ;  $\dot{y}, y$  – эталонные фазовые координаты. В принципе, эталонные модели могут быть и нелинейными, и нестационарными. В этих случаях разработчик СУ должен четко представлять влияние упомянутых особенностей на эталонные процессы. Одной из возможных реализаций нелинейных эталонных моделей может стать простейшая замкнутая система с линейным объектом и квазиоптимальным управлением.

Далее рассмотрим способ формирования управляющих ускорений. Для этого кроме эталонной модели зададим закон изменения ошибки по

параметрам управления. Этот закон наиболее удобно задать линейным дифференциальным уравнением требуемого порядка с постоянными коэффициентами:

$$\begin{aligned} e &= y - x, \\ \dot{e} &= \dot{y} - \dot{x}, \\ \ddot{e} &= \ddot{y} - \ddot{x}, \\ \ddot{e} + k_2 \dot{e} + k_1 e &= 0. \end{aligned}$$

Коэффициенты закона выбираются с учетом динамики эталонных моделей. Не исключаются случаи использования нелинейного или нестационарного закона изменения ошибки. С учетом этого закон управления будет выглядеть следующим образом:

$$\ddot{x}^{\mathcal{J}} = \ddot{y} - \ddot{e} = \ddot{y} + k_2 \dot{e} + k_1 e,$$

где  $\ddot{x}^{\mathcal{J}}$  – управляющее ускорение. Естественно, что программная составляющая, рассчитанная по эталонной модели системы, не будет обеспечивать нулевые ошибки при работе с реальным объектом. Поэтому программная составляющая дополняется обратной связью, которая обеспечивает малую динамическую ошибку отклонения от программной траектории.

С учетом значения  $\ddot{x}^{\mathcal{J}}$  и на основе уравнения, описывающего движение объекта управления (ОУ):

$$\ddot{x} + q(x) = F(x, \dot{x}, u), \quad (1)$$

где  $q(x)$  – некоторое переносное ускорение, а  $F$  – обобщенная внешняя сила, вычислим управляющую силу  $F^{\mathcal{J}}$ :

$$F^{\mathcal{J}} = \ddot{x}^{\mathcal{J}} + q(x).$$

Вид уравнения объекта (1) является характерным при описании движения твердого тела под действием внешней силы в подвижной системе координат (СК).

Следующим шагом по величине управляющей силы вычисляем уставку на отклонение органов управления ( $u^{\mathcal{J}}$ ). Естественно, что уставка должна удовлетворять следующему равенству:

$$F(x, \dot{x}, u^{\mathcal{J}}) = F^{\mathcal{J}},$$

которое в общем случае является нелинейным нестационарным алгебраическим уравнением относительно  $u^{\mathcal{J}}$ . Для этого уравнения за редким исключением не может быть получено аналитическое решение. Поэтому для его решения в общем случае целесообразно применять численные методы. Наиболее органично в концепцию ОЗД вписывается метод, в котором в качестве приближенного значения  $u^{\mathcal{J}}$  используется численное решение нелинейного нестационарного дифференциального уравнения [2–4]:

$$\begin{aligned} \dot{u}^{\mathcal{J}} &= k \int (F^{\mathcal{J}} - F(x, \dot{x}, u^{\mathcal{J}})) dt, \\ \dot{u}^{\mathcal{J}} + kF(x, \dot{x}, u^{\mathcal{J}}) &= kF^{\mathcal{J}}. \end{aligned} \quad (2)$$

При выполнении определенных ограничений на уравнение (1) полученному закону управления соответствует следующая структурная схема [2, 3], представленная на рис. 1.

В этом случае закон управления можно записать в достаточно простой форме

$$u^{\text{ж}} = k \int (\dot{x}^{\text{ж}} - \ddot{x}) dt = k \left( \dot{y} - \dot{x} + k_1 (y - x) + k_0 \int (y - x) dt \right). \quad (3)$$

Таким образом, выражение (3) реализует достаточно простой и компактный алгоритм управления движением. Аналогичное выражение для закона управления (3) можно получить и другим способом, основанным на минимизации функционала, характеризующего отклонение энергии ускорения объекта управления от энергии ускорения эталонной модели [2]:

$$I(u) = \frac{1}{2} (\dot{y} - \ddot{x})^2.$$

Однако упомянутые выше ограничения, наложенные на уравнение (1), исходя из которых получен закон управления (3), не могут быть реализованы для реального объекта управления. Это заставляет отойти от классической формы закона управления и некоторым образом модифицировать подход, предлагаемый концепцией ОЗД.

## 2. Применение модифицированного подхода ОЗД к управлению полетом СА

Рассмотрим построение модифицированного закона управления на основе концепции ОЗД применительно к управлению пространственной ориентацией СА, совершающего спуск в плотных слоях атмосферы Земли. Необходимо отметить, что многие СА и рассматриваемый в частности не обладают полной управляемостью угловым движением. Это является следствием стремления к минимизации объемно-массовых характеристик исполнительных органов. Рассматриваемый СА характерен тем, что его аэродинамическая схема позволяет одновременно вести независимое управление только в двух каналах из трех и реализует связь между управляющими моментами. По результатам анализа характеристик СА в каждом из

каналов и требуемых типов траекторий, в качестве каналов управления выбраны тангаж и крен.

Система уравнений, описывающая движение СА в атмосфере, в связанной СК выглядит следующим образом:

$$\begin{aligned} \dot{\bar{\omega}}_1 &= J^{-1} (\bar{M}_1 + J \bar{\omega}_1 \times \bar{\omega}_1), \\ \dot{\bar{V}}_1 &= \frac{1}{m} (\bar{F}_1 + \bar{V}_1 \times \bar{\omega}_1) + \bar{G}. \end{aligned}$$

При этом  $F_1$  и  $M_1$  являются аэродинамическими силой и моментом, записанными в связанной СК. Коэффициенты аэродинамических моментов и сил для тел с ярко выраженной осевой асимметрией, как правило, задаются с помощью разложения в ряд в цилиндрической связанной СК на основе таблиц. Для рассматриваемого СА объем таблиц, описывающих только лишь стационарные характеристики, достигает порядка 1500 коэффициентов.

Структурная схема системы управления пространственной ориентацией в соответствии с концепцией ОЗД, представлена на рис. 2, где  $M_1^{\text{ж}}$  – уставка по управляющему моменту;  $\delta^{\text{ж}}$  – уставка на отклонение органов управления ориентацией СА;  $i_k$  – командный ток на рулевые машины (РМ).

Структурная схема блока управления пространственной ориентацией в соответствии с рассмотренной классической концепцией ОЗД представлена на рис. 3, где приняты следующие обозначения:  $y_i, \dot{y}_i$  – фазовые координаты эталонных моделей каналов управления;  $\ddot{\varphi}^{\text{ж}}, \ddot{\vartheta}^{\text{ж}}, \ddot{\psi}^{\text{ж}}$  – управляющие ускорения;  $\varphi, \dot{\varphi}, \vartheta, \dot{\vartheta}$  – величины параметров управления ( $\varphi$  – оценка угла скоростного крена,  $\vartheta$  – оценка угла атаки) и их производных.

До этого момента рассматриваемая СУ пространственной ориентацией СА полностью соответствует рассмотренной классической концепции ОЗД. При вычислении углов отклонения органов управления для реального СА уже невозможно реализовать замыкание через объект и применить контур по ускорению (см. рис. 1). Невозможно это сделать, по крайней мере, по двум причинам: во-первых, в этом контуре появляется рулевой привод, вносящий задержки инерционности и допол-

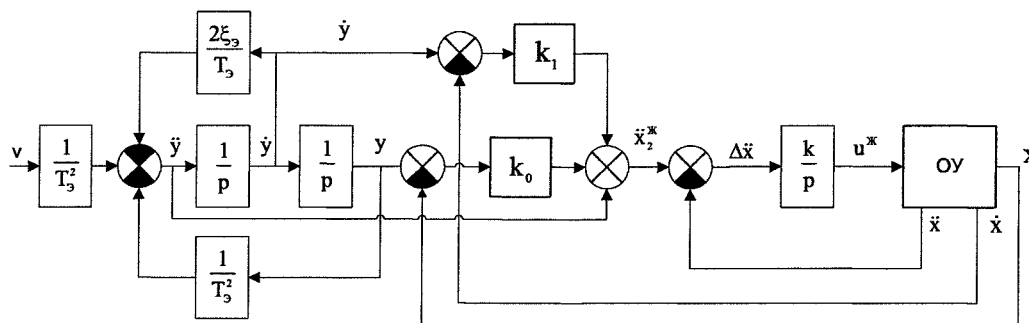


Рис. 1. Преобразованная структурная схема алгоритма управления

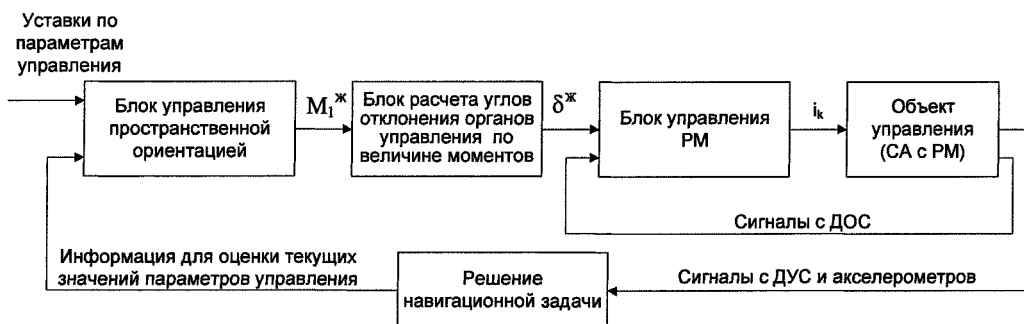


Рис. 2. Структура СУ пространственной ориентацией

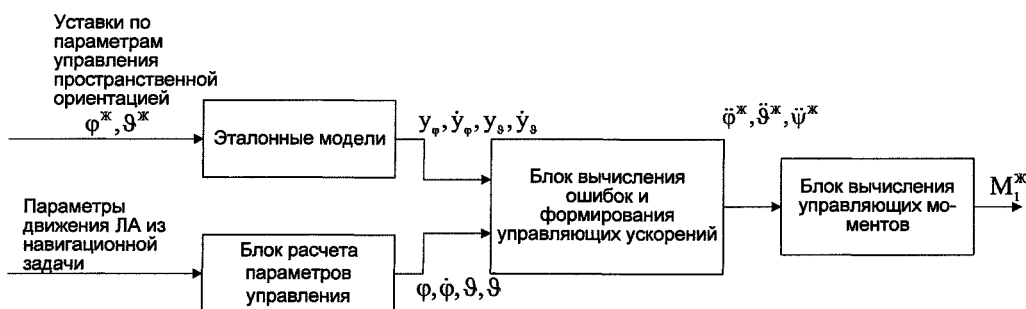


Рис. 3. Структурная схема блока управления пространственной ориентацией

нительные нелинейности, а во-вторых, необходимо наличие датчиков, измеряющих угловые ускорения, что представляет собой отдельную проблему. Для решения этой проблемы предложено реализовывать решение уравнения (2) алгоритмически в СУ без замыкания через объект. На рис. 4 приведена структурная схема контура для алгоритмического решения уравнения (2) в бортовых алгоритмах СУ, работающей в реальном времени. В качестве функции  $F$ , описывающей зависимость внешней силы от фазовых координат, используются аэродинамические характеристики (АДХ) СА.

На рис. 4 приняты следующие обозначения:  $\vartheta$  – величина параметра управления, используемая в качестве оценки угла атаки;  $M$  – число Маха;  $q$  – величина скоростного напора;  $\hat{M}_1$  – оценка аэродинамического момента. Дополнительно к интегратору появляется схема учета приоритетов каналов управления и связанная с ней

реализация ограничений. Этот механизм необходим для возможности вести управление СА с учетом завязки управляющих моментов и частичной управляемости, вызванной особенностями аэродинамической схемы. По результатам исследования характеристик СА в каждом из каналов управления наиболее приоритетным каналом выбран канал крена.

При функционировании контура (см. рис. 4) в реальном времени невозможно использовать полные АДХ СА из-за большого их объема. Это обстоятельство требует сформировать упрощенную эквивалентную модель АДХ. Процедура упрощения полных АДХ проводится путем исключения из расчета АДХ тех составляющих, влияние которых мало (например, динамические моменты) или не существенно (например, моменты в рыскании), а также путем сокращения интерполяционной сетки по отдельным параметрам. Адекватность проведенного упрощения показана путем моделирова-

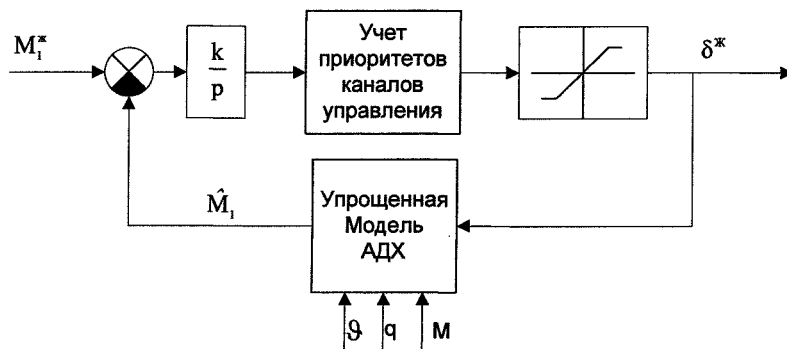


Рис. 4. Структурная схема расчета углов отклонения органов управления

ния работы СУ пространственной ориентации в целом. Результаты моделирования приводятся ниже. В итоге упрощенная модель АДХ включает в себя таблицы, содержащие порядка 100 коэффициентов. Это позволяет не только сократить объем вычислений, но и сделать систему менее чувствительной к изменению отдельных характеристик СА.

Несмотря на малую чувствительность системы к изменению отдельных характеристик, существуют параметры СА, малое изменение которых приводит к существенному изменению управляемости СА. И эти существенные параметры СА имеют как начальные разбросы, так и сильно изменяются в процессе движения СА вследствие обгара и уноса теплозащитного покрытия (ТЗП), а также расхода рабочего тела РП. К этим существенным параметрам можно отнести те, которые определяют балансировочную зависимость СА и возмущения в канале крена. В итоге эти параметрические возмущения оказывают заметное влияние на работу СУ пространственной ориентацией и приводят к большой динамической ошибке как по угловым параметрам движения (порядка 30–50°), так и по линейным параметрам в процессе наведения СА. При этом итоговый динамический промах наведения возрастает в десятки раз. Для решения этой проблемы авторами предложено в рамках системы, построенной в соответствии с концепцией ОЗД, реализовать дополнительный контур адаптации упрощенной модели АДХ к изменениям существенных параметров СА. Структура, предлагаемая ОЗД, позволяет сохранить физический смысл используемых в контурах СУ величин. Это дает возможность достаточно легко выделить параметры упрощенной модели АДХ, за счет изменения которых возможна эффективная адаптация СУ. Оказывается, что в качестве подстраиваемых параметров упрощенной модели наиболее целесообразно использовать эквивалентное относительное положение центра масс СА, которое возможно получить, анализируя оценки аэродинамических моментов. Исходя из этого, разработана система адаптации СУ, которая позволяет в процессе полета в реальном времени проводить оценку эквивалентного положения центра масс СА и производить подстройку упрощенной модели АДХ. Структурная схема системы адаптации СУ показана на рис. 5.

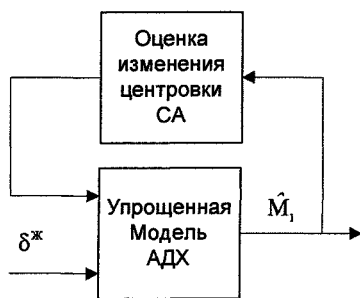


Рис. 5. Структурная схема включения системы адаптации СУ к изменению параметров СА

### 3. Результаты моделирования полета СА

Далее рассмотрим результаты моделирования работы СУ пространственной ориентацией при полете СА на отдельном участке траектории с учетом обгара и уноса ТЗП, а также расхода рабочего тела РП. Моделирование проводилось с использованием среды Matlab в пакете Simulink [7]. На рис. 6, 7 представлены переходные процессы по параметрам управления пространственной ориентацией при работе системы адаптации СУ к изменению характеристик СА, а на рис. 8 – угол скоростного крена без адаптации.

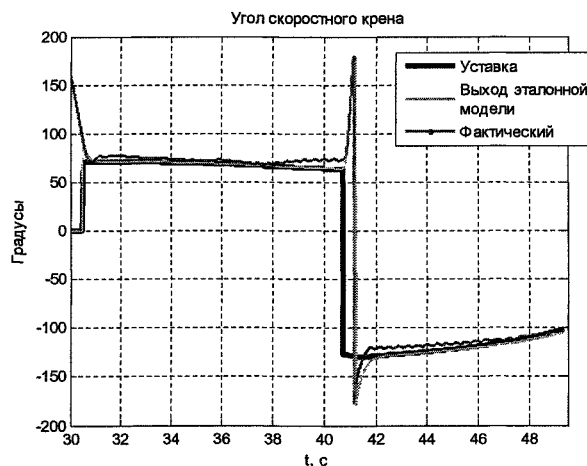


Рис. 6. Угол скоростного крена с адаптацией

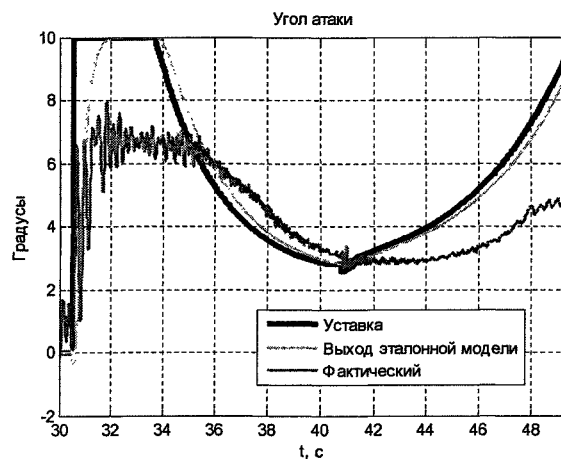


Рис. 7. Угол атаки с адаптацией

По результатам моделирования достигается заданная точность реализации траектории и терминальных условий. Результаты, приведенные на рис. 6, 7, показывают возможность управлять угловым движением СА на рассматриваемом участке траектории с учетом изменения характеристик СА. Процессы по углу атаки в случае с работающей системой адаптации и без нее расходятся незначительно. В канале крена (рис. 8) без системы адаптации СУ реализуется достаточно большая динамическая ошибка по углу скоростного крена, вызванная поперечным смещением центровки СА.



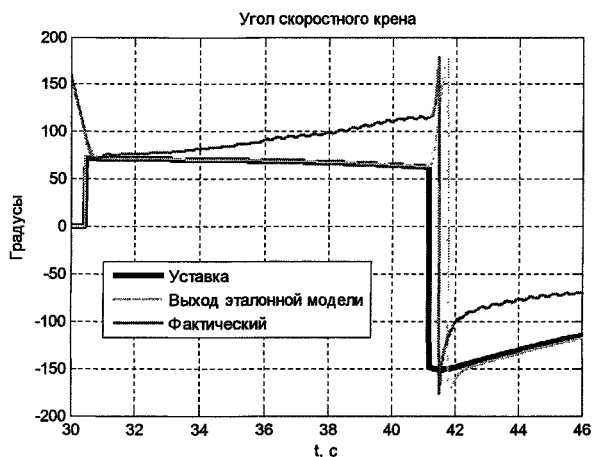


Рис. 8. Угол скоростного крена без адаптации

Отдельно стоит сказать, что отличие фактического угла атаки от уставки (см. рис. 7) вызвано особенностями аэродинамической схемы СА: смещением центровки вдоль продольной оси СА (см. рис. 8), что сокращает максимальные реализуемые углы атаки, и завязкой между управляющими моментами в каналах крена и тангажа. Реализованная схема приоритетов и система адаптации СУ позволяют стабилизировать СА в крене и тангаже в условиях существенного смещения центра масс СА (см. рис. 8) и минимизировать динамическую ошибку в канале крена. При этом естественно повышается нагрузка органов управления и осознанно приходится жертвовать динамической точностью обработки угла атаки в канале тангажа, так как она оказывает меньшее влияние на точность наведения. При высокой динамической точности в канале крена (см. рис. 6), несмотря на заметную ошибку в канале тангажа (см. рис. 7) система управления обеспечивает высокую итоговую точность управления линейным движением СА по траектории, которая оказывается недостижимой при отсутствии системы адаптации СУ к изменению характеристик СА.

#### Заключение

По результатам разработки СУ пространственной ориентацией СА на основе концепции обратных задач динамики в первую очередь выделяем

реализуемость и работоспособность такой системы. Полученная структура позволяет эффективно реализовать систему, которая действительно обладает слабой чувствительностью к параметрическим и координатным возмущениям, а также возможностью эффективно создать на ее основе адаптивную СУ. В заключение необходимо отметить, что приведенные в статье наработки в области ОЗД успешно применяются в СУ различных СА, в том числе и тех, которые имеют жесткие полетные условия [8].

#### Литература

1. Петров, Б.Н. Построение алгоритмов управления как обратная задача динамики / Б.Н. Петров, П.Д. Крутько, Е.П. Попов // ДАН СССР. – 1979. – Т. 247, № 5.
2. Крутько, П.Д. Обратные задачи динамики в теории автоматического управления / П.Д. Крутько. – М.: Машиностроение, 2004.
3. Крутько, П.Д. Обратные задачи динамики управляемых систем. Нелинейные модели / П.Д. Крутько. – М.: Наука, 1988.
4. Крутько, П.Д. Обратные задачи динамики управляемых систем. Линейные модели / П.Д. Крутько. – М.: Наука, 1987.
5. Методы классической и современной теории автоматического управления: учеб.: в 5 т. / под ред. К.А. Пупкова, Н.Д. Егунова. – М.: Изд-во МГТУ им. Н.Э. Баумана, 2004.
6. Крутько, П.Д. Робастно устойчивые структуры управляемых систем высокой динамической точности. Алгоритмы и динамика управления движением модельных объектов / П.Д. Крутько // Изв. РАН. ТуСУ. – 2005. – № 6.
7. Лавров, Н.Г. Моделирование движения ЛА в среде MATLAB / Н.Г. Лавров // Ракетно-космическая техника: Системы управления ракетных комплексов. – Екатеринбург: НПОА, 2006. – С. 148–158.
8. Лавров, Н.Г. Система управления пространственной ориентацией высокоманевренного летательного аппарата / Н.Г. Лавров // Ракетно-космическая техника: Системы управления ракетных комплексов. – Екатеринбург: НПОА, 2008. – С. 16–33.

Поступила в редакцию 4 мая 2009 г.

## АЛГОРИТМЫ ВЫБОРА И ОБОСНОВАНИЯ МОДЕЛЕЙ ФУНКЦИЙ ПРЕОБРАЗОВАНИЯ ИЗМЕРИТЕЛЬНЫХ ПРЕОБРАЗОВАТЕЛЕЙ ДАВЛЕНИЯ

*А.Л. Шестаков, А.П. Лапин, Е.А. Лапина*

## CHOICE ALGORITHMS AND MODEL SUBSTANTIATION OF MEASURING PRESSURE CONVERTERS CONVERSION FUNCTIONS

*A.L. Shestakov, A.P. Lapin, E.A. Lapina*

Использование алгоритмических методов коррекции погрешности в интеллектуальных датчиках давления предполагает знание функции преобразования измерительного преобразователя. Авторами разработаны алгоритмы выбора и обоснования моделей функций преобразования. Применение этих алгоритмов дает возможность находить наилучшие функции преобразования для датчиков давления разного конструктивного исполнения.

*Ключевые слова: функция преобразования, интеллектуальные датчики давления.*

Usage of algorithmic methods of correction of an error in intellectual gauges of pressure assumes knowledge of function of conversion of the measuring converter. Authors develop choice algorithms and model substantiation of measuring pressure converters conversion function. Application of these algorithms gives the chance to find the best functions of conversion for gauges of pressure of different designs.

*Keywords: conversion function, intellectual gauges of pressure.*

### Введение

Современные научные исследования и промышленное производство предъявляют всё более высокие требования к точности измерения физических величин, в частности, давления. Широкое распространение получили измерительные преобразователи давления на основе тензорезисторов [1]. Однако, данным преобразователям присущи недостатки в виде значительных погрешностей от нелинейности функции преобразования и сильной температурной зависимости.

Появление интеллектуальных датчиков давления [2] позволило перейти от схемотехнических решений к алгоритмическим методам коррекции вышеуказанных погрешностей. Определяющую роль в этом играет правильный выбор функции преобразования (ФП) [3] измерительных преобразователей (ИП).

### 1. Постановка задачи

Основополагающие работы в этой области [4, 5] не дают однозначных путей решения проблемы. Анализ современных научных публикаций [6–8] показал, что в настоящее время при разработке и создании современных датчиков давления, обладающих высоким классом точности, также имеют место различные подходы.

Вся сложность возникающих при этом задач заключается в следующем:

- измерительные преобразователи различных конструкций, имеют различные ФП;
- внутри партии измерительных преобразователей одной и той же конструкции, вследствие различного рода отклонений в технологическом процессе изготовления ИП, имеют место отличающиеся друг от друга ФП отдельных ИП, т.е. речь идет об индивидуальных ФП измерительных преобразователей;

---

**Шестаков Александр Леонидович** – д-р техн. наук, профессор, ректор ЮУрГУ; [admin@urc.ac.ru](mailto:admin@urc.ac.ru)

**Лапин Андрей Павлович** – канд. техн. наук, доцент кафедры «Информационно-измерительная техника» ЮУрГУ; [a\\_lapin@mail.ru](mailto:a_lapin@mail.ru)

**Лапина Екатерина Андреевна** – аспирант кафедры «Информационно-измерительная техника» ЮУрГУ; [a\\_lapin@mail.ru](mailto:a_lapin@mail.ru)

---

**Shestakov Aleksandr Leonidovich** – PhD, professor, rector of SUSU; [admin@urc.ac.ru](mailto:admin@urc.ac.ru)

**Lapin Andrey Pavlovich** – PhD, associate professor of the Equipment for information and measuring department of SUSU; [a\\_lapin@mail.ru](mailto:a_lapin@mail.ru)

**Lapina Ekaterina Andreevna** – postgraduate student of the Equipment for information and measuring department of SUSU; [a\\_lapin@mail.ru](mailto:a_lapin@mail.ru)

– нет единообразных требований к функции преобразования ИП;

– отсутствует единая методика нахождения функции преобразования измерительных преобразователей.

Особую актуальность эта проблема приобрела на крупных приборостроительных предприятиях при промышленном производстве больших партий датчиков давления.

Авторы считают, что функция преобразования должна отвечать следующим требованиям.

1. Функция преобразования строится в виде двухфакторной модели, входными факторами которой выступают давление, поступающее на вход преобразователя, и температура среды, в которой находится преобразователь.

2. Модель ФП должна обеспечивать нужный класс точности датчика давления и иметь при этом 2- или 3-кратный метрологический запас по приведенной погрешности.

3. Модель ФП должна быть как можно более простой по своей структуре. Сложность модели ограничивается числом точек по давлению и температуре, задаваемых при проведении испытаний датчика давления при построении его функции преобразования.

4. Функция преобразования должна отвечать требованиям по верификации, заложенным в конструкторскую документацию на датчики давления.

В качестве основы для изучения ФП измерительных преобразователей давления авторы использовали «обратную» градуировочную характеристику ИП в виде двухфакторной математической модели 5-й степени следующего вида:

$$Y = \beta_0 + \beta_1 t + \beta_2 t^2 + \beta_3 t^3 + \beta_4 t^4 + \beta_5 t^5 + \beta_6 p + \beta_7 p t + \beta_8 p t^2 + \beta_9 p t^3 + \beta_{10} p t^4 + \beta_{11} p^2 + \beta_{12} p^2 t + \beta_{13} p^2 t^2 + \beta_{14} p^2 t^3 + \beta_{15} p^3 + \beta_{16} p^3 t + \beta_{17} p^3 t^2 + \beta_{18} p^4 + \beta_{19} p^4 t + \beta_{20} p^5, \quad (1)$$

где  $Y$  – рассчитанное значение измеренного давления;  $p$  – выходной код АЦП преобразователя по давлению;  $t$  – выходной код АЦП преобразователя по температуре;  $\beta_0, \beta_1, \beta_2, \dots, \beta_{20}$  – коэффициенты функции преобразования (ФП), определяемые по методу наименьших квадратов. Эта модель может быть реализована для шести температурных точек

и шести точек по давлению, что отвечает плану стандартных приемосдаточных испытаний датчиков давления.

Учитывая собственный опыт исследовательской работы в этом направлении [9–13], авторам настоящей статьи видится такой путь решения проблемы.

1. Провести исследования ограниченного количества специально отобранных ИП *определенной конструкции* и на основе этой информации *получить «общую» (единую) ФП, являющуюся наилучшей для ИП данной конструкции*, т.е. создать библиотеку ФП, в которой для каждой конструкции ИП, выпускаемой промышленным предприятием, имеется своя, наилучшая, функция преобразования [14–15].

2. Процедуру нахождения «индивидуальных» ФП ИП, *являющихся наилучшими для конкретных экземпляров ИП*, авторы предлагают свести к *нахождению индивидуальных коэффициентов «общей» ФП*, взятой из библиотеки для данной конструкции ИП, для каждого преобразователя в партии, при градуировке измерительных преобразователей.

## 2. Алгоритмы выбора и обоснования модели функции преобразования

Процедура получения «общей» (единой) ФП для датчиков давления (ДД) разного конструктивного исполнения реализуется в виде последовательно выполняемых этапов в соответствии с таблицей.

Отметим особенности выполнения этапов, представленных в таблице.

Этапы 1, 2 включают исследования партии ИП датчиков одной конструкции, проводимые путем градуировки и верификации преобразователей, и их классификацию по разработанным авторами алгоритмам с целью выделения группы ИП, обладающих наилучшими (наименьшими) показателями погрешности от нелинейности и температурной погрешности.

Этапы 3, 4 включают проведение приемосдаточных испытаний партии ДД, собранных на основе исследованной партии ИП, и анализ их метрологических характеристик, что позволяет нам

Процедура получения «общей» (единой) ФП для ДД, реализованная в виде отдельных этапов

№	Название этапа	Содержание этапа	Требования (критерии)
1	Исследование партии ИП давления одной конструкции	Градуировка [4] партии ИП Верификация партии ИП	Согласно требованиям конструкторской документации
2	Классификация ИП	Формирование внутри партии ИП непересекающихся групп преобразователей	Согласно требованиям разработанных авторами рекомендаций
3	Проведение приемосдаточных испытаний ДД, собранных на основе исследованной партии ИП	Градуировка [4] партии ДД Верификация партии ДД	Согласно требованиям конструкторской документации
4	Выбор ДД с наилучшими метрологическими характеристиками (МХ) из партии датчиков (назовем их «эталонными»)	Анализ метрологических характеристик ДД в партии	Минимальные значения приведенной и температурной погрешности ДД

выделить так называемые «эталонные» датчики, обладающие наилучшими МХ, по которым будет определяться ФП исследуемой конструкции.

На этапе 5 выполняются работы, связанные с построением функций преобразования для «эталонных» датчиков давления. Нами были использованы методы регрессионного анализа: метод всех возможных регрессий, метод исключения коэффициентов модели, метод включения коэффициентов, шаговый метод. На основе разработанных авторами алгоритмов и созданных программ, путем статистического моделирования осуществляется поиск наиболее простых моделей ФП, удовлетворяющих сформулированным в статье требованиям.

На этапе 6 исследуются и анализируются различные модели ФП для каждого «эталонного» датчика, полученные путем объединения ФП, созданных на этапе 5. Полученные на этом этапе модели можно трактовать как индивидуальные ФП «эталонных» датчиков.

Этап 7 предназначен для получения по разработанным авторами алгоритмам «общей» (единой) ФП ДД исследуемой конструкции, которая формируется на основе индивидуальных ФП «эталонных» ДД.

Этап 8 направлен на проверку пригодности полученной «общей» ФП для других (не эталонных) датчиков, входящих в исследуемую партию.

Этапы 9, 10 предполагают дополнительную проверку корректности созданной «общей» ФП путем применения ее к другим партиям ДД исследуемой конструкции.

#### Выводы

Авторами изучена задача выбора и обоснования функции преобразования измерительных преобразователей датчиков давления. Сформулированы новые требования к ФП ИП и определены пути решения поставленной задачи.

Предложена новая оригинальная последовательность (совокупность) математических методов и алгоритмов, обеспечивающая законченную процедуру выбора и обоснования «единой» (общей) модели функции преобразования для каждой из конструкций измерительных преобразователей.

Выполнена практическая реализация предложенной процедуры на реальных экспериментальных данных. Полученные результаты подтверждают правильность предложенных алгоритмов и использованных математических методов.

#### Литература

1. Стучебников, В.М. Тензорезисторные преобразователи на основе гетерозитаксиальных структур «кремний на сапфире» / В.М. Стучебников // *Измерения, контроль, автоматизация*. – 1982. – № 4(44). – С. 15–26.
2. *Официальный сайт ЗАО ПГ «Метран»*. URL: <http://www.metran.ru/home/hart/abouthart.html>
3. ГОСТ 8.009–84 Нормирование и использование метрологических характеристик средств измерений. – М.: Изд-во стандартов, 1998. – 151 с.

4. Семенов, Л.А. Методы построения градуировочных характеристик средств измерений / Л.А. Семенов, Т.Н. Сирая. – М.: Изд-во стандартов, 1986. – 128 с.

5. Грановский, В.А. Методы обработки экспериментальных данных при измерениях / В.А. Грановский, Т.Н. Сирая. – Л.: Энергоатомиздат. Ленинградское отделение, 1990. – 288 с.

6. Удод, Е.В. Исследование и разработка прецизионных математических моделей преобразования и алгоритмов вычисления значений давления: автореф. дис. ... канд. техн. наук / Е.В. Удод. – Таганрог, 2007.

7. Бычков, В.В. Высокоточные аналоговые и цифровые измерительные преобразователи давления: автореф. дис. ... канд. техн. наук / В.В. Бычков. – Томск, 2006.

8. Данилов, Н.А. Синтез функций преобразования измерительных приборов для контроля давления по заданному пределу приведенной погрешности: автореф. дис. ... канд. техн. наук / Н.А. Данилов. – СПб., 2007.

9. Концепция комплексного исследования метрологических характеристик датчиков давления «Метран» / А.П. Лапин, Д.В. Мысляева, Е.Е. Филиппова, В.В. Филимонова // *Практика приборостроения*. – 2002. – № 1. – С. 38–41.

10. Опыт сотрудничества ЗАО «Промышленная группа «Метран» и Южно-Уральского государственного университета в области исследования датчиков давления / А.П. Лапин, А.В. Конобеев, Л.В. Ушаков, А.В. Фетисов // *Практика приборостроения*. – 2003. – № 4. – С. 84–85.

11. Лапин, А.П. Статистическое исследование функций преобразования датчиков давления ПГ «Метран» / А.П. Лапин, Л.Ф. Нигамова, Е.Е. Чипеева // *Информационные, измерительные и управляющие системы и устройства: темат. сб. науч. тр.* – Челябинск: Изд-во ЮУрГУ, 2003. – С. 99–103.

12. Лапин, А.П. Исследование многофакторной функции преобразования датчиков давления ПГ «Метран» / А.П. Лапин, Е.Е. Филиппова // *Приборостроение: тематич. сб. науч. тр.* – Челябинск: Изд-во ЮУрГУ, 2002. – С. 33–36.

13. Лапина, Е.А. Статистическое моделирование функций преобразования датчиков давления типа «Метран» / А.П. Лапин, Ю.Н. Цыпина, Е.А. Лапина // *Вестник ЮУрГУ. Серия «Компьютерные технологии, управление, радиоэлектроника»*. – 2008. – № 3(103).

14. Дрейпер, Н. Прикладной регрессионный анализ: в 2 т. / Н. Дрейпер, Г. Смит. – 2-е изд., перераб. и доп. – М.: Финансы и статистика, 1986. – Т. 1. – 366 с.

15. Дрейпер, Н. Прикладной регрессионный анализ: в 2 т. / Н. Дрейпер, Г. Смит. – 2-е изд., перераб. и доп. – М.: Финансы и статистика, 1987. – Т. 2. – 351 с.

Поступила в редакцию 7 апреля 2009 г.

# МАТЕМАТИЧЕСКАЯ МОДЕЛЬ ИЗОБРАЖЕНИЯ ОТПЕЧАТКА ПАЛЬЦА НА ОСНОВЕ ТОПОЛОГИЧЕСКИХ ВЕКТОРОВ ДЛЯ ЛИНИЙ

*В.Ю. Гудков*

## MATHEMATICAL MODEL OF FINGERPRINT BASED ON TOPOLOGY VECTORS FOR LINES

*V.U. Gudkov*

В работе предлагается математическая модель изображения отпечатка пальца на основе топологических векторов для линий, которая сохраняется в шаблоне изображения. Шаблоны содержат наборы признаков, по которым выполняется идентификация изображений.

*Ключевые слова: отпечаток пальца, модель, контур, частные признаки, топология, вектор, шаблон.*

This paper presents a mathematical model of fingerprint based on the topological vectors for lines. The topological vectors and the list of minutiae are stored in the template. Templates are used to identify the fingerprint.

*Keywords: fingerprint, model, skeleton, minutiae, topology, vector, template.*

### Введение

Идентификация дактилоскопических изображений (ДИ) выполняется на основе шаблонов, содержащих признаки узора. Их базис составляет описание частных признаков в виде начал и окончаний линий, слияний и разветвлений линий [11]. Они могут быть детектированы по серому изображению, однако простых методов для этого недостаточно [3]. Поэтому при построении шаблона обычно опираются на скелет линий [2] (рис. 1). В классическом варианте в шаблоне дополнительно сохраняют данные гребневого счета [11, 12].

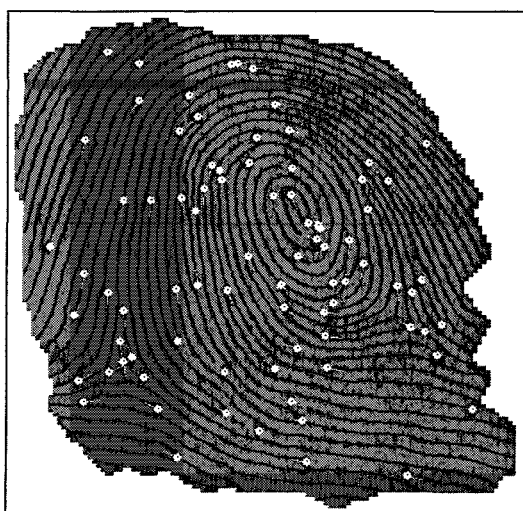


Рис. 1. Скелет и частные признаки

Математическая модель изображения должна содержать необходимое и достаточное количество признаков [1, 3]. В дактилоскопии информативными считают частные признаки и гребневый счет между ними [11, 12]. Однако они не исчерпывают множество математических моделей, используемых для автоматического доказательства индивидуальности узора [5, 7, 9, 12–14]. Каждая из таких моделей нацелена на повышение точности идентификации, однако неизвестно ни одной лучшей, свободной от недостатков модели [12]. Например, гребневый счет, который в криминалистике предписывается измерять по прямой линии [11], не выдерживает критики в области значительного искривления линий узора и в области петель, дельт и завитков [5, 13, 14].

В работе предлагается шаблон изображения, синтезируемый как отображение в виде

$$\Gamma : \{F_0^{(m)}\} \rightarrow \{L_m, L_l\},$$

где  $F_0^{(m)} = [f_0^{(m)}(x, y)]$  – скелет изображения (см. рис. 1);  $L_m$  – список частных признаков;  $L_l$  – список топологических векторов для линий.

### 1. Шаблон изображения

Шаблоны ДИ как набор или множество признаков объектов существенно различаются у разных производителей программного обеспечения. Например, форматы некоторых шаблонов несут ограничения на количество частных признаков:

окончаний и разветвлений [11, 12]. В криминалистике такие ограничения нежелательны, поэтому их рассматривать не будем. Положим, что шаблон содержит все множество распознаваемых частных признаков ДИ.

Введем некоторые определения.

**Определение 1.** Скелетом линии называется простая цепь  $\langle u, v \rangle$  с вершинами  $u$  и  $v$  в 8-смежности, проходящая вблизи геометрического центра линии, причем для каждой вершины  $p_1 \in \langle u, v \rangle$  существует ровно две смежные с ней вершины  $p_2$  и  $p_3$ , при этом вершины  $p_2$  и  $p_3$  несмежные.

**Определение 2.** Окончанием называется такая вершина  $p_1$  скелета, что для вершины  $p_1$  существует ровно одна смежная с ней вершина  $p_2$ .

**Определение 3.** Разветвлением называется такая вершина  $p_1$  скелета, что для вершины  $p_1$  существуют ровно три смежные с ней вершины  $p_2$ ,  $p_3$  и  $p_4$ , при этом любые две вершины из множества  $\{p_2, p_3, p_4\}$  попарно несмежные.

Содержание топологических векторов полностью определяется окончаниями и разветвлениями, которые считываются с вершин скелета и разрезают скелет на отрезки. Совокупность локально расположенных таких отрезков характеризует уникальность фрагмента узора, которая отражается в топологическом векторе.

### 1.1. Список частных признаков

Пусть  $M_i$  – индексированный номером  $i$  частный признак. Тогда список частных признаков  $L_m$  находится в виде

$$L_m = \{M_i = \{(x_i, y_i), \alpha_i, t_i, s_i, v_i, \theta_i, p_i, h_i\} \mid i \in 1..n_1\}, \quad (1)$$

где  $|L_m| = n_1$  – мощность списка;  $(x_i, y_i)$ ,  $\alpha_i$ ,  $t_i$ ,  $s_i$ ,  $v_i$ ,  $\theta_i$ ,  $p_i$  и  $h_i$  – координаты, направление, тип, метка сомнительности, величина и направление кривизны, вероятность и плотность линий в окрестности частного признака. Частные признаки детектируются в информативной области изображения [2, 5] (на рис. 1 информативная область ДИ затемнена, а скелет зачернен).

Координаты  $(x_i, y_i)$  частного признака  $M_i$  определяются координатами вершины скелета [1]. Направление  $\alpha_i$  как угол определяется простой цепью вершин скелета для окончания и тремя простыми цепями для разветвления [2]. Тип  $t_i \in \{0, 1\}$  устанавливается валентностью вершины скелета как вершины графа [4], где 0 – разветвление и 1 – окончание. Координаты  $(x_i, y_i)$ , направление  $\alpha_i$  и тип  $t_i$  являются базовыми параметрами  $M_i$ , используемыми при идентификации ДИ [6, 8, 10, 12].

Метка сомнительности  $s_i \in \{0, 1\}$  определяется положением частного признака, близким к не-

информативным областям (фону), причем 1 означает сомнительность. На границе неинформативных областей частные признаки не детектируются. Величина  $v_i$  и направление  $\theta_i$  кривизны определяются по разностям направлений линий в окрестности  $\varepsilon$  частного признака  $M_i$  [5, 7, 9]. Вероятность  $p_i$  рассчитывается как отношение средней величины оценок качества изображения в окрестности  $\varepsilon$  к наилучшей оценке качества в информативной области ДИ [5]. Плотность линий  $h_i$  рассчитывается как среднее количество линий, умещающихся в окрестности  $\varepsilon$  на прямой, проведенной перпендикулярно линиям [12]. Размер окрестности  $\varepsilon$  определяется алгоритмом обработки ДИ.

### 1.2. Список топологических векторов для линий

Список топологических векторов для линий  $L_i$  находится на основе списка частных признаков  $L_m$ , слоя скелета  $F_0^{(m)}$  и других вспомогательных матриц, элементы которых отображают локальные свойства линий ДИ. Эти матрицы формируются в пирамиде  $\mathcal{R}$ , представляющей информационные слои данных, распределенные по иерархиям пирамиды.

Список топологических векторов для линий синтезируется на основе всех вершин скелета, исключая вершины частных признаков, и записывается в виде

$$L_i = \{V_i = \{(e_j, n_j, l_j)\} \mid i \in 1..n_2, j \in 1..m_i\}, \quad (2)$$

где  $V_i$  – топологический вектор для группы вершин скелета;  $|L_i| = n_2$  – мощность списка и  $n_2 > n_1$ ;  $i$  – индекс как номер топологического вектора;  $j$  – номер связи в топологическом векторе;  $e_j$  – событие, а  $l_j$  – длина связи, сформированная частным признаком с номером  $n_j$ ;  $m_i$  – количество связей с учетом центральной линии в виде

$$m_i = 4m + 2. \quad (3)$$

Опишем процедуру построения топологических векторов для линий. В информативной области ДИ [12] выделяют линии и формируют стилизованное изображение в виде скелета [1]. По скелету детектируют два типа частных признаков: окончания и разветвления [2, 13] (см. рис. 1). Направление частного признака (угол) указывает в область увеличения числа линий [12]. Оно параллельно касательной к папиллярной линии в малой окрестности частного признака  $M_i$ . Каждый частный признак нумеруют и описывают координатами, направлением, типом, меткой сомнительности, величиной и направлением кривизны, вероятностью и плотностью линий в окрестности  $\varepsilon$  частного признака (1). Затем от каждого частного признака фиксируют две проекции: вправо и влево

перпендикулярно направлению частного признака на смежные скелетные линии. На рис. 2 проекции показаны пунктиром, а две соответствующие вершины скелета на линиях 1 и 2 закрашиваются «цветом проекции».

Выберем вершину скелета  $p_i$  (но не частный признак) и проведем через ее координаты  $(x, y_i)$  вправо и влево сечение на глубину нескольких линий  $m$  перпендикулярно касательным к пересекаемым линиям и пронумеруем по спирали, разворачивающейся по часовой стрелке, рассеченные линии, которые назовем связями. Сечение проходит, отслеживая направление кривизны линий [5, 7, 9]. Глубина сечения  $m$  обычно варьируется от одной до восьми линий вправо и также влево. Одна линия в сечении образует две связи. Количество связей в топологическом векторе рассчитывается по (3).

Топологический вектор определяют по сечению методом слежения за ходом каждой связи от сечения до встречи с другим частным признаком, расположенным на связи, или с проекцией от него на связь. При этом на связях детектируются следующие возможные события, показанные на рис. 3 и представленные в двоичном коде:

1101 – на связи проекция от окончания, расположенного справа по ходу связи, направление окончания ориентировано навстречу ходу связи;

1001 – на связи проекция от окончания, расположенного справа по ходу связи, направление окончания ориентировано по ходу связи;

1110 – на связи проекция от окончания, расположенного слева по ходу связи, направление окончания ориентировано навстречу ходу связи;

1010 – на связи проекция от окончания, расположенного слева по ходу связи, направление окончания ориентировано по ходу связи;

0101 – на связи проекция от разветвления, расположенного справа по ходу связи, направление разветвления ориентировано навстречу ходу связи;

0001 – на связи проекция от разветвления, расположенного справа по ходу связи, направление разветвления ориентировано по ходу связи;

0110 – на связи проекция от разветвления, расположенного слева по ходу связи, направление разветвления ориентировано навстречу ходу связи;

0010 – на связи проекция от разветвления, расположенного слева по ходу связи, направление разветвления ориентировано по ходу связи;

0011 – разветвление на связи, направление которого ориентировано по ходу связи;

0111 – разветвление на связи, образованной линией, касательная к которой образует минимальный угол при повороте направления разветвления против часовой стрелки;

1011 – разветвление на связи, образованной линией, касательная к которой образует минимальный угол при повороте направления разветвления по часовой стрелке;

1111 – окончание на связи, направление которого ориентировано навстречу ходу связи;

1100 – связь по линии замыкается, частный признак или проекция от него отсутствует;

0000 – на связи нет ни частного признака, ни проекции от него (обрыв скелета).

С событием как двоичным числом (их всего 14), детектированным на связи, ассоциируют номер

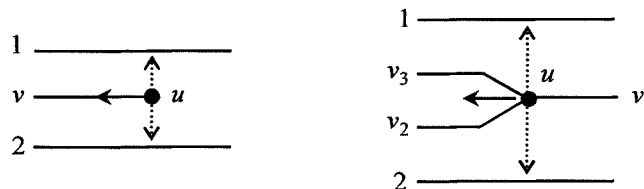


Рис. 2. Направления и проекции для окончания и разветвления

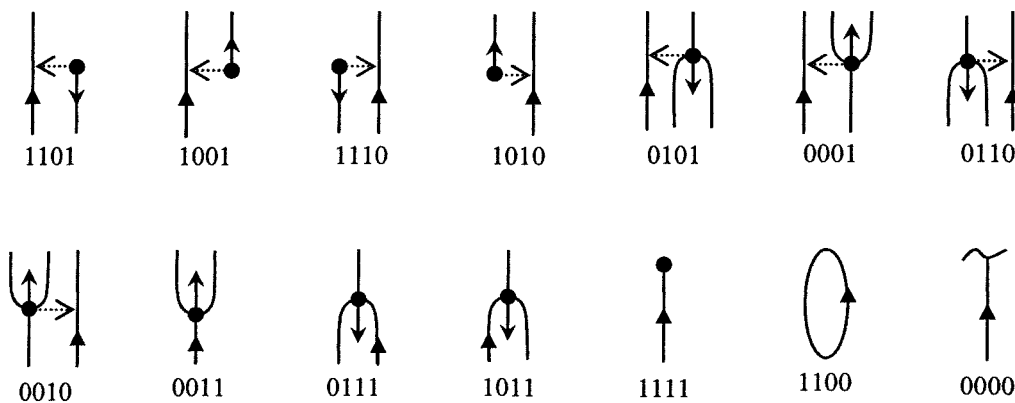


Рис. 3. События, детектируемые на связи

Таблица 1  
Топологический вектор для А

Номер	Событие	Индекс	Длина
0	1110	22	$l_0$
1	1111	19	$l_1$
2	1110	19	$l_2$
3	1111	22	$l_3$
4	0001	21	$l_4$
5	1101	19	$l_5$
6	1010	24	$l_6$
7	0010	25	$l_7$
8	0011	21	$l_8$
9	1111	23	$l_9$
10	1010	26	$l_{10}$
11	0011	25	$l_{11}$
12	0010	21	$l_{12}$
13	1010	20	$l_{13}$
14	1111	27	$l_{14}$
15	0001	25	$l_{15}$
16	0000	—	—
17	1001	20	$l_{17}$

Таблица 2  
Топологический вектор для В

Номер	Событие	Индекс	Длина
0	1110	22	$l_0$
1	1011	19	$l_1$
2	0111	19	$l_2$
3	1111	22	$l_3$
4	0001	21	$l_4$
5	0101	19	$l_5$
6	0110	19	$l_6$
7	0010	25	$l_7$
8	0011	21	$l_8$
9	1111	23	$l_9$
10	1010	26	$l_{10}$
11	0011	25	$l_{11}$
12	0010	21	$l_{12}$
13	1010	20	$l_{13}$
14	1111	27	$l_{14}$
15	0001	25	$l_{15}$
16	0000	—	—
17	1001	20	$l_{17}$

частного признака, инициирующего это событие. Событие привязано к номеру связи. Для событий 0000 и 1100 номера частных признаков отсутствуют. Нумерованный набор связей с оформленными событиями и номерами частных признаков есть *базовый топологический вектор* (экономный). Событие и номер частного признака образуют упорядоченную пару  $(e_j, n_j)$ . Событие может быть

дополнено длиной связи от сечения до позиции, в которой детектируется это событие. Так формируется *расширенный топологический вектор*. Событие, номер частного признака и длина связи образуют упорядоченную тройку  $(e_j, n_j, l_j)$ . Для событий 0000 и 1100 длины связей могут быть и описывать информативные области, свободные от частных признаков! Это существенно [7, 8]. При этом длины связей, обрывающихся на краю отпечатка, устойчивы в том смысле, что не укорачиваются при полной прокатке пальца [7].

Местоположение бита в событии определяет тип частного признака, его направление по отношению к направлению хода связи, его местоположение по отношению к связи и др. (см. рис. 3). События позволяют оперативно сопоставить *базовые топологические векторы* и ускорить процедуру идентификации узора [10].

Топологические векторы строят для каждой вершины скелета  $p_i$  (но не частного признака). Сечение разрезает линии на связи, пронумерованные по спирали, разворачивающейся по часовой стрелке. На рис. 4 в сечении для вершины А скелетной линии, завершающейся окончанием 19, пронумерованы связи 0–17, а соответствующий топологический вектор представлен в табл. 1. На

рис. 5 в сечении для вершины В скелетной линии, замыкающей в разветвление 19, пронумерованы связи 0–17, а соответствующий топологический вектор представлен в табл. 2. Сечения показаны пунктиром, а рисунки представляют обычную мутацию [9] окончания 19 в разветвление 19 (из-за грязной кожи). При этом вершины А и В скелета по сути одни и те же.

Начало нумерации связей в сечении для вершин А и В (связи № 0) несущественно, так как при развороте ДИ формируется зеркальное отображение номеров связей в сечении, которое легко распознается и учитывается при идентификации ДИ. Это аналог игры «Puzzle», где сборка выполняется стыковкой соответствующих разрезов. При глубине сечения  $m = 4$  для линии формируется восемнадцать связей  $m_i = 18$  по (3).

Количество топологических векторов счетное. На рис. 4 внизу двунаправленной стрелкой показана зона, вмещающаяся между частными признаками 19 и 25, в пределах которой для точки А при смещении ее по скелету синтезируется один и тот же *базовый топологический вектор*. Аналогичная зона между частными признаками 19 и 25 для точки В указана внизу на рис. 5. Топологические векторы с равными *базовыми топологическими векторами* объединяются [9]. Их количество при этом на порядок уменьшается и достигает величины 100–1000 для контактного отпечатка пальца с количеством частных признаков 20–100. Синтезируются уникальные топологические характеристики, соответствующие автоматически определяемым отрезкам линий, а не частным признакам.

На этом построение списка топологических векторов для линий завершается.



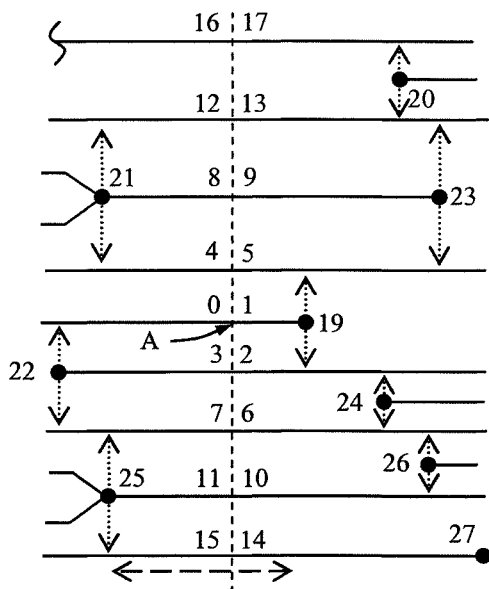


Рис. 4. Сечение для линии с окончанием

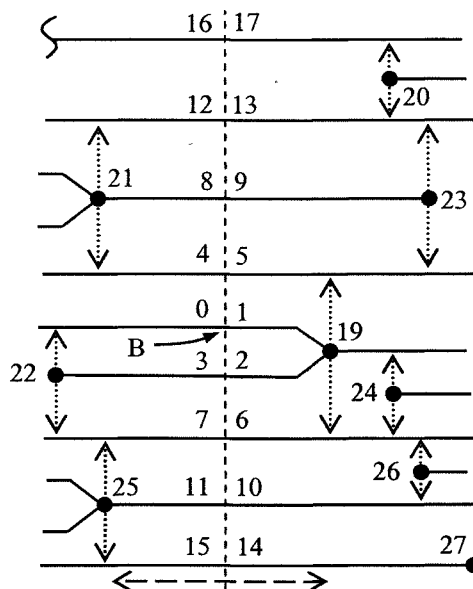


Рис. 5. Сечение для линии с разветвлением

### Заключение

Предложена топологическая модель ДИ для линий, которая сохраняется в шаблоне [12]. Топологические векторы образуют связанный граф с высокой степенью избыточности. Это позволяет соединить подграфы [4] фрагментарных следов отпечатков пальцев. Список  $L_1$  по (2) может быть представлен в эконономном формате (без длин связей).

В отличие от гребневого счета [9, 14] мутация окончания в разветвление в сечении для вершин А и В частично изменяет события и номера особенностей на связях, но не изменяет нумерацию связей и очередность перечисления частных признаков (см. табл. 1, 2). Это повышает устойчивость математической модели [9].

Устойчивость математической модели дополнительно повышается при объединении базовых топологических векторов выбором такого соответствующего расширенного топологического вектора, в котором минимальная длина связи максимальна [9].

Повышенная устойчивость математической модели улучшает точность идентификации ДИ, которая может быть выполнена с высокой производительностью [6, 8, 10, 13, 14].

### Литература

1. Гонсалес, Р. Цифровая обработка изображений / Р. Гонсалес, Р. Вудс; пер. с англ.; под ред. П.А. Чочиа. – М.: Техносфера, 2006. – 1072 с.
2. Гудков, В.Ю. Новая технология формирования скелетов дактилоскопических изображений / В.Ю. Гудков, А.А. Коляда, А.В. Чернявский; под ред. акад. БАН А.Ф. Чернявского и д.т.н. В.В. Ревинского // Методы, алгоритмы и программное обеспечение гибких информационных технологий

для автоматизированных идентификационных систем: сб. науч. ст. – Минск: БГУ, 1999. – С. 71–82.

3. Ковалевский, В.А. Методы оптимальных решений в распознавании изображений / В.А. Ковалевский. – М.: Наука, 1976. – 328 с.

4. Новиков, Ф.А. Дискретная математика для программистов: учеб. / Ф.А. Новиков. – СПб.: Питер, 2001. – 304 с.

5. Пат. 2185660 Российская Федерация, МПК G 06 K 9/52. Способ кодирования отпечатка папиллярного узора / В.Ю. Гудков, А.С. Боков. – № 2000118065/09; заявл. 07.07.2000; опубл. 20.07.2002; Бюл. № 20. – 13 с.

6. Пат. 2185661 Российская Федерация, МПК G 06 K 9/62. Способ сравнения папиллярных узоров пальцев / В.Ю. Гудков, А.С. Боков. – № 2000121735/09; заявл. 14.08.2000; опубл. 20.07.2002; Бюл. № 20. – 10 с.

7. Пат. 2298828 Российская Федерация, МПК G 06 K 9/52. Способ кодирования отпечатка папиллярного узора / В.Ю. Гудков, А.С. Боков. – № 2005132146/09; заявл. 17.10.2005; опубл. 10.05.2007; Бюл. № 13. – 17 с.

8. Пат. 2305317 Российская Федерация, МПК G 06 K 9/62. Способ сравнения папиллярных узоров / В.Ю. Гудков, А.С. Боков. – № 2005132145/09; заявл. 17.10.2005; опубл. 27.04.2007; Бюл. № 24. – 13 с.

9. Пат. 2321057 Российская Федерация, МПК G 06 K 9/52, А 61 В 5/117. Способ кодирования отпечатка папиллярного узора / В.Ю. Гудков. – № 2006142831/09; заявл. 04.12.2006; опубл. 27.03.2008; Бюл. № 9. – 13 с.

10. Пат. 2331108 Российская Федерация, МПК G 06 K 9/62. Способ сравнения отпечатков папиллярных узоров / Д.И. Аркабаев, В.Ю. Гудков. – № 2006140939/09; заявл. 20.11.2006; опубл. 10.08.2008; Бюл. № 22. – 17 с.

11. Komarinski P. *Automated fingerprint identification systems* / P. Komarinski. – New York: Academic Press, 2005. – 295 p.

12. Maltoni, D. *Handbook of fingerprint recognition* / D. Maltoni, D. Maio, A.K. Jain. – New York: Springer-Verlag, 2003. – 348 p.

13. Pat. 5631971 USA, Int.Cl. G 06 K 9/00. *Vector based topological fingerprint matching/*

*M.K. Sparrow (Winchester). – Field: Jul. 15, 1994; Date of patent: May. 20, 1997; U.S.Cl. 382/125. – 17 p.*

14. Sparrow, M.K. *A topological approach to the matching of single fingerprints: development of algorithms for use on latent finger marks* / M.K. Sparrow, P.J. Sparrow // *US dep. comer. nat. bur. stand. spec. pub.* – 1985. – № 500–126. – 61 p.

*Поступила в редакцию 27 апреля 2009 г.*

# ОЦЕНКА ТОЧНОСТИ АВТОМАТИЧЕСКОГО КОДИРОВАНИЯ ЧАСТНЫХ ПРИЗНАКОВ ИЗОБРАЖЕНИЙ ОТПЕЧАТКОВ ПАЛЬЦЕВ

*В.Ю. Гудков*

## ACCURACY RATING OF AUTOMATIC MINUTIAE RECOGNITION OF FINGERPRINTS

*V.U. Gudkov*

В статье рассматривается оценка точности автоматического распознавания частных признаков изображений отпечатков пальцев. Предлагается способ, учитывающий неравномерное распределение ошибок автоматического распознавания частных признаков по изображению. Данная оценка служит критерием для обучения алгоритмов классификационного анализа дактилоскопических изображений и выбора их функционального наполнения.

*Ключевые слова: отпечаток пальца, точность, рейтинг, неравномерное, распределение, частные признаки, распознавание.*

This paper presents a method of accuracy rating of automatic minutiae recognition. The method takes into account no uniform distribution of minutiae extraction errors. This accuracy rating serves as criterion to train the classification algorithms and to select their functionality.

*Keywords: fingerprint, accuracy, rating, no uniform, distribution, minutiae, recognition.*

### Введение

В постановке большинства задач распознавания результаты распознавания должны быть отнесены к определенным классам объектов [1, 3]. Формально это отношение задается пятеркой

$$\langle \Omega, C, W, W_t, f_c \rangle,$$

где  $\Omega$  – множество образцов;  $C = \{c_k, c^*\}$  – множество классов, причем  $c^*$  – дополнение до  $\{c_k\}$  и обозначает неизвестный класс;  $W = \{\omega_k \mid \omega_k \in \Omega\}$  – обучающая выборка образцов;  $W_t = \{\omega_k \mid \omega_k \in \Omega\}$  – тестовая выборка образцов;  $f_c$  – экспертный классификатор. Требуется найти функцию классификации  $f: \Omega \rightarrow P^{|C|}$ , которая наилучшим образом приближает  $f_c$  на тестовой выборке  $W_t$ , где вероятность

$$P^n = \{p \in R^n \mid p = (p_1, \dots, p_n), p_k \in R, \sum_{k=1}^n p_k = 1\}.$$

Оценить степень приближения можно при помощи функции ошибки, например, числа ошибочных классификаций на тестовой выборке [5, 10]. Оптимальным классификатором является тот, у которого функция ошибки минимальна [1, 8, 10]. Для стабилизации решений образцы, которые экспертно не классифицируются однозначно, удаляются из  $\Omega$ .

В задачах классификации опираются на понятие признака образца [3–5]. В дактилоскопии образцом является физический объект, например палец. Но компьютерные технологии имеют дело не с пальцами, а с их изображениями, которые получают с устройств ввода [6, 10–12]. На таких дактилоскопических изображениях (ДИ) гребни кожи обычно отображаются темными линиями с контрастно выделяющимися просветами между ними. Линии могут начинаться или оканчиваться, соединяться или расщепляться, образуя частные признаки ДИ. В компьютерных технологиях частные признаки приравниваются к признакам образца [4, 9–12].

Понятие частного признака хорошо проработано и при ручном кодировании оператор просто указывает его местоположение и ориентацию [6, 11]. Это классический случай однозначной классификации и именно так формируются выборки  $W$  и  $W_t$ . Небольшое смещение частного признака от его истинного местоположения (до одной линии) не считается за грубую ошибку, поскольку не оказывает существенного влияния на результаты идентификации [4, 12]. Однако влияние человеческого фактора при ручном кодировании частных признаков затрудняет оценку точности их автоматического распознавания. В работе [4] эта проблема частично решена. Но метод, предложенный в [4], не позволя-

ет акцентировать внимание на компактно расположенных ложных частных признаках. Именно они значительно ухудшают результаты идентификации. Действительно, след «родного» отпечатка пальца, приходящийся на фрагмент ДИ с этими ложными частными признаками, не идентифицируется [5, 6]. Поэтому необходимо обеспечить чувствительность метода оценки точности автоматического распознавания частных признаков к компактности расположения ложных частных признаков.

### 1. Метрика для сопоставления частных признаков

В суде при доказательстве идентичности двух отпечатков пальцев частные признаки являются основными признаками образца. Пусть одному и тому же образцу  $\omega$  функциями поставлены в соответствие шаблоны  $T_1 = f_1(\omega)$  и  $T_2 = f_2(\omega)$  в виде  $T_1 = \{m_i \mid i \in I = 1..n_1\}$ ,  $m_i = \{x_i, y_i, \theta_i\}$ , (1)  
 $T_2 = \{m_j \mid j \in J = 1..n_2\}$ ,  $m_j = \{x_j, y_j, \theta_j\}$ , (2)  
 где  $n_1$  – количество частных признаков на  $\omega$  при ручной кодировке ДИ функцией  $f_1$ ;  $n_2$  – количество частных признаков на  $\omega$  при автоматической кодировке ДИ функцией  $f_2$ ;  $(x, y)$  – координаты;  $\theta$  – направление как угол;  $I$  – множество индексов, причем для любого  $i \in I$  элемент  $m_i$  известен;  $J$  – множество индексов, причем для любого  $j \in J$  элемент  $m_j$  известен. Шаблоны (1) и (2) есть способ простого описания ДИ [4, 10, 12].

Введем функцию метрики

$$\mu(r, \alpha, m_i, m_j) \in \{0, 1\}, \quad (3)$$

где  $r$  – радиус как допуск на величину отклонения координат  $m_i$  и  $m_j$ ;  $\alpha$  – угол как допуск на величину отклонения углов направлений  $m_i$  и  $m_j$ ;  $m_i$  – частный признак, закодированный вручную,  $m_i \in T_1$ ;  $m_j$  – частный признак, закодированный автоматом,  $m_j \in T_2$ .

Функция метрики  $\mu$  определяет сопоставимость двух частных признаков, полученных разным образом, по формуле

$$\mu(r, \alpha, m_i, m_j) = \begin{cases} 1, & \text{если } \sqrt{(x_i - x_j)^2 + (y_i - y_j)^2} \leq r \wedge \text{turn}(\theta_i, \theta_j) \leq \alpha, \\ 0, & \text{иначе,} \end{cases}$$

где функция наименьшего угла поворота  $\text{turn}(\alpha, \beta) =$

$$= \begin{cases} \alpha - \beta, & \text{если } |\alpha - \beta| \leq 180, \\ 360 \text{sign}(\beta - \alpha) + \alpha - \beta, & \text{если } |\alpha - \beta| > 180 \end{cases}$$

и знаковая функция

$$\text{sign}(a) = \begin{cases} +1, & \text{если } a \geq 0, \\ -1, & \text{если } a < 0. \end{cases}$$

Если  $m_i$  сопоставляется с  $m_j$ , то отклонение координат  $m_i$  и  $m_j$  в допуске  $r$  и отклонение направлений  $m_i$  и  $m_j$  в допуске  $\alpha$  и  $\mu = 1$ , иначе  $\mu = 0$ .

Задавая  $r = 0$ , требуем, чтобы частные признаки ручной и автоматической кодировки располагались точка в точке. Это идеальный случай, недостижимый в действительности. Частные признаки не только смещаются, но появляются ложные и исчезают истинные [4, 6, 12]. Вопрос оценки качества признакового пространства объекта вообще актуален [2, 10], а в криминалистике в особенности из-за сильно загрязненных изображений следов отпечатков пальцев с мест преступлений. Если материал, на котором оставлен след, текстурирован, количество ложных частных признаков может возрастать и группироваться. Функция метрики  $\mu$  это не учитывает. Она выносит булево решение о сопоставимости только двух частных признаков, выбранных из шаблонов  $T_1$  и  $T_2$ , на основе указанных допусков, причем шаблон  $T_1$  считается эталонным. Однако это позволяет построить оценку, учитывающую как точность распознавания частных признаков [4], так и компактность расположения ложных частных признаков.

### 2. Оценка точности

Оценка точности автоматического распознавания частных признаков основана на функции массовой ранговой вероятности. Зададим допуск  $\alpha$  константой и ранжируем по допуску  $r$  сопоставляемые по (3) частные признаки  $(m_i, m_j)$ , которые назовем парой. Для допуска  $r$  количество сопоставляемых пар  $(m_i, m_j)$  находится как

$$n(r, \alpha) = \sum_{\substack{m_i \in T_1 \\ m_j \in T_2}} \mu(r, \alpha, m_i, m_j), \quad (4)$$

где  $T_1$  и  $T_2$  – шаблоны по (1) и (2).

**Определение 1.** Рангом пары  $r(m_i, m_j)$  называется минимальная величина допуска  $r \in 1..r_{\max}$ , при которой эта пара сопоставляется.

Количество пар ранга  $r(m_i, m_j)$  находится в виде  $\delta(r) = n(r, \alpha) - n(r - 1, \alpha)$ .

**Определение 2.** Функцией массовой ранговой вероятности  $P(r)$  (МРВ) называется вероятность того, что ранг пары  $r(m_i, m_j)$ .

Шаблоны  $T_1$  и  $T_2$  могут образовать не более  $\min(n_1, n_2)$  сопоставляемых пар. С позиций шаблона  $T_i$ ,  $i \in \{1, 2\}$ , функция МРВ находится в виде

$$P_i(r) = \frac{\delta(r)}{n_i}. \quad (5)$$

Обычно в шаблонах  $n_1 \neq n_2$ . Тогда первичный классификатор на основе функции МРВ (5) строится в виде

$$\hat{f}_c(r) = P_1(r)P_2(r) = \frac{\delta^2(r)}{n_1 n_2}, \quad (6)$$

где произведение в знаменателе (6) максимизирует  $\hat{f}_c(r)$  при  $n_1 \equiv n_2$ . Фактически величина  $\hat{f}_c(r)$  уменьшается при неравенстве количества частных признаков в шаблонах.

Оценку точности автоматического распознавания частных признаков как алгоритмический метод [7] построим следующим образом.

Зададим допуск  $r=1$ . Рассчитаем (6), а сопоставляемые по (3) частные признаки  $(m_i, m_j)$  удалим из шаблонов:  $T_1 = T_1 \setminus \{m_i \mid i \in I\}$  и  $T_2 = T_2 \setminus \{m_j \mid j \in J\}$ . Выберем частные признаки из шаблона  $T_2$  в качестве вершин полного графа  $G_r$ , количество вершин которого  $p(G_r) = |T_2|$ . В графе все вершины связаны. Рассчитаем сумму длин ребер в графе  $G_r$  от частного признака  $m_s \in T_2 \mid s \in J$  в виде

$$L_r(s) = \sum_{j \in J} l_{sj}, \quad (7)$$

где

$$l_{sj} = \sqrt{(x_s - x_j)^2 + (y_s - y_j)^2}.$$

Для исходного графа  $|T_2| = n_2$  и  $L_r(s) = L_0(s)$ .

Инкрементируем допуск  $r$  и рассчитаем  $\hat{f}_c(r)$  по (6), удаляя сопоставляемые частные признаки из шаблонов  $T_1$  и  $T_2$ . Выберем частные признаки из шаблона  $T_2$  в качестве вершин полного графа  $G_r$  и рассчитаем  $L_r(s)$  по (7). Последовательно применяя инкрементацию допуска  $r$ , оценим точность автоматического распознавания частных признаков на образце  $\omega_k$ , являющимся  $k$ -м по счету в тестовой выборке  $W_t$ , при помощи ошибки вероятностной массы по формуле

$$e_k = \max_s \sum_{r=1}^{r_{\max}} \exp\left(-\frac{L_r(s)}{L_0(s)}\right) \hat{f}_c(r) \frac{r}{r_{\max}}, \quad (8)$$

где  $r_{\max}$  – максимальный допуск на величину отклонения координат (до 15). В ошибке вероятностной массы  $e_k$  допуск  $\alpha$  задан константой, а величина ошибки отражает взвешенную вероятность несовпадения шаблонов  $T_1$  и  $T_2$  для образца  $\omega_k$ . Метрика  $\mu$  по (3) устанавливает критерий, по которому частные признаки  $m_i$  и  $m_j$  определяются сопоставляемыми. Чем ближе они располагаются, тем меньше величина  $e_k$ , так как уменьшается величина  $\hat{f}_c(r)$ . Заметим, что (6) не зависит от того, какое множество  $T_1$  или  $T_2$  является действительно полученным автоматом. Множитель  $r/r_{\max}$  в формуле (8) дополнительно увеличивает

чувствительность  $e_k$  к удаленным друг от друга частным признакам из  $T_1$  и  $T_2$ , а отношение сумм длин ребер по (7) обеспечивает чувствительность к компактному расположению ложных частных признаков. Действительно, для малых допусков  $r$  выполняется  $L_r(s) \approx L_0(s)$ , а с увеличением  $r$  сумма длин ребер  $L_r(s)$  снижается. Однако компактно расположенные ложные частные признаки не сопоставляются и можно найти такой частный признак  $m_s$ , который минимизирует  $L_r(s)$  (как центр тяжести). Это позволяет выбором наиболее «неудобного» частного признака  $m_s \mid s \in J$  максимизировать ошибку вероятностной массы  $e_k$ .

Точность автоматического распознавания частных признаков на тестовой выборке образцов  $W_t = \{\omega_k \mid \omega_k \in \Omega\}$ , ранжированной по величине ошибки вероятностной массы  $e_k$ , можно оценить при помощи суммарной ошибки

$$e = \frac{1}{|W_t|} \sum_{\omega_k \in W_t} e_k, \quad (9)$$

где  $|W_t|$  – мощность тестовой выборки образцов. Величина допуска  $\alpha$  фиксирована (30–45). На рис. 1 показана кривая ошибки вероятностной массы (8). Образцы  $\omega_k$  ранжированы по величине  $e_k$ . Старшим индексам  $k$  соответствуют ДИ с успешно распознанными частными признаками. При удалении хороших изображений выделяется множество «плохих лидеров». Ограничивая тестовую выборку образцов  $W_t$  множеством «плохих лидеров», можно определить индекс  $k'$ , начиная с которого оценка (9) наиболее критична, а изображения представляют наибольший интерес для развития методов классификационного анализа.

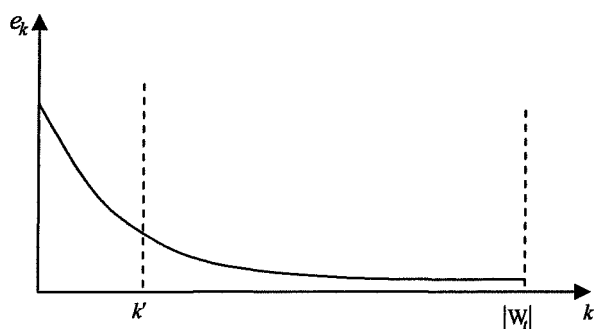


Рис. 1. Кривая ошибки вероятностной массы

### Выводы

Оценка точности автоматического распознавания частных признаков ДИ на основе МРВ (5) и суммарной ошибки (9) реализована в рамках автоматизированной дактилоскопической информационной системы СОНДА и использована для подготовки к тестам, проводимым независимыми международными организациями. Обучающая и тестовая

вая выборки образцов сформированы вручную на основе изображений, полученных от этих организаций в качестве образцов для пробных предварительных тестов.

Оценки (8) и (9) показали высокую эффективность при подготовке к международным тестам. Они служат удобным средством для предварительного обучения алгоритмов распознавания частных признаков, по которым выполняется идентификация изображений отпечатков пальцев. Выявлена зависимость качества обучения алгоритмов и выбора их функционального наполнения от мощности обучающей выборки образцов.

#### Литература

1. Аркадьев, А.Г. Обучение машины классификации объектов / А.Г. Аркадьев, Э.М. Браверман. – М.: Наука, 1971. – 192 с.
2. А.с. 1652984 СССР, МКИ G 06 K 9/00. Способ формирования признаков при распознавании изображений объектов / Г.Е. Баскин, В.И. Гордиенко, Л.С. Королюк, Б.П. Русын. – № 4468868/24; заявл. 01.08.88; опубл. 30.05.91, Бюл. № 20. – 9 с.
3. Боннер, Р.Э. Некоторые методы классификации / Р.Э. Боннер; под ред. Э.М. Бравермана // Автоматический анализ сложных изображений. – М.: Мир, 1969. – С. 209–234.
4. Гудков, В.Ю. Оценка точности автоматического кодирования особенностей отпечатков пальцев / В.Ю. Гудков; под ред. член-корр. РАН В.Л. Арлазарова и д.т.н. проф. Н.Е. Емельянова // Обработка изображений и анализ данных: Труды ИСА РАН. – М.: Книжный дом «ЛИБРОКОМ», 2008. – Т. 38. – С. 357–363.
5. Гуревич, И.Б. Проблемы распознавания изображений: распознавание, классификация, прогноз: математические методы и их применение / И.Б. Гуревич. – М.: Наука, 1982. – Вып. 1. – 237 с.
6. Дактилоскопическая экспертиза: современное состояние и перспективы развития / В.Е. Корноухов, В.К. Анциферов, Г.П. Морозов и др.; под ред. Г.Л. Грановского. – Красноярск: Изд-во Красноярского университета, 1990. – 416 с.
7. Журавлев, Ю.И. Об алгоритмических методах в задачах распознавания и классификации: распознавание, классификация, прогноз: математические методы и их применение / Ю.И. Журавлев. – М.: Наука, 1989. – Вып. 1. – С. 9–16.
8. Кирилов, С.Н. Оптимизация признакового пространства в задачах распознавания элементов речи / С.Н. Кирилов, А.С. Шелудяков // Автоматика и телемеханика. – 1998. – № 5. – С. 157–162.
9. Орлов, П.Г. Статистический метод оценки идентификационных признаков в дактилоскопической экспертизе: применение теории вероятностей и математической статистики в судебной экспертизе: материалы к теоретической конференции / П.Г. Орлов. – М., 1963. – С. 22–28.
10. Bolle, R.M. Guide to biometrics / R.M. Bolle, J.Y. Connel, S. Pankanti, N.K. Ratha. – New York: Springer-Verlag, 2004. – 368 p.
11. Komarinski, P. Automated fingerprint identification systems / P. Komarinski. – New York: Academic Press, 2005. – 295 p.
12. Maltoni, D. Handbook of fingerprint recognition / Davide Maltoni, Dario Maio, Anil K. Jain. – New York: Springer-Verlag, 2003. – 348 p.

Поступила в редакцию 27 апреля 2009 г.

# ПРИМЕНЕНИЕ МОДЕЛЬНОГО ПРОГНОЗИРУЮЩЕГО УПРАВЛЕНИЯ ДЛЯ РЕГУЛИРОВАНИЯ АКТИВНОЙ МОЩНОСТИ ЭНЕРГОБЛОКА ТЕПЛОВОЙ ЭЛЕКТРОСТАНЦИИ

*А.А. Лебедев*

## MODEL PREDICTIVE CONTROL OF THE THERMAL POWER PLANT GENERATING UNIT OUTPUT

*A.A. Lebedev*

В данной статье описана процедура синтеза алгоритма модельного прогнозирующего управления для регулирования активной мощности энергоблока тепловой электростанции, позволяющего повысить качество регулирования по сравнению с классическим линейным управлением.

*Ключевые слова:* тепловая электростанция, энергоблок, модельное прогнозирующее управление.

This paper reports the development of model predictive control for improved control of a thermal power plant generating unit output. As the results of modeling demonstrate, this control strategy provides more efficient control than the standard linear one.

*Keywords:* thermal power plant, power generating unit, model predictive control.

### Введение

Традиционные алгоритмы управления в электроэнергетических системах сложились более полувека назад и предполагают проектирование системы управления в виде отдельных линейных подсистем, использующих классические линейные регуляторы. Однако линейные регуляторы не позволяют учитывать ряд важных нелинейностей объекта управления: ограничение амплитуды и скорости ее нарастания для исполнительных механизмов, ограничения на некоторые параметры (например, уровень воды в барабане котла). Кроме того, энергоблок является сложной и многосвязной системой, поэтому согласованная настройка всех регуляторов требует больших временных затрат. Также очевидно, что такой подход не гарантирует получения оптимального и экономичного управления [1].

В отличие от классического линейного управления, модельное прогнозирующее управление (МПУ) позволяет учитывать характерные для энергоблоков нелинейности, а также обеспечить возможность оптимизации процесса. Данные особенности алгоритмов МПУ позволяют значительно повысить эффективность работы системы управления.

В данной работе представлена процедура синтеза алгоритма МПУ для регулирования активной мощности энергоблока тепловой электростанции,

учитывающего ограничения исполнительных механизмов и позволяющего повысить качество регулирования мощности по сравнению с классическим линейным управлением.

### 1. Математическая модель энергоблока

Нелинейная модель энергоблока была получена из векторных уравнений балансов массы, энергии и импульсов для элементов парового котла, представленных в виде рекуперативных теплообменников, а также из уравнений материального баланса отсеков паровой турбины и уравнения ее вращательного движения [1, 2]. Линейная модель энергоблока получена из нелинейной с помощью процедуры линеаризации «в малом» относительно номинального режима работы энергоблока и с учетом допущений, принятых в [3]. Входами модели являются входные сигналы, подаваемые на исполнительные механизмы, управляющие перемещением клапана турбины, расходом питательной воды и расходом топлива ( $u_{кл}$ ,  $u_{п.в.}$  и  $u_{т}$  соответственно). Выходы – активная мощность  $N$ , температура пара в промежуточной точке тракта  $t_{п}$  и температура факела в топке  $t_{ф}$ . Записанная в пространстве состояний, модель имеет вид

$$\begin{aligned} \dot{\mathbf{x}}(t) &= \mathbf{A}\mathbf{x}(t) + \mathbf{B}\mathbf{u}(t), \\ \mathbf{y}(t) &= \mathbf{C}\mathbf{x}(t), \end{aligned} \quad (1)$$

где  $\mathbf{u}(t) = [u_{кл} \ u_{п.в.} \ u_{т}]^T$ ,  $\mathbf{y}(t) = [t_{п} \ t_{ф} \ N]^T$ ,  $\mathbf{x}(t) \in \mathbf{R}^{30}$ .

## 2. Синтез алгоритма модельного прогнозирующего управления

Типовая структурная схема системы управления энергоблоком представлена на рис. 1 [2]. Введены следующие обозначения (см. рис. 1): ТРМ – турбинный регулятор мощности, КРМ – котельный регулятор мощности, РТ – регулятор температуры, ИМК – исполнительный механизм клапана турбины, ИМПВ – исполнительный механизм подачи питательной воды, ИМТ – исполнительный механизм подачи топлива, Д – дифференциатор,  $N_3$  – задание по мощности,  $t_3$  – задание по температуре,  $D_{п.в.}$  – расход питательной воды,  $D_t$  – расход топлива,  $h_{кл}$  – относительное перемещение клапана турбины,  $p_x$  – давление пара на выходе котла,  $D_x$  – расход пара на выходе котла,  $u_{кл}$ ,  $u_{п.в.}$  и  $u_t$  – соответственно управляющие воздействия действующие на исполнительные механизмы клапана турбины, клапана подачи питательной воды и подачи топлива.

Модельное прогнозирующее управление – алгоритм управления, основанный на решении задачи оптимального управления в реальном масштабе времени и состоящий из следующих шагов.

1. В момент времени  $k$  для текущего вектора состояния  $x(k)$  решить в реальном времени для разомкнутого контура задачу оптимального управления для некоторого будущего интервала (горизонт предсказания  $n$ ), учитывая ограничения.

2. Применить первый шаг последовательности оптимального управления.

3. Повторить процедуру для момента  $k+1$ , используя текущее состояние  $x(k+1)$ .

Наиболее приемлемым вариантом применения алгоритма МПУ в задаче регулирования активной мощности энергоблока является замена одного или нескольких линейных регуляторов системы управления на регуляторы с МПУ. Такой подход позволяет сохранить существующую структуру системы управления, оставляя возможность непосредственного регулирования других парамет-

ров энергоблока (давление пара на выходе котла, температура в промежуточной точке тракта и др.).

Основным управляющим воздействием при регулировании активной мощности энергоблока является перемещение регулирующего клапана турбины  $h_{кл}$  [2]. Формирование сигналов, управляющих исполнительным механизмом клапана турбины, осуществляется ТРМ. Поэтому использование алгоритма МПУ в ТРМ представляется наиболее эффективным.

Таким образом, в качестве объекта управления для регулятора с МПУ выступает канал  $u_{кл} - N$ . Выделяя вышеуказанный канал из модели (1) и переходя к дискретной форме с шагом 0,1 с, получим

$$\begin{aligned} x(k+1) &= A_1 x(k) + B_1 u(k), \\ y(k) &= C_1 x(k), \end{aligned} \quad (1)$$

где  $u = u_{кл}$ ,  $y = N$ ,  $x \in R^{14}$ .

Задача МПУ формулируется как задача ограниченного квадратичного программирования, учитывающая нелинейность исполнительного механизма:

$$\begin{aligned} J(U, \Delta U, E) &= \Delta U^T Q_{\Delta U} \Delta U + E^T Q_E E \rightarrow \min, \\ x(k) - Ax(k-1) - Bu(k-1) &= 0, \\ y(k) - Cx(k) &= 0, \\ 0 \leq u(k) \leq 1, \end{aligned}$$

где

$$\begin{aligned} U &= [u(k+1), \dots, u(k+n)], \\ \Delta U &= [u(k+1) - u(k), \dots, u(k+n) - u(k+n-1)], \\ E &= [y(k+1) - y_3(k+1), \dots, y(k+n) - y_3(k+n)], \\ Q_{\Delta U} &= \text{diag}(0, 1 \dots 0, 1), \quad Q_E = \text{diag}(1 \dots 1). \end{aligned}$$

Здесь  $y_3(k)$  – заданное значение (уставка).

## 3. Результаты моделирования

Для моделирования использовался пакет программ MATLAB 2006a / SIMULINK. На рис. 2 представлены результаты моделирования при уве-

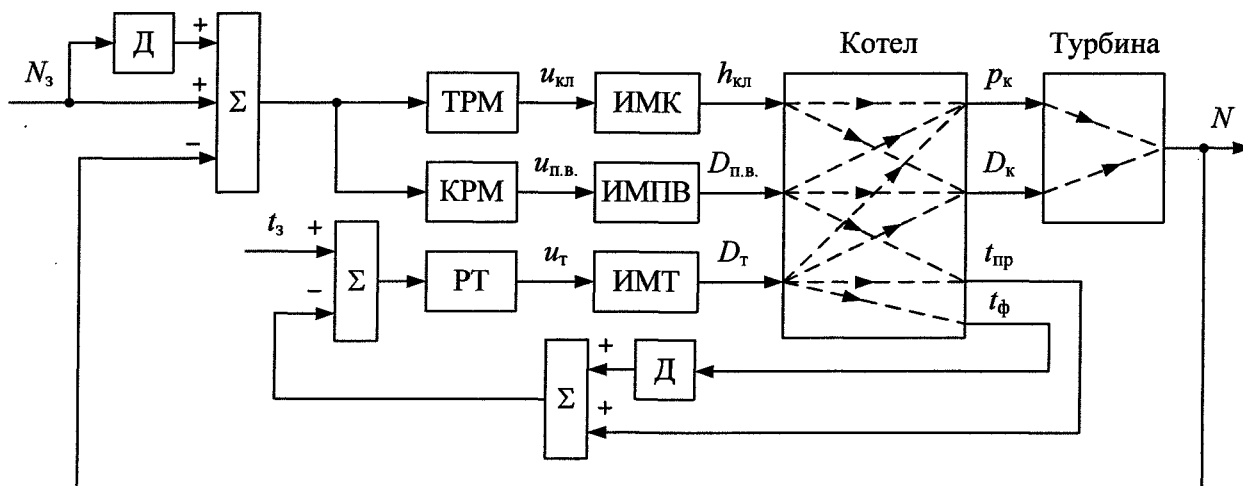
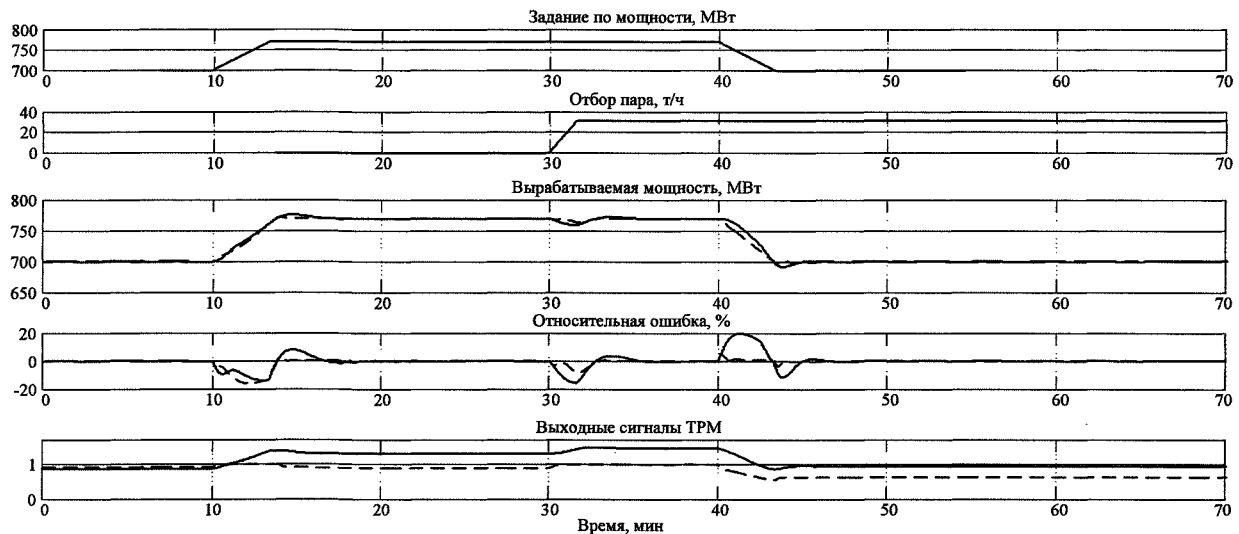


Рис. 1. Структурная схема системы управления мощностью энергоблока





**Рис. 2. Результаты моделирования**

личении задания по мощности на 10 %. При этом на 30-й минуте на систему начинает действовать внешнее возмущение – включается отбор пара на собственные нужды. На трех последних графиках сплошная линия соответствует системе с ПИД-регуляторами, настроенными на технический оптимум, пунктирная линия соответствует МПУ.

Используемые модели исполнительных механизмов содержат естественные ограничения по амплитуде, поэтому при подаче входного сигнала выход исполнительного механизма попадает в зону насыщения, что приводит к ухудшению качества регулирования. В отличие от классического управления, алгоритм МПУ учитывает данное ограничение, что позволяет обеспечить работу механизма в зоне линейности. Также наблюдается увеличение быстродействия и повышение качества регулирования.

#### **Заключение**

В рамках проведенных исследований рассмотрена возможность применения алгоритмов модель-

ного прогнозирующего управления. Экспериментальные результаты показывают увеличение быстродействия и улучшение качества регулирования мощности при модельном прогнозирующем управлении по сравнению с классическим управлением. При этом более высокое качество регулирования достигается при меньшей интенсивности управляющих воздействий.

#### **Литература**

1. Синергетические методы управления сложными системами: Энергетические системы / под ред. А.А. Колесникова. – М.: КомКнига, 2006. – 248 с.
2. Плетнев Г.П. Автоматизированное управление объектами тепловых электростанций: учеб. пособие для вузов. / Г.П. Плетнев. – М.: Энергоиздат, 1981. – 367 с.
3. Чертков Н.К. Система моделирования энергетических установок / Н.К. Чертков, Г.Б. Барменков, Б.Б. Мешков // Теплоэнергетика. – 2006. – № 5. – С. 62–65.

*Поступила в редакцию 14 мая 2009 г.*

## ИСПОЛЬЗОВАНИЕ НЕЙРОСЕТЕВОГО МОДЕЛИРОВАНИЯ ДЛЯ ДИАГНОСТИКИ СТРУКТУРЫ ЭМОЦИОНАЛЬНОЙ КОМПЕТЕНТНОСТИ РЕБЕНКА ПО ФАКТОРАМ, ВКЛЮЧАЮЩИМ ФИЗИОЛОГИЧЕСКИЕ ПАРАМЕТРЫ

*Е.В. Вишневетская, И.А. Горбунов*

## USE OF NEURAL MODELING STRUCTURE FOR DIAGNOSTICS OF EMOTIONAL COMPETENCE OF A CHILD WITH FACTORS INCLUDING PHYSIOLOGICAL PARAMETERS

*E.V. Vishnevetskaya, I.A. Gorbunov*

В работе рассматривается практическое применение математической модели нейронной сети (многослойного перцептрона) для диагностики эмоциональной компетентности детей дошкольного возраста по физиологическим параметрам. Эмоциональная компетентность оценивалась с помощью психологического теста, включающего в себя 6 шкал, определяющих узнавание и выражение эмоций с помощью мимики, голоса и рисунка. В качестве физиологических параметров использовались показатели спектральных плотностей ЭЭГ, в стандартных диапазонах частот, а также изменение показателей вегетативной регуляции (ЧСС, АД) в ответ на эмоциогенные и интеллектуальные нагрузки. В результате построения диагностической модели были проверены показатели качества диагностики, а также описаны возможные физиологические механизмы, влияющие на особенности выраженности различных показателей эмоциональной компетентности.

*Ключевые слова: нейросетевое моделирование, эмоциональная компетентность.*

We consider the practical application of mathematical models of neural networks (multi-layer perceptron) for the diagnosis of emotional competence of preschool children on the physiological parameters. Emotional competence was assessed using the psychological test, which consists of 6 scales that determine the recognition and expression of emotion through facial expressions, voice and image. The physiological parameters were used as indicators of EEG spectral density in the standard frequency ranges, as well as indicators of change in autonomic regulation (HR, BP) in response to emotion-generating and intellectual burden. As a result of constructing diagnostic models quality of diagnosis was tested, and describes the possible physiological mechanisms that influence the particular expression of various indicators of emotional competence.

*Keywords: neural modeling, emotional competence.*

**Актуальность.** В дошкольном периоде психофизиологическое развитие ребенка происходит быстрыми темпами. При этом закладываются основы для успешной адаптации при освоении новой ведущей деятельности (учебы). При освоении учебной деятельности ребенок попадает в новую жизненную ситуацию, в которой наработанные стереотипы поведения, связанные с поведением в семье и дошкольном учреждении, могут быть неприменимы. В таких ситуациях ведущую роль в организации поведения будет играть эмоциональ-

ная сфера ребенка, так как она содержит в себе не только онтогенетический опыт, но и базовые поведенческие программы, сформированные в процессе филогенеза, а следовательно более универсальные (Ч. Дарвин, К. Изард, П.К. Симонов). Поэтому в дошкольном развитии ребенка очень важное значение имеет эмоциональная сфера. Однако, для коррекции и оптимизации развития ребенка необходима диагностика особенностей эмоциональной регуляции ребенка, которая бы учитывала не только качественные характеристики эмоций, прояв-

**Вишневетская Елена Васильевна** – канд. психол. наук, доцент кафедры прикладной психологии ЮУрГУ; Elena--VV@yandex.ru

**Горбунов Иван Анатольевич** – канд. психол. наук, зав. учеб. лаб. фак. психологии СПбГУ; Jean@psy.pu.ru

**Vishnevetskaya Eleva Vasilevna** – PhD, associate professor of Applied psychology department of SUSU; Elena--VV@yandex.ru

**Gorbunov Ivan Anatolievich** – PhD, head of the laboratory of psychology of Saint-Petersburg state university; Jean@psy.pu.ru

ляемые в поведении, но и особенности организации физиологических механизмов, стоящих за этими внешними проявлениями [1–4].

Особое место в эмоциональной регуляции имеет коммуникативная сфера, которая, с одной стороны, охватывает практически все аспекты эмоциональной регуляции, а с другой стороны, является одним из важных компонентов социальной адаптации ребенка. Следовательно, актуальным является построение диагностической модели, способной диагностировать успешность распознавания и выражения эмоций, учитывая функциональное состояние мозга ребенка и вегетативные функции, отраженные в ЭЭГ (электроэнцефалограмма мозга) и параметрах изменения ВНС (вегетативная нервная система) при эмоциогенной нагрузке.

**Цель исследования:** построение диагностической модели распознавания и выражения эмоций по характеристикам функционирования ЦНС (центральная нервная система) и ВНС.

**Задачи:**

1) выбор адекватной диагностической модели, позволяющей учесть всю сложность взаимодействия между психическими явлениями и физиологическими функциями;

2) построение диагностических правил, с помощью которых можно успешно диагностировать особенности узнавания и выражения эмоций;

3) описать функциональные системы, деятельность которых стоит за успешностью узнавания и выражения эмоций.

**Гипотеза:** 1) адекватной физиологической моделью, с помощью которой можно успешно диагностировать эмоциональную компетентность ребенка, а также позволяющей описать ее физиологические механизмы являются математические модели нейронных сетей; 2) в основе физиологических механизмов эмоциональной компетентности лежат активирующие системы мозга, а также особенности вегетативной регуляции, связанные с базовыми биологическими программами.

По данным многочисленных авторов в дошкольном возрасте наблюдается улучшение распознавания эмоций, а также значительные качественные изменения эмоциональной регуляции детей. Эмоциональная компетентность с возрастом значительно увеличивается [1, 5, 8].

На первом году жизни регуляция эмоционального напряжения осуществляется опекуном ребенка, создавая почву для «направляемой саморегуляции». Регуляция с этой точки зрения представляет собой ряд качественных преобразований, затрагивающих созревание ЦНС и формирование более сложных поведенческих программ [7]. В результате этого эмоции становятся менее подверженными воздействиям окружающей среды и контролируются благодаря пониманию личных переживаний.

Формирование такого сложного поведения, подразумевающего взаимодействие различных

систем (когнитивной, эмоциональной, поведенческой, системы отношений), позволяет говорить об усложнении, а также об увеличении количества ФС (функциональных систем), обеспечивающих эмоциональную сферу ребенка. Такие изменения могут отражаться на усложнении динамики фоновых характеристик активности мозга ребенка.

Взаимосвязи между физиологическими и психологическими явлениями основаны на взаимодействии различных систем организма и отличаются нелинейностью и многоуровневостью. Следовательно, чтобы распознать психические явления по физиологическим параметрам, требуется модель, учитывающая нелинейные взаимосвязи и взаимозависимость параметров входного образа. Обычно для распознавания сложных образов используются многослойные нейронные сети обратного распространения [6].

Несмотря на большое разнообразие вариантов нейронных сетей, все они имеют общие черты. Так все они, как и мозг человека, состоят из большого числа однотипных элементов – нейронов, которые имитируют клетки головного мозга, связанные между собой. На рисунке показана схема нейрона.

Из рисунка видно, что искусственный нейрон, так же как и живой, состоит из синапсов, связывающих входы нейрона с телом, которое осуществляет обработку входных сигналов и аксона, который связывает нейрон с нейронами следующего слоя. Каждый синапс имеет вес, определяющий, насколько соответствующий вход нейрона влияет на его состояние. Состояние нейрона определяется по формуле

$$S = \sum_{i=1}^n x_i w_i, \quad (1)$$

где  $n$  – число входов нейрона,  $x_i$  – значение  $i$ -го входа нейрона,  $w_i$  – вес  $i$ -го синапса.

Затем определяется значение аксона нейрона по формуле

$$y = f(S), \quad (2)$$

где  $f$  – некоторая функция, которая называется активационной.

Наиболее часто в качестве активационной функции используется так называемый сигмоид, который имеет следующий вид:

$$f(x) = \frac{1}{1 + e^{-ax}}. \quad (3)$$

Основное достоинство этой функции в том, что она дифференцируема на всей оси абсцисс и имеет очень простую производную

$$f'(x) = af(x)(1 - f(x)). \quad (4)$$

При уменьшении параметра  $a$ , сигмоид становится более пологим, вырождаясь в горизонтальную линию на уровне 0,5 при  $a = 0$ . При увеличении  $a$  сигмоид все больше приближается к функции единичного скачка.

Нейронные сети обратного распространения – это мощнейший инструмент поиска закономерности

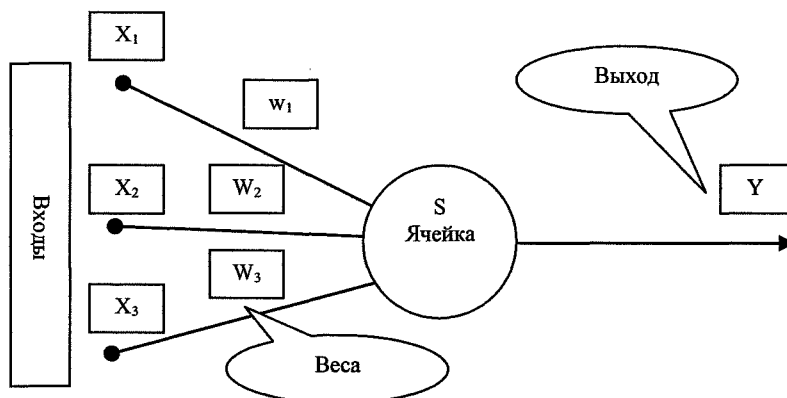


Схема одной ячейки нейронной сети «нейрона»

стей, прогнозирования, качественного анализа. Такое название – сети обратного распространения (back propagation) – они получили из-за используемого алгоритма обучения, в котором ошибка распространяется от выходного слоя к входному, т.е. в направлении, противоположном направлению распространения сигнала при нормальном функционировании сети [6].

Нейронная сеть (НС) обратного распространения состоит из нескольких слоев нейронов, причем каждый нейрон слоя  $i$ , связан с каждым нейроном слоя  $i+1$ , т.е. речь идет о полносвязной НС.

В общем случае задача обучения нейронной сети сводится к нахождению некоей функциональной зависимости  $Y=F(X)$  где  $X$  – входной, а  $Y$  – выходной векторы. В общем случае такая задача, при ограниченном наборе входных данных имеет бесконечное множество решений. Для ограничения пространства поиска при обучении ставится задача минимизации целевой функции ошибки нейронной сети, которая находится по методу наименьших квадратов:

$$E(w) = \frac{1}{2} \sum_{j=1}^p (y_j - d_j)^2, \quad (5)$$

где  $y_j$  – значение  $j$ -го выхода нейросети,  $d_j$  – целевое значение  $j$ -го выхода,  $p$  – число нейронов в выходном слое, Обучение нейросети производится методом градиентного спуска, т.е. на каждой итерации изменение веса производится по формуле

$$\Delta w_{ij} = -\eta \cdot \frac{\partial E}{\partial w_{ij}}, \quad (6)$$

где  $\eta$  – параметр определяющий скорость обучения

$$\frac{\partial E}{\partial w_{ij}} = \frac{\partial E}{\partial y_j} \cdot \frac{\partial y_j}{\partial S_j} \cdot \frac{\partial S_j}{\partial w_{ij}}. \quad (7)$$

где  $y_j$  – значение выхода  $j$ -го нейрона,  $S_j$  – взвешенная сумма входных сигналов, определяемая по формуле (1). При этом множитель

$$\frac{\partial S_j}{\partial w_{ij}} \equiv x_i, \quad (8)$$

где  $x_i$  – значение  $i$ -го входа нейрона.

Далее рассмотрим определение первого множителя формулы (7)

$$\frac{\partial E}{\partial y_j} = \sum_k \frac{\partial E}{\partial y_k} \cdot \frac{\partial y_k}{\partial S_k} \cdot \frac{\partial S_k}{\partial y_j} = \sum_k \frac{\partial E}{\partial y_k} \cdot \frac{\partial y_k}{\partial S_k} \cdot w_{jk}^{(n+1)}, \quad (9)$$

где  $k$  – число нейронов в слое  $n+1$ .

Введем вспомогательную переменную

$$\delta_j^{(n)} = \frac{\partial E}{\partial y_j} \cdot \frac{\partial y_j}{\partial S_j}. \quad (10)$$

Тогда мы сможем определить рекурсивную формулу для определения  $\delta_j^n$   $n$ -го слоя если нам известно  $\delta_j^{n+1}$  следующего  $n+1$ -го слоя.

$$\delta_j^n = \left[ \sum_k \delta_k^{n+1} \cdot w_{jk}^{n+1} \right] \frac{\partial y_j}{\partial S_j}. \quad (11)$$

Нахождение же для последнего слоя нейронной сети не представляет трудности, так как нам известен целевой вектор, т.е. вектор тех значений, которые должна выдавать НС при данном наборе входных значений

$$\delta_j^N = (y_j^N - d_j) \frac{\partial y_j}{\partial S_j}. \quad (12)$$

И, наконец, запишем формулу (6) в раскрытом виде

$$\Delta w_{ij}^n = -\eta \cdot \delta_j^n \cdot x_i^n. \quad (13)$$

Рассмотрим теперь полный алгоритм обучения нейросети.

1. Подать на вход нейронной сети один из требуемых образов и определить значения выходов нейронов нейросети.

2. Рассчитать  $\delta^N$  для выходного слоя нейронной сети по формуле (12) и рассчитать изменения весов  $\Delta w_{ij}^N$  выходного слоя  $N$  по формуле (13).

3. Рассчитать по формулам (11) и (13) соответственно  $\delta^n$  и  $\Delta w_{ij}^n$  для остальных слоев НС,  $n=(N-1)\dots 1$ .

4. Скорректировать все веса нейронной сети

$$w_{ij}^n = w_{ij}^n(t-1) + \Delta w_{ij}^n t. \quad (14)$$

5. Если ошибка существенна, то перейти на шаг 1.

На этапе 2 сети поочередно в случайном порядке предъявляются вектора из обучающей последовательности.

Простейший метод градиентного спуска, рассмотренный выше, очень неэффективен в случае, когда производные по различным весам сильно отличаются. Это соответствует ситуации, когда значение функции  $S$  для некоторых нейронов близко по модулю к 1, или когда модуль некоторых весов много больше 1. В этом случае для плавного уменьшения ошибки надо выбирать очень маленькую скорость обучения, но при этом обучение может занять неоправданно много времени.

Простейшим методом усовершенствования градиентного спуска является введение момента  $m$ , когда влияние градиента на изменение весов изменяется со временем. Тогда формула (13) примет вид

$$\Delta w_{ij}^n = -\eta \cdot \delta_j^n x_i^n + \mu \Delta w_{ij}^n (t-1). \quad (13.1)$$

Дополнительным преимуществом от введения момента является способность алгоритма преодолевать мелкие локальные минимумы.

Учитывая все преимущества нейронной сети для распознавания сложных образов, данная модель чрезвычайно подходит для обеспечения психодиагностики по физиологическим параметрам.

Очевидно, что диагностика эмоциональной компетентности по физиологическим параметрам составляет большие трудности по многим причинам. Во первых, это нелинейность взаимосвязей между физиологическими параметрами и психологическими характеристиками эмоциональной компетентности, представляющей один из высших механизмов регуляции целостного поведения человека. Гипотетически можно предположить, что различные показатели вегетативной и центральной регуляции отражают деятельность определенных функциональных систем, входящих в состав множества функций регуляции поведения, в которые, в частности, входит и эмоциональная регуляция. Таким образом, создается некая иерархическая система, нижним этажом которой представлены элементарные физиологические параметры, на промежуточном этапе находятся физиологические функциональные системы, обеспечивающие различные регуляторные и когнитивные функции, на верхнем же этапе находятся различные когнитивные и регуляторные способности которые можно измерить с помощью психологических тестов. Особенности такой модели будет то, что средний этаж этой схемы будет зависеть от группы физиологических переменных, в то же время любое психологическое свойство также будет зависеть от группы функциональных систем, находящихся на промежуточном этапе. Построение модели такой системы возможно с помощью математического моделирования нейронных сетей. При этом нейронная сеть будет представлять собой трехслойный перцептрон, на входной слой которого подаются значения физиологических переменных, выходным слоем перцептрона будут психо-

логические характеристики. Промежуточный же слой будет представлять физиологические факторы или функциональные системы, лежащие в основе регуляции поведения человека и отражения окружающего мира. Последовательно подавая на вход этого перцептрона физиологические переменные и обучая его методом обратного распространения ошибки рассогласования между результатом распознавания психологической характеристики и самой психологической характеристикой, измеренной по тесту, можно построить такую модель, которая позволит достаточно успешно диагностировать профиль психологических характеристик эмоциональной компетентности по физиологическим параметрам. При этом, структура такой обученной нейронной сети, отраженная в матрице связей между ячейками разных слоев (нейронами), может достаточно адекватно отражать реальные взаимосвязи между функциональными системами (факторами) и психологическими и физиологическими параметрами.

В качестве входных параметров данной модели мы выбрали факторные оценки, вычисленные из первичных физиологических параметров, рассмотренные в работе [1]. Большим преимуществом использования таких факторных оценок в нейронной модели является их независимость между собой (ортогональность факторов) и близость к нормальному распределению (нормализованность) этих оценок. В указанной работе с помощью факторного анализа проведенного над матрицей корреляций, связывающих показатели функционального состояния КГМ (кора головного мозга) (выраженные в спектральных плотностях различных ритмов ЭЭГ) и параметры вегетативных реакций (отраженные в изменении ЧСС (частота сердечных сокращений), АД (артериальное давление) и КГР (кожно-гальваническая реакция) на эмоциогенные и интеллектуальные нагрузки). В результате можно выделить следующие факторы.

1. Снижения активации кожно-кинестетического анализатора на интеллектуальную нагрузку.
2. Ритмичности передних и латеральных отделов коры мозга.
3. Реактивности НС (нервной системы).
4. Сосудистых реакций на вербальные интеллектуальные задачи.
5. Распространения низкочастотных ритмов в медиальных лобных и затылочных отделах.
6. Изменения АД на эмоциогенную нагрузку.
7. Изменения ЧСС на эмоциогенную нагрузку.
8. Распределения альфа ритма в КГМ.
9. Активации мозга за счет высокочастотных ритмов с акцентом на левое полушарие.
10. Кожно-гальванической реакции на эмоциогенное воздействие.
11. Тахикардии во время выполнения интеллектуальных задач.
12. Повышения высокочастотных ритмов во фронтальных областях КГМ.

Данные факторы отражают основные тенденции изменения функционального состояния ЦНС и ВНС (вегетативная нервная система) у детей от 4 до 6 лет. Факторные оценки, вычисленные для каждого ребенка являлись исходной физиологической информацией, которая загружалась при обучении в 1 слой нейронной сети обратного распространения ошибки, 3 слой представлял параметры эмоциональной компетентности детей.

В табл. 1 представлены результаты построения такой модели с точки зрения качества распознавания. В первой колонке таблицы, отражены психологические характеристики эмоциональной компетентности, которые диагностировались с помощью нейронной сети. Вторая колонка представляет средний квадрат ошибки диагностики показателя, отраженного в первой колонке. Третья колонка отражает максимальный квадрат ошибки, допущенной нейронной сетью при диагностике одного из испытуемых. В четвертой колонке таблицы находятся коэффициенты корреляции, между спрогнозированным нейронной сетью значением, и данной способностью ребенка, измеренной психологическими тестами.

Как видно из табл. 1, квадрат ошибки варьирует от 0,558 до 1,286, учитывая, что разброс показателей по тестам от 0 до 5. Корреляция варьирует от 0,629 до 0,776.

Кроме этого проводилась кросс-проверка надежности диагностики, представленной в табл. 1 в 4–6 колонках.

При кросс-проверке математическая модель нейронной сети обучалась на выборке, исключаяющей одного из испытуемых, а диагноз строился по этому испытуемому. Такая процедура была проведена 48 раз с исключением каждого испытуемого. Таким образом, корреляции и средние

значения считались для испытуемых, исключенных из группы обучения, то есть фактически для внешней выборки. Качество распознавания, при кросс-проверке, варьирует по квадрату ошибки от 0,646 до 1,573, а по коэффициенту корреляции от 0,447 до 0,706. Все корреляции значимы на уровне  $p < 0,01$ .

В табл. 2 отражена структура модели, объединяющая входной слой физиологических параметров и второй слой сети, отражающий функциональные системы, регулирующие эмоциональную компетентность ребенка и представленные нормализованными весовыми коэффициентами межнейронных связей. Близкое к нулю значение, будет отражать отсутствие связи входного физиологического параметра с данной функциональной системой.

Положительный вес (выше 1) отражает сильную взаимосвязь данного показателя с активацией той или иной функциональной системы (связь с этим показателем больше среднего на 1 сигму). Отрицательное значение связи (меньше -1) отражает сильную обратную зависимость от этой системы. Следовательно, интерпретация сущности функциональной системы, отраженной тем или иным нейроном промежуточного слоя, как и в факторном анализе, должна осуществляться с учетом не только весовых коэффициентов, но и теоретических знаний о закономерностях физиологической регуляции поведения человека, а также уточняться на основе связей с выходным психологическим слоем данной модели. В первом столбце (см. табл. 2) отражены входные факторы, отражающие совокупности физиологических параметров, описанных выше. В первой строке таблицы находятся предположительные названия функциональных систем, осуществляющих регуляцию эмоциональной компетентности ребенка (табл. 3).

Таблица 1

Показатели качества диагностики эмоциональной компетентности с помощью математической модели нейронной сети

Показатели ЭК	Качество распознавания в выборке обучения			Качество распознавания при кросс-проверке		
	Средний квадрат ошибки	Максимальный квадрат ошибки	Корреляция с диагнозом	Средний квадрат ошибки	Максимальный квадрат ошибки	Корреляция с диагнозом
Узнавание эмоций по рисунку	1,286	10,361	0,629	1,573	10,347	0,503
Выражение эмоций мимикой	0,691	2,911	0,664	0,933	2,911	0,540
Узнавание эмоций по голосу	1,434	8,577	0,440	1,481	8,577	0,483
Выражение эмоций голосом	0,583	4,742	0,741	1,193	5,302	0,447
Выражение эмоций рисунком	0,685	3,521	0,514	0,746	3,521	0,465
Эмоциональная компетентность	0,558	2,187	0,776	0,646	3,076	0,706

Структура взаимосвязей между входным слоем нейронной сети, отражающей физиологические параметры с промежуточным слоем нейронной сети

Нейроны	Гомеостатическая регуляция сердечной ритмики созревающими мозговыми механизмами	Отражение гомеостатической регуляции сосудистых реакций с созревающими мозговыми механизмами	Снижение доминирования КГМ у старших девочек	Преобладание гормональных подкорковых влияний на формирование эмоциональной регуляции у мальчиков	Общее снижение активации НС у некоторых детей	Снижение ориентировочной реакции при снижении ритмичности КГМ	Снижение ФС мозга в медиальных отделах и повышения вегетативных	Десинхронизация медиальных медленных ритмов и повышение вегетативных у мальчиков
Снижение активации кожно-кинестетического анализатора на интеллектуальную нагрузку	0,15	-0,38	-0,06	-0,46	-0,48	-1,22	0,51	0,52
Ритмичности передних и латеральных отделов коры головного мозга	-0,58	0,36	0,77	1,45	0,89	-1,01	0,71	0,81
Реактивность нервной системы	-0,33	-1,38	-0,95	-1,34	-0,96	0,25	-2,35	-0,58
Сосудистые реакции на вербальные интеллектуальные задачи	-0,49	-0,42	-0,70	0,54	-1,11	0,64	1,59	1,10
Распространение низкочастотных ритмов в медиальных лобных и затылочных отделах	-0,29	0,74	0,48	0,54	1,36	0,59	1,01	-1,12
Изменение артериального давления на эмоциогенную нагрузку	-0,57	-1,08	-0,59	0,60	-0,98	-0,63	0,58	1,34
Изменение ЧСС на эмоциогенную нагрузку	-0,20	0,27	0,83	0,33	0,14	0,48	-0,07	-0,54
Распределение альфа ритма в коре головного мозга	1,98	1,63	-1,28	-0,15	-0,42	0,73	0,69	-0,90
Активация мозга за счет высокочастотных ритмов с акцентом на левое полушарие	-0,27	-0,15	0,01	-1,32	0,71	-1,22	-0,29	-0,74
Кожно-гальваническая реакция на эмоциогенное воздействие	0,22	0,35	-0,75	-1,75	-0,52	-1,55	-0,50	-1,22
Тахикардия во время выполнения интеллектуальных задач	-1,09	-0,74	0,39	1,04	0,83	0,77	0,12	1,15
Повышение высокочастотных ритмов во фронтальных областях коры головного мозга	-0,76	-1,23	0,12	-0,66	-1,19	-0,07	-1,60	-1,25
Возраст	2,63	2,22	2,72	-0,35	2,17	2,24	-0,63	-0,14
Пол	-0,38	-0,19	-1,00	1,54	-0,45	0,01	0,21	1,56

1 НЕЙРОН. Вторая колонка таблицы отражает взаимосвязи физиологических параметров с первой функциональной системой. Основные влияния эта система испытывает от ритмичности различных областей КГМ в области альфа ритмов, а так же возраста и характеристик тахикардии на интеллектуальную нагрузку (с отрицательным весом). Можно сказать, что остальные параметры отражают

снижение спектральной плотности как медленных, так и быстрых ритмов КГМ, уменьшение сосудистых реакций, некоторое повышение ЧСС и КГР на эмоциогенную нагрузку. Эту функциональную систему можно назвать гомеостатическая регуляция сердечной ритмики созревающими мозговыми механизмами (созревание КГМ без централизации регуляции сердечной ритмики).

Таблица 3

Структура взаимосвязей между промежуточным слоем нейронной сети и выходным слоем, отражающим параметры эмоциональной компетентности

Нейроны	Узнавание эмоций по рисунку	Выражение эмоций мимикой	Узнавание эмоций по голосу	Выражение эмоций голосом	Выражение эмоций рисунком	Эмоциональная компетентность
Гомеостатическая регуляция сердечной ритмики созревающими мозговыми механизмами	0,86	0,05	0,04	0,23	1,59	0,57
Отражает гомеостатическую регуляцию сосудистых реакций с созревающими мозговыми механизмами	-0,25	-0,39	0,68	1,00	0,51	0,14
Снижение доминирования КГМ у старших девочек	0,69	1,38	0,84	1,65	1,08	0,73
Преобладание тормозящих подкорковых влияний на формирование эмоциональной регуляции у мальчиков	-0,64	-0,28	-0,85	-1,27	-0,85	-0,26
Общее снижение активации НС у некоторых детей	0,88	1,30	1,05	0,28	-1,00	1,00
Снижение ориентировочной реакции при снижении ритмичности КГМ	1,25	0,65	0,88	0,10	0,27	1,00
Снижение ФС мозга в медиальных отделах и повышения вегетативных	-1,64	-1,52	-0,73	-0,48	-1,49	-1,61
Десинхронизация медиальных медленных ритмов и повышения вегетативных у мальчиков	-1,15	-1,19	-1,92	-1,51	-0,12	-1,57

**2 НЕЙРОН.** Отражает функциональную систему, активность которой связана с несколькими показателями центральной и вегетативной регуляции. Как и предыдущая функциональная система, данная система тесно связана с созреванием КГМ, отражающимся, в увеличении спектральной плотности в области альфа ритмов при увеличении возраста. Однако большую роль в этой системе играет снижение реактивности НС, отраженное в понижении спектральной плотности в области бета ритмов и уменьшении реакции артериального давления. Следовательно, данный нейрон отражает гомеостатическую регуляцию сосудистых реакций с созревающими мозговыми механизмами (созревание КГМ без централизации сосудистых реакций).

**3 НЕЙРОН.** Отражает активацию функциональной системы у девочек старшего возраста, в отличие от мальчиков у которых данная система в этом возрастном диапазоне не формируется.

Данная система характеризуется низким созреванием КГМ (снижение спектральной плотности альфа-ритмов), возможно связанным с сильным влиянием подкорковых структур, которое отражается в незначительном увеличении спектральной плотности в области дельта- и тета-ритмов как в латеральных и передних, так и в центральных и задних отделах. Следовательно, данную систему можно назвать снижением доминирования КГМ у старших девочек.

**4 НЕЙРОН.** Отражает функциональную систему, формирующуюся у мальчиков. Эта функциональная система связана в большей степени с из-

менениями спектральных плотностей в области высокочастотных (бета- и гамма-ритмов) и низкочастотных (дельта- и тета-ритмов). При этом увеличение активации данного нейрона связано с увеличением медленно-волновой ритмики и снижением высокочастотной. Таким образом, эту систему можно назвать преобладание тормозящих подкорковых влияний на формирование эмоциональной регуляции у мальчиков.

**5 НЕЙРОН.** Отражает формирование функциональной системы, связанной с преобладанием медленноволновой ритмики в медиальных, лобных и затылочных отделах мозга, снижением высокочастотной ритмики в лобных отделах, а также снижением сосудистых реакций на вербальные интеллектуальные задачи. Влияние этих отличий на эмоциональную компетентность происходит при увеличении возраста. Такие изменения могут отражать общее снижение активации НС у некоторых детей старшего дошкольного возраста.

**6 НЕЙРОН.** Отрицательно связан с 2 факторами, включающими КГР, также данная функциональная система тесно связана со снижением спектральной плотности в области бета 2 и гамма-ритмов, а также в диапазоне дельта-, тета- и в меньшей степени альфа-ритмов в латеральных и передних отделах КГМ. Учитывая большое влияние этой функциональной системы на КГР и зависимость от функционального состояния мозга, данную систему можно назвать снижением ориентировочной реакции при снижении ритмичности КГМ у детей старшего дошкольного возраста.



**7 НЕЙРОН.** Тесно связан с факторами отражающими медленноволновую ритмику и отрицательно связан с факторами, отражающими высокочастотную ритмику. Также он имеет положительную взаимосвязь с фактором сосудистых реакций на вербальные задачи. Следовательно, увеличение активации этого нейрона отражает снижение функционального состояния мозга при увеличении сосудистых реакций. Деактивация этого нейрона будет отражать противоположные явления.

**8 НЕЙРОН.** Связан со снижением спектральных мощностей в области медленно-волновой ритмики и высокочастотной ритмики в сагиттальных, лобных и затылочных отделах. Кроме этого активация данного нейрона связана со снижением сосудистых реакций на эмоциогенные нагрузки, при уменьшении таковых на вербальные интеллектуальные задачи и со снижением КГР на эмоциогенную нагрузку.

**Описание нейронов промежуточного слоя с точки зрения взаимосвязи с характеристиками эмоциональной компетентности**

**1 НЕЙРОН.** Эта функциональная система положительно влияет практически на все показатели эмоциональной компетентности с особым акцентом на выражении эмоций рисунком. При выражении через рисунок обычно используется слабо ассоциированная позиция, что позволяет решать поставленную задачу с большим сохранением гомеостаза.

**2 НЕЙРОН.** Данный нейрон связан максимальными связями с выражением, голосом и узнаванием по голосу, имеет небольшую связь с выражением эмоций рисунком. Следовательно, данная функциональная система отражает симпатическую регуляцию сосудистых реакций параллельно с созреванием КГМ, и снижение реактивности НС влияет в большей степени на ЭК, связанную с вербальной коммуникацией. Возможно, снижение реактивности НС и парасимпатическая регуляция НС позволяют более спокойно вести себя в процессе речевой коммуникации, т.е. в более ассоциированной позиции в ситуациях непосредственно вызывающих эмоции.

**3 НЕЙРОН.** Сильно связан с успешностью эмоционального выражения голосом, мимикой и рисунком. Следовательно, данная ФС (функциональная система), отражающая снижение влияния КМ с некоторым преобладанием подкорки увеличивающееся с возрастом у некоторой группы девочек эта ФС может отражать преобладание подкорковых влияний в регуляции поведения, что будет способствовать лучшему выражению своего эмоционального состояния за счет лучшей репрезентации своих потребностей при увеличении влияния подкорковых структур на поведение.

**4 НЕЙРОН.** Данная ФС формируется в большей степени у мальчиков за счет отрицательной взаимосвязи с возрастом. Если инвертировать веса данного нейрона, то можем более адекватно пред-

ставить себе эмоциональную компетентность, обеспечивающую успешное выражение голосом, выражение рисунком и узнавание по голосу. Инверсия весов, соединяющих данный нейрон с предыдущим слоем, показывает, что данная ФС реализуется в большей степени у девочек с незначительным влиянием возраста, и связана с процессами активации КГМ отраженных в повышении высокочастотных ритмов и снижении низкочастотных. А также с одновременным повышением реактивности НС, но при преобладании парасимпатической регуляции сердечной ритмики.

**5 НЕЙРОН.** Сильно связан с выражением эмоций мимикой, а также с узнаванием эмоций по голосу. Наблюдается обратная корреляция с выражением эмоций рисунком. Таким образом, дети с преобладанием данной ФС более компетентны в ассоциированной позиции и при этом с акцентом на невербальные формы выражения эмоций. За этими способностями стоит ФС, которая отражает снижение возбуждения в лобных отделах КГМ, снижение высокочастотных и повышение низкочастотных ритмов, а также парасимпатические сосудистые реакции на интеллектуальную нагрузку. Очевидно, что такая ФС отражает доминирование подкорковых структур при повышении эмоциональной компетентности с возрастом.

**6 НЕЙРОН.** Данная ФС по своим характеристикам отражает снижение реакции на внешние стимулы при увеличении процесса переработки информации в НС, что может соответствовать углубленности в себя и интравертным признакам, что соответствует признакам ЭК, проявляющимся в узнавании эмоций и не выражении их.

**7 НЕЙРОН.** Связан с показателями ЭК отрицательно, следовательно, чтобы выявить ФС, отражающую ЭК, необходимо инвертировать веса данного нейрона как в последнем слое, так и в первом. Таким образом, оказывается, что все веса связаны с данным нейроном положительно и имеют выраженные связи. Этот показатель ЭК будет реализовывать ФС, отраженная повышением ФС мозга при снижении сосудистых реакций на интеллектуальные задачи, т.е. парасимпатическая регуляция сосудистых реакций при интеллектуальных нагрузках. Данная ФС, видимо, отражает оптимальное состояние ребенка для осуществления ЭК (эмоциональной компетентности).

**8 НЕЙРОН.** Чтобы выявить ФС, отражающую ЭК, так же инвертируем веса данного нейрона и в первом и в последнем слое. В результате мы получаем, что данный нейрон связан со всеми характеристиками ЭК с большим акцентом на речевую коммуникацию. ФС, отражающая данный нейрон, характерна в большей степени для девочек и отражает повышение амплитуды как высокочастотных, так и низкочастотных ритмов при парасимпатической регуляции сердечно-сосудистых реакций как на интеллектуальные, так и на эмоциогенные нагрузки.

**Выводы.** Результатом построения предложенной модели было описание 8 функциональных систем, тем или иным способом реализующих эмоциональную компетентность детей. Основными факторами, определяющими развитие эмоциональной компетентности детей, являются возраст и пол ребенка, которые могут отражаться как на процессе созревания коры головного мозга, так и в оптимизации вегетативных реакций на интеллектуальные и эмоциогенные нагрузки в сторону парасимпатической регуляции. При этом существуют различные способы реализации эмоциональной компетентности, которые можно разделить на два больших класса: 1) эмоциональная компетентность с повышением реактивности нервной системы, требующая высокой активации коры головного мозга с преобладанием в большей степени когнитивных функций в реализации эмоциональной компетентности; 2) преобладание подкорковых влияний на организацию эмоционального поведения ребенка, отражающимся в лучшем выражении эмоций, особенно в непосредственной вербальной и невербальной коммуникации.

Особым механизмом реализации эмоциональной компетентности является преобладание процессов переработки информации в коре головного мозга и снижение реактивности нервной системы на внешние стимулы, возможно коррелирующим с таким психологическим способом адаптации, как интроверсия.

Обращает внимание, что большинство описанных функциональных систем реализуют эмоциональную компетентность у мальчиков и девочек различными способами.

Данный метод диагностики, при детальной его проверке, позволит формализовать диагностику эмоциональной компетентности, учитывая физиологические характеристики ребенка, что в свою очередь позволит определять зону ближайшего развития, учитывая физиологические механизмы, стоящие за этими явлениями.

### Литература

1. Вишневетская, Е.В. Психофизиологические корреляты эмоциональной компетентности детей дошкольного возраста / Е.В. Вишневетская, И.А. Горбунов // *Материалы 4 съезда РПО.* – Ростов на Дону, 2007.
2. Вишневетская, Е.В. Физиологические механизмы эмоциональной компетентности детей дошкольного возраста / Е.В. Вишневетская, И.А. Горбунов // *20 съезд физиологического общества им. И.П. Павлова.* – М., 2007. – С. 179.
3. Горбунов, И.А. *Диагностические возможности психофизиологических характеристик человека: автореф. дис. ... канд. психол. наук / И.А. Горбунов.* – СПб., 2005.
4. *Статистические модели нейронных сетей / В.С. Доценко, Л.Б. Иоффе, М.В. Фейгельман, М.В. Цодыкс // Итоги науки и техники. Серия «Физические и математические модели нейронных сетей». Т. 1, ч. 1: Спиновые стекла и нейронные сети / ред. А.А. Веденов.* – М.: ВИНТИ, 1990. – С. 4–43.
5. Лафренье, П. *Эмоциональное развитие детей и подростков / П. Лафренье.* – СПб.: Прайм-Еврознак, 2004.
6. Осовский, С. *Нейронные сети для обработки информации / С. Осовский.* – М.: Финансы и статистика, 2002.
7. Cicchetti, D., & Ganiban, J. *The organization and coherence of developmental processes in infants and children with Down syndrome. In R.M. Hodapp, J. Burack & E. Zigler (Eds.), Issues in the developmental approach to mental retardation. New York: Cambridge University Press, 1990.* – P. 169–225.
8. Scherer, K.R. *Vocal correlates of emotional arousal and affective disturbance. In H. Wagner & A. Manstead (Eds.), Handbook of social psychophysiology. New York: Wiley, 1989.* – P. 165–197.

Поступила в редакцию 10 декабря 2008 г.

# СТАТИЧЕСКИЕ МЕТОДЫ ОЦЕНКИ НАДЕЖНОСТИ ПРОГРАММНОГО ОБЕСПЕЧЕНИЯ

О.О. Павловская

## STATIC METHODS OF ASSESSMENT OF SOFTWARE

O.O. Pavlovskaya

В статье показаны особенности использования математического аппарата прикладной теории надежности технических систем для определения показателей надежности программного обеспечения (ПО). Этот материал необходим студентам, аспирантам, практикам для получения возможности оценки влияния различных внутренних и внешних факторов на качество функционирования ПО.

*Ключевые слова:* оценка надежности, теория надежности, программное обеспечение.

The article shows the features of the use of mathematical apparatus applied the theory of reliability of technical systems for determining the reliability of software (SW). This material is needed for students, graduate students, practitioners to be able to assess the impact of various internal and external factors on the quality of the software.

*Keywords:* reliability estimation, reliability theory, software.

### Введение

В настоящее время актуальной является проблема исследования надежности программного обеспечения (ПО). В рамках данной проблемы можно выделить ряд частных задач, таких как:

- определение основных факторов, влияющих на надежность ПО;
- разработка методов оценки надежности ПО;
- разработка методов, обеспечивающих достижение заданного уровня надежности ПО.

Под надежностью ПО понимается его способность безотказно выполнять определенные функции при заданных условиях в течение заданного периода времени с достаточно большой вероятностью.

*Составляющие функциональной надежности ПО*

1. Безотказность – свойство программы выполнять свои функции во время эксплуатации.

2. Работоспособность – свойство программы корректно (как ожидает пользователь) работать весь заданный период эксплуатации.

3. Безопасность – свойство программы быть неопасной для людей и окружающих систем.

4. Защищенность – свойство программы противостоять случайным или умышленным вторжениям в нее.

*Факторы, влияющие на надежность ПО, делят на 2 группы.*

1. Внутренние (ошибки проектирования при постановке задач; ошибки алгоритмизации задач; ошибки программирования; недостаточное качество средств защиты).

2. Внешние (ошибки персонала при эксплуатации; искажения информации в каналах связи; сбой и отказы аппаратуры ЭВМ).

Процентные частоты появления ошибок в ПО по типам ошибок представлены в таблице [2, с. 103].

Частота появления ошибок в ПО

Тип ошибки	Частота появления, %
Неполная или ошибочная спецификация	28
Отклонение от спецификации	12
Пренебрежение правилами программирования	10
Ошибочная выборка данных	10
Ошибочная логика или последовательность операций	12
Ошибочные арифметические операции	9
Нехватка времени для решения	4
Ошибка обработки прерываний	4
Ошибка в исходных данных	3
Неточная запись	8

Как видно из таблицы, основное количество ошибок делается из-за неверной спецификации и следует обращать особое внимание на проведение тестирования ПО с намерением, во-первых, найти и исправить ошибки в ПО, снижающие надежность последнего, а во-вторых, определить надежность ПО.

Оценка надежности ПО, как и любого другого объекта, начинается с определения признаков, по которым будет оцениваться надежность объекта (так называемых критериев надежности).

### 1. Критерии надежности ПО

При анализе надежности ПО используются традиционные для технических систем критерии надежности.

1. Вероятность безотказной работы  $P(t_s)$ .
2. Вероятность отказа  $Q(t_s)$ .
3. Интенсивность отказов  $\lambda(t)$ .
4. Средняя наработка на отказ  $T_0$ .
5. Среднее время восстановления  $T_B$ .
6. Коэффициент готовности  $K_T$ .

Традиционные критерии надежности ПО характеризуют наличие ошибок программы (производственных дефектов), но ни один из них не характеризует характер этих ошибок и возможные их последствия. Поэтому вводится дополнительный критерий надежности ПО – *средняя тяжесть ошибок* (СТО) [1, с. 122], определяемый выражением

$$\text{СТО} = \frac{1}{Q} \sum_{i=1}^m b_i p_i z_i,$$

где  $Q$  – вероятность сбоя ПО;  $b_i$  – функция принадлежности тяжести последствий ошибки, возникшей при  $i$ -м наборе входных данных, к максимально тяжелым последствиям;  $p_i$  – вероятность ввода  $i$ -го набора входных данных при эксплуатации ПО;  $z_i$  – дихотомическая переменная, равная 1, если при  $i$ -м наборе входных данных был зафиксирован сбой, и 0 в противном случае;  $m$  – общее число наборов входных данных.

Значение показателя надежности СТО лежит на интервале [0; 1]. Чем ближе значение СТО к единице, тем тяжелее последствия ошибок ПО, и тем менее надежна программа. Близость СТО к нулю показывает незначительность последствий ошибок программы.

Введение данного критерия надежности ПО позволяет характеризовать не столько безошибочность ПО, сколько его безопасность. Однако следует помнить, что значение этого критерия субъективно и может быть различным для одного и того же программного продукта в зависимости от области его применения. Это объясняется тем, что при использовании конкретного ПО, например для выполнения студенческих расчетов и для выполнения конструкторских расчетов в космической промышленности последствия ошибок программы – несопоставимы. В ряде случаев, если к ПО предъявляются жесткие требования, лучше оценивать максимальную тяжесть ошибок ПО.

Основным средством определения показателей надежности являются *модели надежности*, под которыми понимают математическую модель, построенную для оценки зависимости надежности от заранее известных или оцененных в ходе создания ПО параметров.

### 2. Модели надежности ПО

Различают модели ПО статические и динамические. Статические модели принципиально отличаются от динамических прежде всего тем, что в них появление отказов не связывают со временем появления ошибок в процессе тестирования, а учитывают только зависимость количества ошибок от числа тестовых прогонов (по области ошибок) или зависимость количества ошибок от характеристики входных данных (по области данных). В динамических же моделях поведение ПО (появление отказов) рассматривается во времени.

Рассмотрим статические модели надежности ПО [3, с. 167].

#### 2.1. Статическая модель надежности Миллса

Использование этой модели предполагает необходимость перед началом тестирования искусственно вносить в программу (засорять) некоторое количество известных ошибок. Ошибки вносятся случайным образом и фиксируются в протоколе искусственных ошибок. Специалист, проводящий тестирование, не знает ни количества ошибок, ни характера внесенных ошибок до момента оценки показателей надежности по модели Миллса. Предполагается, что все ошибки (как естественные, так и искусственно внесенные) имеют равную вероятность быть найденными в процессе тестирования.

Тестируя программу в течение некоторого времени, собирается статистика об ошибках. В момент оценки надежности по протоколу искусственных ошибок все ошибки делятся на собственные и искусственные.

Модель надежности Миллса образуется двумя связанными между собой по смыслу соотношениями.

*Первое соотношение*  $N = S \cdot n / V$  предсказывает  $N$  – первоначальное количество ошибок в программе. В данном соотношении, которое называется формулой Миллса,  $S$  – количество искусственно внесенных ошибок,  $n$  – число найденных собственных ошибок,  $V$  – число обнаруженных к моменту оценки искусственных ошибок.

Предположим, что в программе имеется  $K$  собственных ошибок, и внесем в нее еще  $S$  ошибок. Если в процессе тестирования были обнаружены все  $S$  внесенных ошибок и  $n$  собственных ошибок, то по формуле Миллса мы предполагаем, что первоначально в программе было  $N = n$  ошибок.

*Второе соотношение* используется для установления доверительного уровня прогноза  $C$  (вероятности, с которой можно высказать предположение об  $N$ ). Величина  $C$  является мерой доверия к модели.

Если обнаружены все искусственно рассеянные ошибки ( $V = N$ ) вероятность того, что значение  $N$  найдено правильно, можно рассчитать по следующему соотношению:

$$C = S / (S + K + 1).$$

В случае, когда оценка надежности производится до момента обнаружения всех  $S$  рассеянных ошибок, величина  $C$  рассчитывается по модифицируемой формуле

$$C = \left( \frac{S}{V-1} \right) / \left( \frac{S+K+1}{V+K} \right),$$

где числитель и знаменатель формулы являются биномиальными коэффициентами вида

$$\frac{a}{b} = \frac{a!}{b!(a-b)!}.$$

Например, если утверждается, что в программе нет ошибок, а к моменту оценки надежности обнаружено 5 из 10 рассеянных ошибок и не обнаружено ни одной собственной ошибки, то вероятность того, что в программе действительно нет ошибок, будет равна

$$C = \left( \frac{10}{4} \right) / \left( \frac{11}{5} \right) = \frac{10! \cdot 5! \cdot 6!}{4! \cdot 6! \cdot 11!} = 0,45.$$

Если при тех же исходных условиях оценка надежности производится в момент, когда обнаружены 8 из 10 искусственных ошибок, то вероятность того, что в программе не было ошибок, увеличивается до 0,73.

В действительности модель Миллса можно использовать для оценки  $N$  после каждой найденной ошибки. Предлагается во время всего периода тестирования отмечать на графике число найденных ошибок и текущее значение для  $N$ .

Достоинством модели является простота применения математического аппарата, наглядность и возможность использования в процессе тестирования. Однако модель не лишена и ряда недостатков, самые существенные из которых – это необходимость внесения искусственных ошибок (этот процесс плохо формализуется) и достаточно вольное допущение величины  $K$ , которое основывается исключительно на интуиции и опыте человека, проводящего оценку, т.е. допускается большое влияние субъективного фактора.

## **2.2. Статическая модель надежности Липова**

Липов модифицировал модель Миллса, сделал же предположение, что и Миллса, т.е. что собственные и искусственные ошибки имеют равную вероятность быть найденными.

Тогда вероятность обнаружения  $n$  собственных и  $V$  внесенных ошибок равна

$$Q(n, V) = \frac{m}{n+V} \cdot q^{n+V} \cdot (1-q)^{m-n \frac{V \cdot \frac{N \cdot S}{n \cdot V}}{(N+S)}},$$

где  $m$  – количество тестов, используемых при тестировании;  $q$  – вероятность обнаружения ошибки в каждом из  $m$  тестов, рассчитываемая по формуле  $q = (n+V)/n$ ;  $S$  – общее количество искусственно внесенных ошибок;  $N$  – количество собственных ошибок, имеющихся в ПС до начала тестирования.

Для использования модели Липова должны выполняться следующие условия:  $N \geq n \geq 0$ ,  $S \geq V \geq 0$ ,  $m \geq n+V \geq 0$ .

### **Заключение**

1. Особенностью оценки надежности ПО является использование наряду с традиционными критериями надежности, такого специфического критерия, как СТО. Введение данного критерия надежности ПО позволяет характеризовать не только безошибочность ПО, но и его безопасность.

2. Основным средством определения показателей надежности являются модели надежности.

3. Статические модели не связывают появление отказов со временем появления ошибок в процессе тестирования, а учитывают только зависимость количества ошибок от числа тестовых прогонов (по области ошибок) или зависимость количества ошибок от характеристики входных данных (по области данных).

4. Модель Липова дополняет модель Миллса, дав возможность оценить вероятность обнаружения определенного числа ошибок к моменту оценки.

### **Литература**

1. Игнатъев, М.Б. *Активные методы обеспечения надежности алгоритмов и программ* / М.Б. Игнатъев, В.В. Фильчаков, Л.Г. Осовецкий. – СПб.: Политехника, 1992.
2. Липаев, В.В. *Надежность программных средств* / В.В. Липаев. – М.: СИНТЕГ, 1998.
3. Половко, А.М. *Основы теории надежности* / А.М. Половко, С.В. Гуров. – 2-е изд., перераб. и доп. – СПб.: БХВ-Петербург, 2006.

*Поступила в редакцию 2 октября 2007 г.*

# НЕПАРАМЕТРИЧЕСКОЕ ОБНАРУЖЕНИЕ УЗКОПОЛОСНОГО РАДИОСИГНАЛА С НЕИЗВЕСТНОЙ НЕСУЩЕЙ ЧАСТОТОЙ НА ФОНЕ ГАУССОВСКИХ ШУМОВ

Ю.Т. Карманов, А.Н. Николаев

## NONPARAMETRIC NARROWBAND DETECTION OF THE SIGNAL WITH UNKNOWN CARRIER FREQUENCY WITH A BACKGROUND GAUSSIAN NOISE

Y.T. Karmanov, A.N. Nikolaev

Предложен алгоритм непараметрического обнаружения узкополосного сигнала в условиях гауссовских шумов. Основными достоинствами алгоритма являются отсутствие необходимости в априорной информации о несущей частоте сигнала и независимость порога обнаружения от уровня шумов. При помощи моделирования получены характеристики обнаружения предлагаемого алгоритма.

*Ключевые слова:* обнаружение радиосигнала, узкополосные радиосигналы.

An algorithm for nonparametric detection of narrow-band signal in a Gaussian noise is presented. The main advantages of the algorithm is a no need for a priori information about the carrier frequency signal and the independence of the threshold of detection of the level of noise. With the help of simulation were obtained characteristics of the proposed detection algorithm.

*Keywords:* signal detection, narrow-band radio-signals.

### Введение

Оптимальное обнаружение узкополосного радиосигнала на фоне гауссовских шумов требует априорного знания параметров обнаруживаемого радиосигнала, в частности его несущей частоты  $f_c$ , либо требует ее предварительной оценки. Кроме того, необходимо знать спектральную плотность действующих шумов.

Это не всегда приемлемо, поэтому представляют интерес непараметрические алгоритмы обнаружения сигнала, которые не требуют априорного знания характеристик обнаруживаемого радиосигнала и шумов, хотя они имеют худшие характеристики обнаружения сигналов, чем оптимальные.

В [1] рассмотрен класс «фазовых» обнаружителей, в которых для обнаружения используются «текущая фаза» наблюдаемой смеси сигнала и шума  $x(t)$  равная:

$$\varphi(t) = \arctg \frac{x^Q(t)}{x^I(t)},$$

где  $x^I(t), x^Q(t)$  – синфазная и квадратурная составляющие  $x(t)$ , полученные относительно опорного колебания с частотой  $f_0$ .

Показано, что в таких обнаружителях уровень вероятности ложных тревог  $P_F$  не зависит от спектральной плотности шума  $N_0$ .

Однако сам алгоритм обнаружения зависит от параметров обнаруживаемого сигнала, в частности от его несущей частоты  $f_c$ .

Рассматривались алгоритмы обнаружения, основанные на простом накоплении текущей фазы  $\varphi(t)$ . Их серьезным недостатком является зависимость вероятности правильного обнаружения  $P_D$  от значения несущей частоты сигнала. В частности при малых расстройках  $\Delta f_c = f_c - f_0$   $P_D \approx 0$ .

Актуальным является поиск алгоритмов непараметрического обнаружения узкополосных сигналов на фоне шумов, свободных от указанных недостатков.

Карманов Юрий Трофимович – д-р техн. наук, профессор, заведующий кафедрой цифровых радиотехнических систем ЮУрГУ; ea@drts.susu.as.ru

Николаев Андрей Николаевич – старший преподаватель кафедры цифровых радиотехнических систем ЮУрГУ; ea@drts.susu.ac.ru

Karmanov Yuri Trofimovich – PhD, professor, head of Digital radio-technical systems department of SUSU; ea@drts.susu.as.ru

Nikolaev Andrey Nikolaevich – senior lecturer of Digital radio-technical systems department of SUSU; ea@drts.susu.ac.ru

**Постановка задачи**

Пусть наблюдается колебание  $x(t)$ , состоящее из шумового колебания  $n(t)$  и появляющегося на его фоне в момент времени  $t = t_0$  узкополосного радиосигнала  $S(t)$  длительностью  $T_c$  в виде

$$S(t) = A_c(t) \cos[2\pi f_c t + \varphi_c(t)],$$

где  $A_c(t)$  – амплитуда сигнала.

Будем полагать.

1. Колебание  $n(t)$  является белым гауссовским шумом со спектральной плотностью  $N_0$ .

2. Значение несущей частоты  $f_c$  обнаруживаемого сигнала  $S(t)$  лежит в диапазоне  $\left[ f_0 - \frac{\Delta F}{2}, f_0 + \frac{\Delta F}{2} \right]$ , точное значение  $f_c$  неизвестно.

3. Существует такой интервал  $\Delta\tau_0$ , в течение которого значение внутрисигнальной модуляции не изменяется и  $\varphi_c(t) \equiv \varphi_{\Delta\tau_0}$  при  $t \in \Delta\tau_0$  и  $A_c(t) \equiv A_{\Delta\tau_0}$ .

4. Предварительно проведено выделение квадратур  $x^I(t), x^Q(t)$  колебания  $x(t)$  относительно несущей частоты  $f_0$ . Они имеют вид:

$$\begin{aligned} x^I(t) &= A_{\Delta\tau_0} \cos[2\pi \Delta f_c t + \varphi_{\Delta\tau_0}] + n^I(t) \\ x^Q(t) &= A_{\Delta\tau_0} \sin[2\pi \Delta f_c t + \varphi_{\Delta\tau_0}] + n^Q(t) \end{aligned} \quad \text{если}$$

сигнал есть ( $H_1$ );

$$\begin{aligned} x^I(t) &= n^I(t) \\ x^Q(t) &= n^Q(t) \end{aligned} \quad \text{если сигнала нет } (H_0).$$

Здесь  $n^I(t), n^Q(t)$  – квадратурные составляющие шума.

Задача состоит в обнаружении данного сигнала, т.е. определении момента его появления  $t = t_0$  и окончания  $t = t_0 + T_c$ .

**Алгоритм обнаружения**

Проведем дискретизацию квадратур колебания  $x(t)$  с частотой дискретизации  $f_d \geq \Delta F$  и представим наблюдаемое колебание в виде двух последовательностей отсчетов квадратур:

$$\begin{aligned} x^I(1), x^I(2), \dots, x^I(k), \dots \\ x^Q(1), x^Q(2), \dots, x^Q(k), \dots \end{aligned}$$

Перейдем к отсчетам текущей фазы колебания  $x(t)$ :

$$\varphi(k) = \arctg \frac{x^Q(k)}{x^I(k)},$$

получим последовательность отсчетов:

$$\varphi(1), \varphi(2), \dots, \varphi(k), \dots$$

Вычислим на основании их последовательность  $\Delta\varphi(k)$ :

$$\Delta\varphi(k) = \varphi(k) - \varphi(k - m), \quad (1)$$

где  $m$  – некоторый целый параметр. Получим:

$$\Delta\varphi(1), \Delta\varphi(2), \dots, \Delta\varphi(k), \dots$$

Сформируем из них статистику

$$Z_k = \frac{1}{R} \sum_{i=1}^R |\Delta\varphi(k - i + 1) - \Delta\varphi(k - i)|^2. \quad (2)$$

Решение о наличии ( $H_1$ ) или отсутствии ( $H_0$ ) сигнала в наблюдаемой смеси  $x(t)$  принимается на основе сравнения величины  $Z_k$  с порогом  $g$ :

$$\begin{aligned} H_0 \\ > \\ Z_k &< g. \\ < \\ H_1 \end{aligned} \quad (3)$$

Величина  $R$  в выражении (1) определяется:

$$R \approx \frac{\Delta\tau_0}{\Delta t_d},$$

где  $\Delta t_d = \frac{1}{f_d}$ .

Предлагаемый алгоритм основан на следующем.

1. При наличии обнаруживаемого сигнала, как известно [1], текущая фаза  $\varphi(k)$  случайна и имеет математическое ожидание и дисперсию:

$$M[\varphi(k)] = 2\pi \Delta f_c t_k + \varphi_{\Delta\tau_0};$$

$$D[\varphi(k)] = \frac{1}{d_c^2},$$

где  $d_c^2 = \frac{A_{\Delta\tau_0}^2}{N_0}$  – отношение сигнал-шум.

В таком случае при наличии сигнала величина (1) также случайна и имеет

$$M[\Delta\varphi(k)] = 2\pi \Delta f_c m \Delta t_d;$$

$$D[\varphi(k)] = \frac{2}{d_c^2}.$$

Тогда:

$$M[(\Delta\varphi(k - i + 1) - \Delta\varphi(k - i))] \equiv 0;$$

$$D[(\Delta\varphi(k - i + 1) - \Delta\varphi(k - i))] \equiv \frac{4}{d_c^2}.$$

Накапливая статистику (2), получим:

$$M\left[\frac{Z_k}{H_1}\right] = \frac{4}{d_c^2};$$

$$D\left[\frac{Z_k}{H_1}\right] \approx \frac{2 \cdot 16}{R(d_c^2)^2} = \frac{32}{R(d_c^2)^2}.$$

2. При отсутствии сигнала текущая фаза  $\varphi(k)$  распределена равномерно в интервале  $(-\pi, \pi)$  и не зависит от интенсивности шума  $N_0$ . Следовательно, и  $\Delta\varphi(k)$ , и статистика  $Z_k$  не зависят от  $N_0$ :

$$M\left[Z_k/H_0\right] = \frac{4 \cdot 4\pi^2}{12} = 13,1;$$

$$D\left[Z_k/H_0\right] \approx \frac{2 \cdot (16\pi^2)^2}{R \cdot 12^2} = \frac{346}{R}.$$

Таким образом, вероятность ложной тревоги  $P_F$  в обнаружителе (3) от интенсивности шумов не зависит и определяется только значением порога  $g$ .

Качество алгоритма обнаружения характеризуется семейством кривых  $P_D(d_c^2)$  (зависимости вероятности правильного обнаружения от соотношения сигнал-шум) для различных фиксированных значений  $P_F$ . Вычисление характеристик обнаружения предлагаемого алгоритма проводилось при помощи имитационного моделирования с использованием возможностей пакета Matlab [2].

### Результаты моделирования

Моделирование проводилось для набора исходных данных, приведенных в табл. 1.

Таблица 1  
Исходные данные моделирования

Параметр	$b$	$m$	$R$
Значение	8	1	16

Параметр  $b$  (см. табл. 1) определяет число разрядов, используемых при квантовании отсчетов квадратурных составляющих колебания  $x(t)$ .

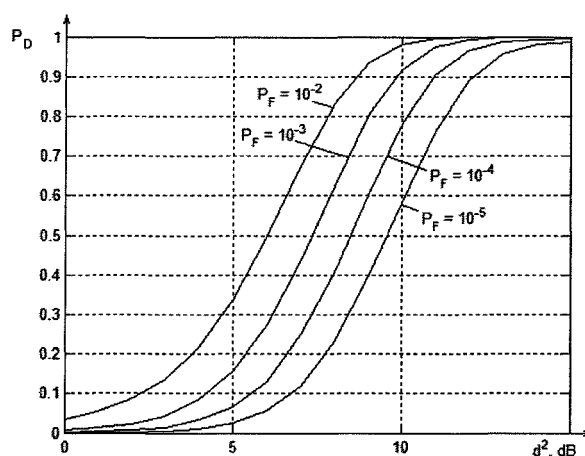
На первом этапе моделирования были получены значения порога для заданных уровней вероятности ложной тревоги  $P_F$  (табл. 2).

Таблица 2  
Результаты расчета порога обнаружения

$P_F$	$10^{-2}$	$10^{-3}$	$10^{-4}$	$10^{-5}$
Нормированный порог $g/g_{\max}$	0,0849	0,0632	0,0472	0,0353

В табл. 1  $g_{\max} = 2^{b-1}$  – максимальное значение порога для заданного числа разрядов, используемых при квантовании.

Для рассчитанных значений порога на основе (2) и (3) было получено семейство кривых обнаружения (см. рисунок).



Кривые обнаружения

### Выводы

Предлагаемый алгоритм обнаружения не требует априорной информации о несущей частоте сигнала. Вероятность ложной тревоги в таком алгоритме не зависит от уровня шумов и определяется только значением порога.

Характеристики обнаружения несущественно отличаются от характеристик оптимального обнаружителя [3]. Это отличие составляет несколько дБ.

### Литература

1. Левин, Б.Р. Теоретические основы статистической радиотехники / Б.Р. Левин. – 3-е изд., перераб. и доп. – М.: Радио и связь, 1989. – 625 с.
2. Гультияев, А. Визуальное моделирование в среде Matlab: учебный курс / А. Гультияев. – СПб.: Питер, 2000. – 432 с.: ил.
3. Перов, А.И. Статистическая теория радиотехнических систем: учебное пособие для вузов / А.И. Перов. – М.: Радиотехника, 2003. – 400 с.: ил.

Поступила в редакцию 16 марта 2009 г.



# СПОСОБ ЗАЩИТЫ РЛС СО СЛОЖНЫМ СИГНАЛОМ ОТ ИМИТИРУЮЩЕЙ ПОМЕХИ

Ю.Т. Карманов, Г.А. Непомнящий

## ONE WAY TO PROTECT THE RADAR OF A COMPLEX SIGNALS FROM SIMULATING INTERFERENCE

Y.T. Karmanov, G.A. Nepomnyashiy

Предлагается статистический способ защиты радиолокационных станций (РЛС) со сложным сигналом от имитирующей помехи. Эффективность способа исследуется для случая воздействия типовой помехи и рэлеевских флуктуаций амплитуд полезных и помехового сигнала.

*Ключевые слова:* защита РЛС, помеха, рэлеевские флуктуации.

A statistical method of protecting the complex radar signal of simulating interference is proposed. The effectiveness of the method is studied for the case of the impact of the model noise and Raleigh fluctuations of amplitudes and useful jamming signal.

*Keywords:* RLS protection, jamming, Raleigh fluctuations.

### 1. Введение

Радиолокационные станции (РЛС), использующие технику сжатия сигналов с внутриимпульсной модуляцией, обладают повышенными скрытностью и помехозащищенностью. Вместе с тем, РЛС данного класса также как и обычные локаторы подвержены воздействию преднамеренных имитирующих помех, причем специфика обработки сигналов с внутриимпульсной модуляцией в ряде случаев упрощают противнику задачу формирования имитирующей помехи. Например, для создания двух и более ложных отметок в дальномерном канале в ретрансляционных станциях помех успешно используется внесение амплитудной модуляции в ретранслируемый зондирующий сигнал РЛС (рис. 1) [1, 2]. Указанные обстоятельства заставляют искать специальные способы защиты РЛС с сжатием импульсов от имитирующих помех.

### 2. Возможные подходы

Исследования возможных путей решения рассматриваемой проблемы проводились с учетом следующих ограничений.

Подавляемой стороне, как правило, не известны параметры имитирующей помехи: форма имитирующего сигнала, его статистические свойства, количество помеховых сигналов. Как правило, не установленными являются сам факт действия по-

мехи и количество реальных целей на анализируемом интервале дальности. Следовательно, синтез оптимального приемника, а также построение адаптивной системы обработки смеси полезных и мешающих сигналов не являются возможными. Поэтому принятие решения о принадлежности отметки в канале вторичной обработки РЛС к истинной или ложной цели можно осуществить только на основе оценок параметров принимаемых сигналов, в том числе и статистических параметров.

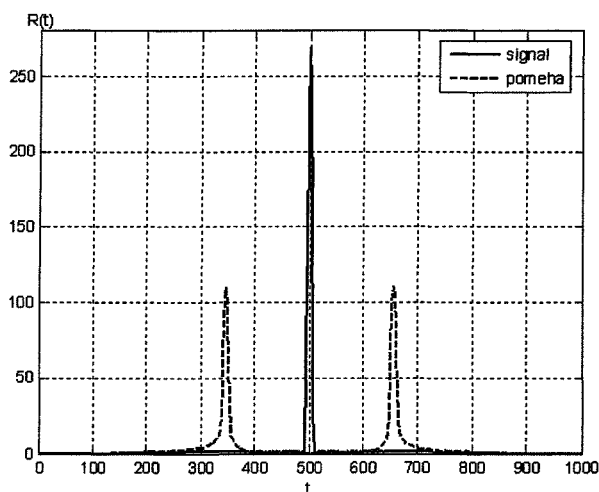


Рис. 1. Отклики СФ на сигналы от цели и помехи

Карманов Юрий Трофимович – д-р техн. наук, профессор, заведующий кафедрой цифровых радиотехнических систем ЮУрГУ; ea@drts.susu.as.ru

Непомнящий Глеб Александрович – старший преподаватель кафедры цифровых радиотехнических систем ЮУрГУ; ea@drts.susu.ac.ru

Karmanov Yuri Trofimovich – PhD, professor, head of Digital radio-technical systems department of SUSU; ea@drts.susu.as.ru

Nepomnyashiy Gleb Aleksandrovich – senior lecturer of Digital radio-technical systems department of SUSU; ea@drts.susu.ac.ru

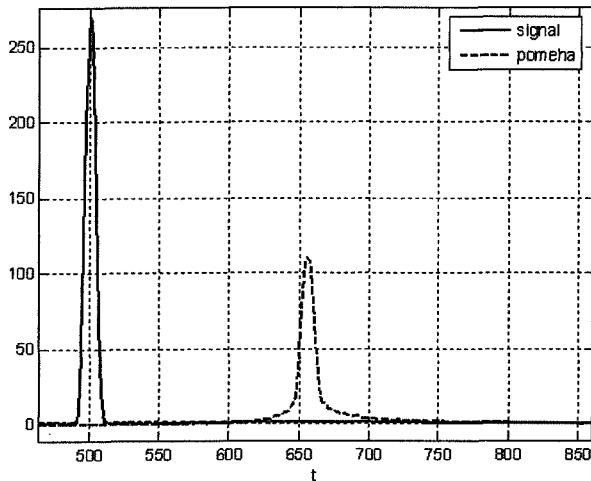


Рис. 2. Расширение длительности ГЛ отклика в случае ЛЧМ сигнала

В качестве оцениваемых нестатистических параметров теоретически можно использовать:

- 1) амплитуду отклика сжимающего фильтра (СФ) на принятый нормированный по энергии сигнал;
- 2) относительный уровень боковых лепестков сжатого сигнала;
- 3) длительность главного лепестка сжатого сигнала.

Уменьшение амплитуды отклика фильтра на помеховый сигнал, как и увеличение уровня боковых лепестков, связано с различием в форме между отраженным от реальной цели сигналом и имитирующим сигналом. Однако, в сложной сигнальной обстановке, когда на вход канала обработки поступает смесь нескольких сигналов и шума, нормирование энергий каждого сигнала по отдельности и измерение уровня боковых лепестков отдельных откликов невозможны.

Увеличение длительности главного лепестка (ГЛ) отклика на имитирующий сигнал помехи характерно для приемника сигналов с линейной частотной модуляцией (ЛЧМ). Исследования показали, что увеличение длительности главного лепестка составляет до 4 раз по уровню половины амплитуды и до 20 раз по уровню одной десятой амплитуды при приеме амплитудно-модулированной имитирующей помехи. Следовательно, для защиты РЛС с ЛЧМ сигналом можно использовать измерение длительности откликов и сравнение их с заданным порогом (рис. 2).

В качестве оцениваемых статистических параметров можно рассматривать параметры амплитудных распределений отметок в дальномерном канале, например, средние значения и дисперсии амплитуд откликов. Те отметки, для которых соотношение между средним и дисперсией не подходит под параметры возможных распределений амплитуд отметок реальных целей, следует считать ложными.

### 3. Предлагаемое решение

В ходе исследований был предложен следующий статистический способ определения принадлежности отметок к ложным целям.

Если уровень боковых лепестков отклика фильтра на сигнал помехи значителен и превышает порог обнаружения (как в случае амплитудно-модулированной помехи), то в ответ на один помеховый сигнал формируется сразу несколько ложных отметок. В таком случае флуктуации амплитуд ложных отметок должны быть достаточно сильно коррелированы на протяжении интервала наблюдения и для принятия решений можно использовать оценки коэффициентов корреляции между амплитудами отметок. На рис. 3 изображена

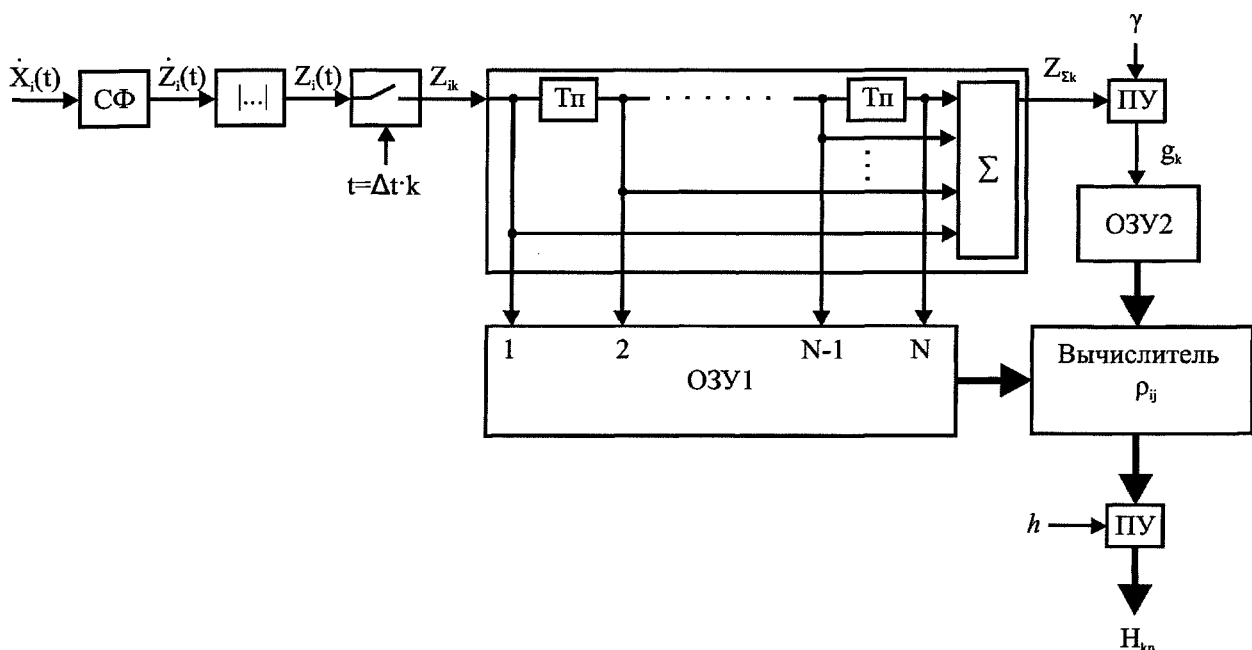


Рис. 3. Структурная схема приемника

структурная схема приемника, в которой реализуется предлагаемый способ.

Входной сигнал подвергается сжатию в СФ, затем формируется модуль отклика фильтра (огibaющая отклика), подвергаемый дискретизации с интервалом  $\Delta t$ . В накопителе происходит некогерентное накопление дискретизованных модулей отклика  $Z_{ik}$  для каждого излученного импульса. Число импульсов  $N, i=1 \dots N$ . Результат накопления  $Z_k^N$  сравнивается с порогом  $\gamma$ , значение которого определяет вероятность ложной тревоги. Все  $N$  векторов  $Z_{ik}$  записываются в ОЗУ1 и далее поступают на вычислитель коэффициента корреляции  $\rho$  между столбцами матрицы с элементами  $Z_{ik}$  для тех значений  $k$ , в которых порог  $\gamma$  превышен ( $g_k=1$ ). Вычисление производится в соответствии с выражением

$$\rho_{nm} = \sum_{j=1}^N \left[ \frac{Z_{jn} - \frac{1}{N} \sum_{i=1}^N Z_{in}}{\left\{ \sum_{l=1}^N \left[ Z_{ln} - \frac{1}{N} \sum_{i=1}^N Z_{il} \right]^2 \right\}^{1/2}} \times \frac{Z_{jm} - \frac{1}{N} \sum_{i=1}^N Z_{im}}{\left\{ \sum_{l=1}^N \left[ Z_{lm} - \frac{1}{N} \sum_{i=1}^N Z_{il} \right]^2 \right\}^{1/2}} \right], \quad n, m \in 1 \dots K, \quad (1)$$

где  $K$  – число анализируемых отсчетов дальности.

Значения коэффициентов  $\rho_{nm}$  сравниваются с заданным порогом  $h$ . Если  $\rho_{nm} \geq h$  то принимается решение, что отметки на дальностях с номерами  $n$  и  $m$  являются ложными (для номеров  $n$  и  $m$  принимается гипотеза  $H_2$ ). Все значения  $k$ , удовлетворяющие условию  $\rho_{km} < h, m \in 1 \dots K, g(m) = 1$ , считаются номерами отсчетов дальности, на которых присутствуют истинные цели (для номеров  $k$  принимается гипотеза  $H_1$ ). Для нахождения вероятностей ошибок  $P_{12}$  (вероятность принятия ложной отметки за истинную) и  $P_{21}$  (вероятность принятия истинной отметки за ложную) необходимо знать плотности распределения вероятности оценок  $\rho_{nm}$  по обеим гипотезам  $W_\rho(\rho|H_1)$  и  $W_\rho(\rho|H_2)$ , которые зависят от параметров отраженных и помеховых сигналов, уровня шума и числа импульсов  $N$ .

#### 4. Результаты исследований

Найдем выражения для фактических значений коэффициентов корреляции между амплитудами отметок для двух случаев:

1) двух независимо флюктуирующих отметок (гипотеза  $H_1$ );

2) одинаково флюктуирующих отметок на фоне шума (гипотеза  $H_2$ ).

Комплексный сигнал на входе приемника в случае двух независимо флюктуирующих целей [3]

$$\dot{X}_i(t) = \dot{A}_1 \dot{S}(t - \tau_1) + \dot{A}_2 \dot{S}(t - \tau_2) + \dot{n}(t), \quad (2)$$

где  $\dot{A}_1$  и  $\dot{A}_2$  – независимые комплексные нормальные величины с нулевыми средними и дисперсиями  $D_1$  и  $D_2$  соответственно;  $\dot{S}(t)$  – комплексный зондирующий сигнал;  $\tau_1$  и  $\tau_2$  – задержки отраженных сигналов;  $\dot{n}(t)$  – комплексный белый шум.

Огибающая отклика СФ на сигнал  $\dot{X}_i(t)$

$$Z_i(t) = \left| \dot{A}_1 \dot{R}(t - \tau_1) + \dot{A}_2 \dot{R}(t - \tau_2) + \dot{v}(t) \right|, \quad (3)$$

где  $\dot{R}(t)$  – отклик СФ на  $\dot{S}(t)$ ;  $\dot{v}(t)$  – комплексный гауссовский шум на выходе СФ с дисперсией  $D_{ш}$ .

Дискретные отсчеты огибающей отклика

$$Z_{ik} = \left| \dot{A}_1 \dot{R}_{k-n} + \dot{A}_2 \dot{R}_{k-m} + \dot{v}_{ik} \right|, \quad (4)$$

где  $k = t/\Delta t, n = \tau_1/\Delta t, m = \tau_2/\Delta t, \Delta t$  – интервал дискретизации, определяемый разрешающей способностью РЛС по дальности.

Таким образом, величина  $Z_{ik}$  – модуль вектора с независимыми нормальными компонентами, имеющими нулевые средние и дисперсии

$$D_k = D_1 R_{k-n}^2 + D_2 R_{k-m}^2 + D_{ш}, \quad (5)$$

где  $R_k$  – отсчеты модуля (огibaющей) отклика СФ.

Коэффициент корреляции между амплитудами отсчетов дальности с номерами  $n$  и  $m$

$$\rho_{nm}^* = \frac{M\{Z_{in} \cdot Z_{im}\} - M\{Z_{in}\} \cdot M\{Z_{im}\}}{[D\{Z_{in}\} \cdot D\{Z_{im}\}]^{1/2}}, \quad (6)$$

где  $M\{*\}$  и  $D\{*\}$  – среднее и дисперсия случайной величины.

Нахождение среднего  $M\{Z_{in} \cdot Z_{im}\}$  затруднительно. В ходе исследований было замечено, что коэффициенты корреляции модулей амплитуд отметок мало отличаются от коэффициентов корреляции квадратов модулей  $\rho_{nm}^{2*}$  (рис. 4).

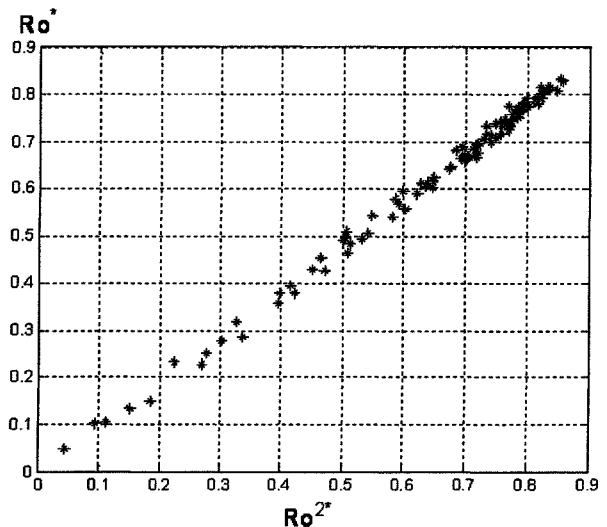


Рис. 4. Зависимость коэффициента корреляции  $\rho_{nm}^*$  от коэффициента корреляции  $\rho_{nm}^{2*}$

Таким образом

$$\rho_{nm}^* \approx \rho_{nm}^{2*} = \frac{M\{Z_{in}^2 \cdot Z_{im}^2\} - M\{Z_{in}^2\} \cdot M\{Z_{im}^2\}}{[D\{Z_{in}^2\} \cdot D\{Z_{im}^2\}]^{1/2}}, \quad (7)$$

Для нахождения (7) используем [4]. Основная трудность здесь связана с вычислением  $M\{Z_{in}^2 \cdot Z_{im}^2\}$ . Для упрощения процедуры вычисления  $M\{Z_{in}^2 \cdot Z_{im}^2\}$  будем считать отсчеты шума  $\dot{v}_{in}$  и  $\dot{v}_{im}$  статистически независимыми. Опуская громоздкие выкладки, запишем окончательный результат

$$\rho_{nm}^* \approx \frac{R_0^2 R_{m-n}^2 (k_1^2 + k_2^2)}{(k_1 R_0^2 + k_2 R_{m-n}^2 + 1)(k_2 R_0^2 + k_1 R_{m-n}^2 + 1)} + \frac{2k_1 k_2}{(k_1 R_0^2 + k_2 R_{m-n}^2 + 1)(k_2 R_0^2 + k_1 R_{m-n}^2 + 1)} \times [\text{Re}^2\{\dot{R}_0\} - \text{Im}^2\{\dot{R}_0\}][\text{Re}^2\{\dot{R}_{m-n}\} - \text{Im}^2\{\dot{R}_{m-n}\}] + 4 \text{Re}\{\dot{R}_0\} \cdot \text{Im}\{\dot{R}_0\} \cdot \text{Re}\{\dot{R}_{m-n}\} \cdot \text{Im}\{\dot{R}_{m-n}\}], \quad (8)$$

где  $k_1 = \frac{D_1}{D_{ш}}$ ,  $k_2 = \frac{D_2}{D_{ш}}$ .

В случае использования в РЛС ФКМ сигналов мнимые составляющие откликов отсутствуют и выражение (8) можно переписать следующим образом:

$$\rho_{nm}^* | H_1 \approx \frac{(q_1 + q_2)^2}{\Delta R^2 \left( q_1 + \frac{q_2}{\Delta R^2} + 1 \right) \left( q_2 + \frac{q_1}{\Delta R^2} + 1 \right)}, \quad (9)$$

где  $q_1 = \frac{D_1 R_0^2}{D_{ш}}$ ,  $q_2 = \frac{D_2 R_0^2}{D_{ш}}$  – отношения сигнал/шум

на выходе СФ для отраженных сигналов,  $\Delta R = \frac{R_0}{R_{m-n}}$  – подавление боковых лепестков отклика СФ на отраженный сигнал.

Из (9) следует, что в случае воздействия на приемник одного сигнала помехи

$$\rho_{nm}^* | H_2 \approx \frac{q_n^2}{(q_n + 1) \cdot (q_n + \Delta R_n^2)}, \quad (10)$$

где  $q_n$  – отношение сигнал/шум на выходе СФ для имитирующего сигнала помехи;  $\Delta R_n$  – подавление боковых лепестков отклика СФ на сигнал помехи.

Выражения (9) и (10) позволяют приближенно определить средние значения оценок  $\rho_{nm} | H_1$  и  $\rho_{nm} | H_2$ , определяемых в (1). Нахождение других параметров распределений  $W_\rho(\rho|H_1)$  и  $W_\rho(\rho|H_2)$  в силу сложности (1) является трудновыполнимой задачей. На рис. 5 изображены результаты статистического расчета плотности  $W_\rho(\rho|H_2)$  для разных значений  $N$ , показывающие, что распределение  $W_\rho(\rho|H_2)$  может существенно отличаться от нормального.

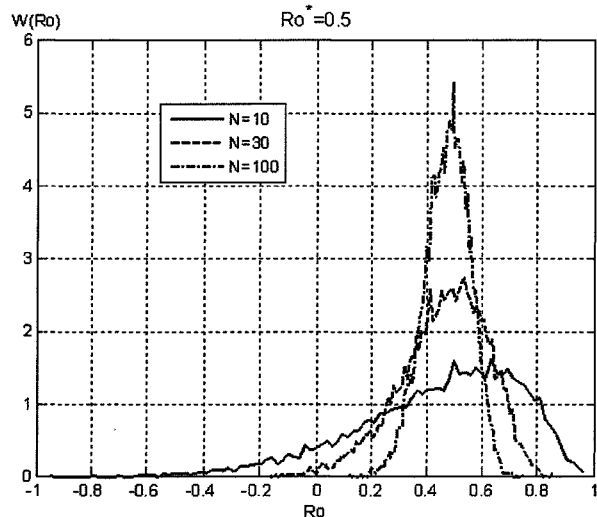


Рис. 5. Распределения  $W_\rho(\rho|H_2)$  при  $\rho_{nm}^*/H_2$  и различных  $N$

По аналогии с [5] введем монотонное преобразование

$$r_{nm} = \frac{1}{2} \ln \left( \frac{1 + \rho_{nm}}{1 - \rho_{nm}} \right). \quad (11)$$

На рис. 6 пунктирными линиями изображены кривые распределений величины  $r_{nm}$ , соответствующих распределениям на рис. 5 и полученных с помощью статистического моделирования. Сплошными линиями на рис. 6 изображены кривые нормального распределения с параметрами:

$$\text{среднее} \quad M_r = \frac{1}{2} \ln \left( \frac{1 + \rho_{nm}^*}{1 - \rho_{nm}^*} \right), \quad (12)$$

$$\text{дисперсия} \quad D_r = \frac{1}{N - 3}. \quad (13)$$

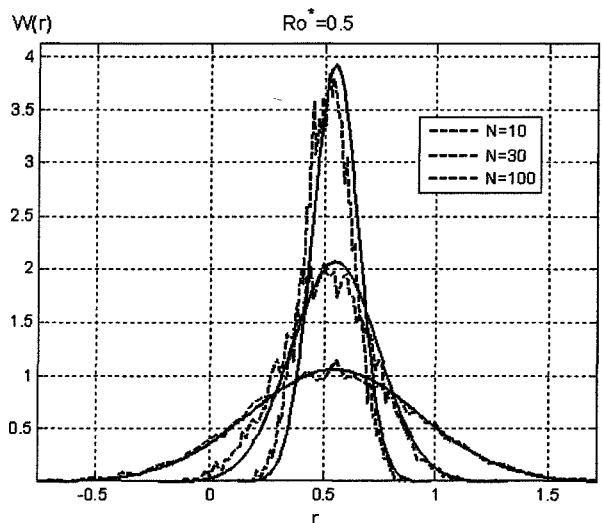


Рис. 6. Аппроксимация распределения  $W_r(r|H_2)$  нормальным законом

Распределения  $W_r(r|H_2)$  мало отличаются от нормальных с указанными параметрами. Результаты моделирования показали, что данная аппрок-

симация справедлива для произвольных значений  $\rho_{nm}^* | H_1$  и  $\rho_{nm}^* | H_2$  и  $N \geq 0$ . С учетом преобразования (11) можно приближенно определить вероятности ошибочных решений

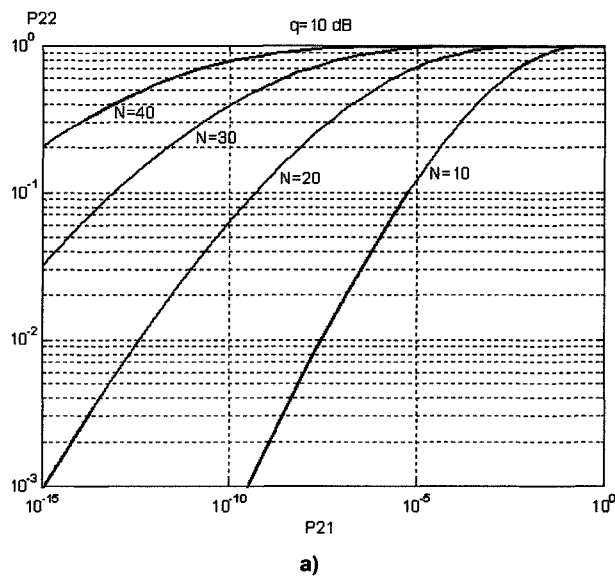
$$P_{12} = P\{\rho_{nm} | H_2 < h\} \approx \frac{1}{2} + \frac{1}{2} \operatorname{erf} \left\{ \frac{\ln \left[ \frac{(1+h)(1-\rho_{nm}^* | H_2)}{(1-h)(1+\rho_{nm}^* | H_2)} \right]}{2\sqrt{\frac{2}{N-3}}} \right\}; \quad (14)$$

$$P_{21} = P\{\rho_{nm} | H_1 > h\} \approx \frac{1}{2} - \frac{1}{2} \operatorname{erf} \left\{ \frac{\ln \left[ \frac{(1+h)(1-\rho_{nm}^* | H_1)}{(1-h)(1+\rho_{nm}^* | H_1)} \right]}{2\sqrt{\frac{2}{N-3}}} \right\}. \quad (15)$$

где  $\rho_{nm}^* | H_1$  и  $\rho_{nm}^* | H_2$  определяются (9) и (10) соответственно.

Соотношения (14) и (15) позволяют построить характеристики распознавания истинных и ложных отметок (зависимость вероятности распознавания ложных отметок  $P_{22} = 1 - P_{12}$  от вероятности ошибки  $P_{21}$ ) при различных параметрах отраженных и помеховых сигналов.

Пусть отношения сигнал / шум для отраженных и помехового сигналов одинаковы и равны  $q$ , подавление боковых лепестков отклика СФ на отраженный сигнал составляет 30 дБ, а на сигнал помехи 0 дБ (рис. 7). Характеристики распознавания, построенные для различных значений отно-



шения  $q$  и числа наблюдаемых (излучаемых РЛС) импульсов  $N$  изображены на рис. 8, а, б. Как видно, при сравнительно небольших значениях  $q$  и  $N$  достоверность принятия решения в рассматриваемом приемнике весьма высока.

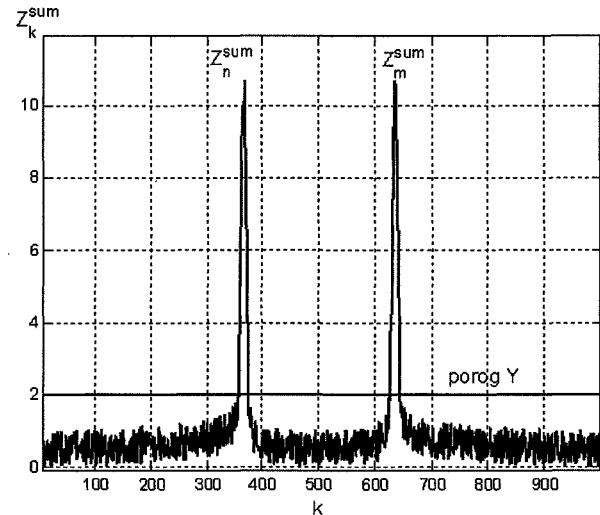


Рис. 7. Наблюдение двух отметок одинаковой интенсивности на выходе накопителя

Результат наблюдения для случая, когда средние амплитуды отметок, формируемых отраженными или помеховым сигналами, различаются (рис. 9, а). Для построения характеристик распознавания в этой ситуации следует задать отношение сигнал шум отметки меньшей амплитуды  $q_2$  и уровень подавления боковых лепестков отклика на сигнал помехи  $\Delta R_{\Pi}$ . При этом  $q_1 \approx q_2 \cdot \Delta R_{\Pi}^2$ . Характеристики распознавания, построенные для  $\Delta R = 30$  дБ,  $q_2 = 10$  дБ,  $N = 30$  и различных значений  $\Delta R_{\Pi}$  изображены на рис. 9, б. Достоверность принятия решений высока и в этом случае.

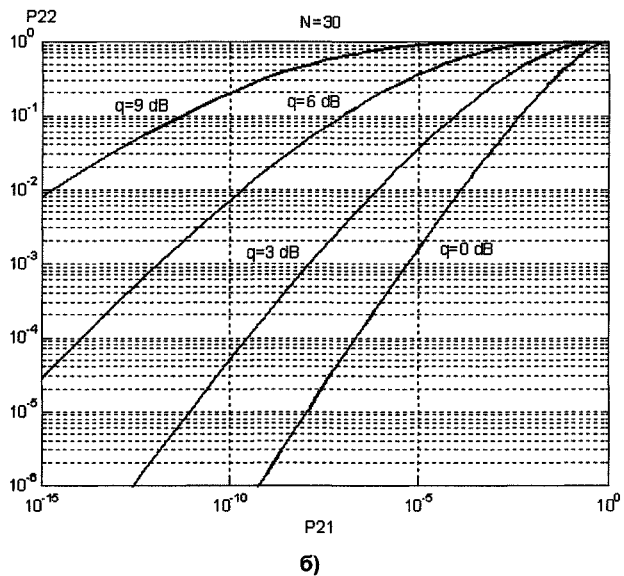


Рис. 8. Характеристики распознавания: а – при различных  $N$ ; б – при различных  $q$

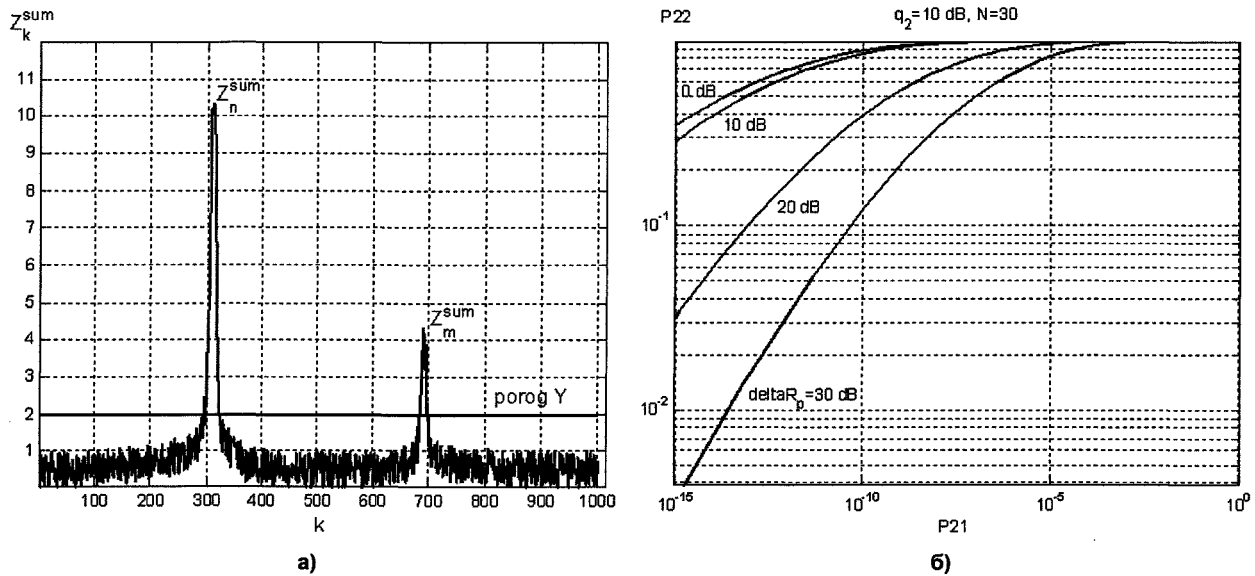


Рис. 9. Наблюдение двух отметок разной интенсивности (а) и соответствующие характеристики распознавания (б)

### 5. Заключение

В ходе исследований способов защиты РЛС сжатием импульсов от имитирующей помехи предложен специальный способ, обладающий высокой эффективностью при воздействии на приемник типовой (амплитудно-модулированной) имитирующей помехи, что подтверждается результатами теоретических расчетов. Данный способ можно применить и при воздействии другого типа помехи, отклик СФ на которую имеет значительный уровень боковых лепестков.

Следует отметить, что предлагаемый способ будет эффективен только в том случае, если флуктуации амплитуд сигналов истинных целей слабо коррелированы в течение одного периода излучения зондирующего импульса РЛС (быстрые флуктуации ЭПР). В случае медленных амплитудных флуктуаций, когда период повторения импульсов невелик, в качестве оцениваемого параметра можно использовать коэффициент корреляции фазы откликов СФ. Такой способ нуждается в проведении отдельного исследования на эффективность.

### Литература

1. Свистов, В.М. Радиолокационные сигналы и их обработка / В.М. Свистов. – М.: Сов. радио, 1977.
2. Перунов, Ю.М. Радиоэлектронное подавление информационных каналов систем управления оружием / Ю.М. Перунов, К.И. Фомичев, Л.М. Юдин; под ред. Ю.М. Перунова. – М.: Радиотехника, 2003.
3. Ван Трис, Г. Теория обнаружения, оценок и модуляции / Г. Ван Трис. – М.: Сов. радио, 1972. – Т. 3.
4. Левин, Б.Р. Теоретические основы статистической радиотехники / Б.Р. Левин. – М.: Радио и связь, 1989.
5. Кендалл, М. Статистические выводы и связи / М. Кендалл, А. Стюарт. – Главная редакция физико-математической литературы издательства «Наука», 1973.

Поступила в редакцию 16 марта 2009 г.

# ПРОТОКОЛ МАРШРУТИЗАЦИИ ДЛЯ AD-HOC СЕТЕЙ

*М.Л. Карманов*

## ROUTING IN AD-HOC NETWORKS

*M.L. Karmanov*

В статье описывается метод существенного повышения защищенности протоколов маршрутизации ad-hoc сетей от внешних воздействий. Ad-hoc сети – это относительно новое направление, предполагающее построение самоорганизующихся вычислительных сетей на базе беспроводных технологий. Изложенный в статье метод позволяет существенным образом повысить устойчивость алгоритмов маршрутизации от внешних воздействий, при этом, практически не увеличивая вычислительную сложность и объем служебного трафика.

*Ключевые слова: вычислительные сети, алгоритмы маршрутизации, ad-hoc сети.*

In given article the method of essential increase of security of reports of routeing ad-hoc of networks from external influences is described. Ad-hoc networks is rather new on-board, assuming construction self-organized computer networks on the basis of wireless technologies. The method stated in article allows to raise essentially stability of algorithms of routeing from external influences, thus, practically without increasing computing complexity and volume of the traffic.

*Keywords: network, routing, ad-hoc networks.*

### Введение

В 1990-х годах широкое распространение получили мобильные компьютеры, оснащенные беспроводными интерфейсами. Именно тогда и зародилась идея создания так называемых ad-hoc сетей. Ad-hoc сеть подразумевает под собой объединение нескольких мобильных устройств с беспроводными сетевыми адаптерами в единую сеть передачи данных. При этом предполагается, что мобильные устройства принадлежат различным людям и единое централизованное административное управление отсутствует, кроме того, такие сети должны быть самоорганизующимися (самонастраивающимися). Также, необходимо учитывать мобильность узлов сети и их ограниченность как в вычислительных, так и в энергоресурсах.

Построение ad-hoc сетей порождает множество самых разнообразных проблем. В данной статье рассматривается только вопрос маршрутизации.

В ходе выполненного анализа существующих протоколов маршрутизации, используемых в таких сетях, выяснилось, что все они обладают слабой устойчивостью к ситуациям, при которых один или несколько узлов сети отклоняются от требований протокола. В ряде случаев такое поведение всего лишь одного узла может парализовать работу всей сети. Кроме того, большинство протоколов маршрутизации было так или иначе унаследовано

из проводных сетей и слабо адаптировано к особенностям беспроводных ad-hoc сетей.

Был проведен детальный анализ причин неустойчивости протокола к отклонениям отдельных узлов от выполнения протокола. Эти причины были выделены и формализованы. Затем один из существующих протоколов динамической маршрутизации был существенным образом модернизирован. Было проведено моделирование работы нового протокола и проведен сравнительный анализ с другими протоколами.

В результате исследований было установлено, что новый протокол обладает свойством локализации воздействий, при этом вычислительные и энергозатраты возрастают незначительно, кроме того, новый протокол оказался более производительным по ряду параметров.

### 1. Обзор существующих протоколов маршрутизации

Был проведен анализ следующих протоколов динамической маршрутизации: OSPF, RIP, EBFRRP, DSDV, DSR, AODV и ряда других. Для проведения анализа была разработана система критериев. Рассматривались следующие критерии: количество узлов, занимающихся маршрутизацией; уровень знания каждым узлом топологии всей сети; наличие у каждого узла маршрута до любого другого

узла; возможность использования различных метрик; возможность возникновения «петель»; знание резервных маршрутов до узла; время реакции на изменение топологии сети; объем данных передаваемых по сети для построения маршрута; количество ресурсов, необходимых узлу для построения таблицы маршрутизации; поддержка multicast (групповой рассылки); устойчивость протокола.

В результате проведенных исследований было выяснено, что все протоколы маршрутизации, использующиеся на сегодняшний день в ad-hoc сетях, являются не устойчивыми к внешним и внутренним воздействиям. Во многих протоколах отклонение нескольких узлов от протокола может полностью парализовать работу сети.

Например, в описании одного из самых популярных протоколов AODV, прямо указано, что разработчики понимают, что их протокол является неустойчивым к внешним и внутренним воздействиям, но менять эту ситуацию пока не собираются [1].

На основании результатов проведенного исследования был выявлен наиболее подходящий протокол для ad-hoc сетей – Ad-hoc On demand Distance-Vector (AODV). Также в пользу этого выбора говорит то, что данный протокол получил отражение в RFC, и то, что многие последние доклады на различных конференциях связаны именно с этим протоколом. В том числе, проведены многочисленные исследования, в которых сравнивалась производительность различных протоколов в смоделированных сетях. По итогам этих исследований протокол AODV по многим параметрам оказался лучше остальных [2, 3, 4]. Кроме того, данный протокол используется во многих программных реализациях для построения ad-hoc сетей [5, 6].

## 2. Описание протокола AODV

Вкратце, работа данного протокола выглядит следующим образом. Если некоторому узлу, назовем его инициатором, требуется передать данные другому узлу, назовем его узлом назначения, а соответствующий маршрут не известен, то узел инициатор посылает широковещательный запрос RREQ для поиска данного маршрута. Этот запрос передается всем соседям данного узла, которые, в свою очередь, пересылают его своим соседям и так далее. Так как в запросе указана информация об узле инициаторе, то все узлы, получившие данный запрос строят маршруты до узла – инициатора. Через некоторое время данный запрос достигает либо узла назначения, либо некоторого промежуточного узла, которому известен необходимый маршрут. При этом узлу инициатору отправляется ответ RREP, который передается уже не широковещательно, а адресно, по маршруту, построенному при распространении запроса. В результате такой операции выстраивается двунаправленный маршрут между заинтересованными узлами. Прав-

да, в случае ответа на запрос промежуточным узлом он обязан отправить ещё один «беспричинный» ответ уже не инициатору, а узлу назначения. Этот ответ необходим для построения обратного маршрута от узла назначения до инициатора.

Заметим, что информация, находящаяся у различных узлов сети, может иметь разную степень давности, при этом понятно, что для построения маршрутов лучше использовать более новую информацию. Для того чтобы можно было определять степень свежести информации, в данном протоколе введены порядковые номера узлов. Эта идея была взята из протокола DSDV, один из авторов которого С.Е. Perkins является также и автором протокола AODV. Суть идеи состоит в следующем. Каждый узел имеет собственный порядковый номер, значением которого он управляет, кроме того, этот номер прикрепляется ко всей информации о маршрутах до данного узла, в том числе хранится и в таблицах маршрутизации. При изменении связей с соседними узлами, узел увеличивает свой порядковый номер. Благодаря этому, появляется возможность выбора более новой информации о маршрутах до данного узла.

При обнаружении нарушения маршрута узел, обнаруживший это нарушение, рассылает информацию об ошибке посредством специального сообщения RERR. При этом маршрут в большинстве случаев заново перестраивается.

На рис. 1 изображена отправка запроса RREQ (пунктирные стрелки) узлом *A*, который желает узнать маршрут до узла *B*. При этом один из промежуточных узлов – узел 5 знает необходимый маршрут и отправляет узлу *A* ответ RREP (сплошные стрелки), он же отправляет ещё один «беспричинный» ответ (сплошные стрелки) узлу *B*, для того, чтобы построился обратный маршрут от узла *B* к узлу *A*.

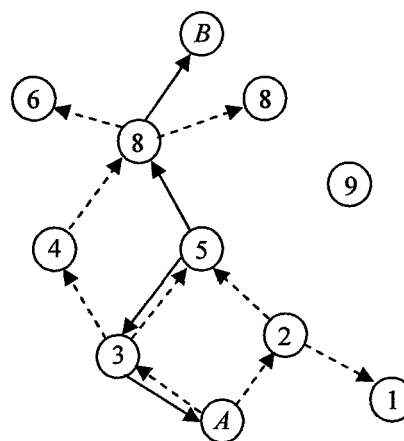


Рис. 1. Установление маршрута от *A* к *B*

При построении маршрутов в ad-hoc сетях, как правило, возможны различные варианты и узел *A* получит несколько ответов с различными маршрутами. При этом в первую очередь выбираются маршруты с большим порядковым номером



узла  $B$  (как маршруты, построенные на основании более новой информации), а среди маршрутов с одинаковым порядковым номером выбирается более короткий. Узел  $A$  не ожидает всех ответов, а начинает передачу данных после прихода первого ответа. Получив дополнительные ответы, он может изменить маршрут и следующие пакеты с данными передавать уже по более новому или короткому маршруту.

Обратим внимание на ряд особенностей данного протокола. Узел инициатор всегда выбирает более новый маршрут (с большим порядковым номером), даже если тот является существенно длиннее. Узел назначения при ответе на запрос может увеличивать свой порядковый номер только на единицу. Протокол AODV работает на прикладном уровне, используя в качестве транспортного протокола UDP. Получение узлом ответа RREP без отправки соответствующего запроса является нормальной ситуацией и получивший такой ответ узел должен его обработать.

### 3. Анализ устойчивости протокола AODV

Теперь перейдем к рассмотрению устойчивости данного протокола. Возможны два вида воздействий на протоколы маршрутизации. Первый вид воздействий может быть отнесен к классу *man-in-the-middle*. Своей целью данные воздействия ставят прокладку маршрута между двумя узлами через определенный третий узел, с целью прослушивания трафика между интересующими узлами. Второй вид воздействий принадлежит классу *deny-of-service* и ставит своей целью нарушение работы сети, путем создания неверных маршрутов, чрезмерного увеличения нагрузки на сеть или блокирования возможности создания верных маршрутов.

После тщательного изучения спецификации протокола AODV были разработаны способы реализации воздействий обоих видов. Важно отметить, что эти воздействия являются возможным не из-за ошибок в конкретных реализациях протокола, а из-за особенностей самого протокола маршрутизации.

Воздействия первого вида реализуются одним из двух способов. Узел  $Z$  отправляет узлам жертвам  $A$  и  $B$  «беспричинные» ответы, содержащие либо заниженное расстояние до второго узла, либо завышенный порядковый номер. При этом маршрут от узла  $A$  к узлу  $B$  выстраивается через узел  $Z$  и наоборот. Более выигрышной является стратегия завышения порядкового номера узла, так как при этом созданные маршруты обладают большей устойчивостью.

Поясним это с помощью рис. 2. На нем сплошными стрелками изображен «беспричинный» ответ RREP узлу  $A$ , содержащий завышенный порядковый номер узла  $B$ . То есть узлы 2 и  $A$  будут иметь неверную завышенную информацию о порядковом номере узла  $B$ . Также, сплошными стрелками обо-

значен «беспричинный» ответ RREP узлу  $B$ . При этом узлы 9, 8 и  $B$  будут иметь неверную завышенную информацию о порядковом номере узла  $A$ . Так как сами узлы  $A$  и  $B$  могут увеличивать свои порядковые номера не более чем на единицу при каждом запросе, то в течение некоторого времени вся информация, приходящая узлам 2 и  $A$ , относительно маршрута до узла  $B$  будет иметь порядковый номер меньше, чем записан у этих узлов и, соответственно, будет игнорироваться. Аналогично с маршрутом от  $X$  до  $B$ . Заметим, что при реализации данного воздействия нарушается одно из двух соотношений, которые выполняются для произвольного маршрута ( $A \rightarrow \dots \rightarrow N \rightarrow \dots \rightarrow B$ ) при нормальном функционировании протокола AODV:

$$D_B(A) > \dots > D_B(N) > \dots > D_B(B),$$

$$S_B(A) \leq \dots \leq S_B(N) \leq \dots \leq S_B(B),$$

где  $D_B(N)$  – это расстояние до узла  $B$ , известное узлу  $N$ , а  $S_B(N)$  – это порядковый номер узла  $N$ , известный узлу  $N$ .

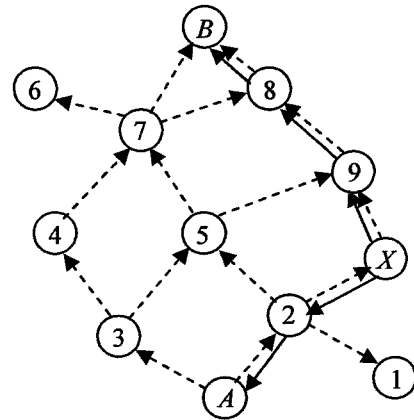


Рис. 2. Узел  $B$  отправляет ответы RREP<sub>A</sub> и RREP<sub>B</sub>

Если модернизировать протокол таким образом, что бы он мог контролировать выполнение данных соотношений, то такие воздействия станут невозможными.

Воздействия типа *deny-of-service* могут быть реализованы двумя способами. Первый способ состоит в регулярной рассылке «беспричинных» ответов RREP различным узлам. При этом все маршруты будут выстраиваться к одному узлу. Второй способ состоит в отправке сообщений об ошибке RERR различным узлам от имени других узлов сети.

Нужно отметить, что с 2000 года на различных конференциях стали появляться работы, связанные с безопасностью и устойчивостью протоколов маршрутизации в ad-hoc сетях [7, 8]. В них предлагаются различные модификации, призванные увеличить защищенность протоколов. Однако все предложенные модернизации опираются на криптографию с асимметричными ключами и предполагают наличие некоторого удостоверяющего центра, либо регулярную генерацию сеансо-

вых ключей. При этом, как признаются сами авторы модификаций, возможности для воздействий все равно остаются. Кроме того, применение криптографических средств, особенно асимметричной криптографии, существенно увеличивает требования к вычислительным мощностям узлов, а также увеличивает энергопотребление.

В данной статье рассматривается другой подход – модернизировать протокол так, чтобы воздействия на него если и были возможны, то носили локальный характер. Локальный, либо по времени воздействия, либо по области воздействия.

#### 4. Модернизированный протокол AODV

Для решения проблем устойчивости протокола AODV в него был внесен ряд изменений. В данной статье приводятся только наиболее существенные из них.

1. Все служебные пакеты инкапсулируются не в протокол UDP, а в протокол канального уровня. Это позволяет локализовать область, в которой могут подделываться адреса отправителей в сообщениях.

2. Введен механизм периодического обмена таблицами маршрутизации между соседними узлами. Тем самым к знаниям узла о топологии сети добавляются также знания его соседей. Это позволяет в большинстве случаев выявлять нарушение равенств (1) и (2) и блокировать связь с узлом-нарушителем на длительное время.

3. Благодаря периодическому обмену таблицами маршрутизации между соседними узлами, был также введен механизм быстрого восстановления маршрутов в случае нарушения одной из промежуточных связей. При этом многократно сократилось время на восстановление нарушенного маршрута, а также уменьшился объем трафика, передаваемого по сети для восстановления нарушенного маршрута.

4. Изменена метрика сети. Под длиной связи понимается остаточная пропускная способность канала. Это позволило ввести в протокол AODV функции балансировки нагрузки. И избавиться от проблемы протокола AODV, связанной с повышенной нагрузкой на каналы передачи данных, находящиеся в центральной части вычислительной сети.

5. Добавлена возможность включения блока данных в пакет запроса построения маршрута, что позволило в случае использования протокола TCP передавать в месте с запросом на построение маршрута и запрос на установление соединения, что примерно на 35 % снизило задержку при установлении соединения с новым узлом, маршрут до которого был ранее не известен.

Также было внесено множество других изменений, некоторые из которых не влияли на защищенность протокола, но увеличивали его эффективность. Общее количество внесенных изменений равно шестнадцати.

#### 5. Основные результаты и достижения

Была проведена классификация большинства протоколов динамической маршрутизации. Протоколы были проанализированы на предмет их устойчивости к внутренним и внешним воздействиям.

На основании проведенного анализа был модернизирован протокол AODV, что позволило сделать его существенно более устойчивым к внутренним и внешним воздействиям, а также оптимизировать его работу.

Было проведено изучение модернизированного протокола в смоделированной вычислительной сети. Наиболее показательными являются следующие результаты, полученные в сравнении со стандартным протоколом AODV.

1. Объем служебного трафика увеличился всего на 13 %.

2. Среднее время построения маршрута осталось неизменным.

3. Среднее время восстановления маршрута сократилось на 81 %.

4. Количество операций полного восстановления маршрута (иницированных одним из конечных узлов) сократилось на 98 %.

5. Средняя задержка в установлении TCP-соединения с новым узлом сократилась на 36 %.

6. Воздействие типа Deny-of-service на вычислительную сеть локализовано зоной покрытия приемопередающего устройства, осуществляющего воздействие.

7. Воздействие типа man-in-the-middle локализовано зоной диаметром в 1–2 узла от источника воздействия.

#### Заключение

В ходе проделанной работы был существенно модернизирован протокол динамической маршрутизации для ad-hoc сетей.

Было проведено моделирование работы протокола, результаты которого показали более высокую эффективность модернизированного протокола по сравнению с базовым. Также моделирование показало более высокую устойчивость протокола к внешним воздействиям и локализацию последствий воздействия относительно небольшой областью.

В настоящее время ведутся работы по реализации данного протокола и его полевым испытаниям.

#### Литература

1. Perkins, C.E. *Ad hoc On-Demand Distance Vector (AODV) Routing* / C.E. Perkins, C.E. Belding-Royer // RFC 3561. – July 2003.

2. Shin, K.G. *Performance Analysis of Distributed Routing Strategies Free of Ping-Pong-Type Looping* / K.G. Shin, M. Chen // *IEEE Trans. Computers*. – February 1987. – V. COMP-36, № 2. – P. 129–137.

3. Broch, J. *A performance comparison of multi-hop wireless ad hoc network routing protocols* /

J. Broch, D.A. Maltz, D.B. Johnson, Y.-C. Hu, J. Jetcheva // *Proc. of MOBICOM'98*. – 1998.

4. Basagni, S. *Mobility- Adaptive Protocols for Managing Large Ad Hoc Network* / S. Basagni, D. Turgut, S.K. Das // *Proceedings of the IEEE International Conference on Communications (ICC)*. – 2001. – P. 1539–1543.

5. Kawadia, V. *System services for implementing Ad-Hoc routing: Architecture, Implementation and Experiences* / V. Kawadia, Y. Zhang, B. Gupta // *Proceedings of the 1st International Conference on Mobile Systems, Applications and Services (MobiSys)*. San Francisco, CA. – June 2003. – P. 99–112.

6. Kernel AODV from National Institute of Standards and Technology (NIST). Режим доступа: [http://w3.antd.nist.gov/wctg/aodv\\_kernel/](http://w3.antd.nist.gov/wctg/aodv_kernel/), свободный.

7. Zhou, L. *Securing Ad Hoc Networks* / L. Zhou, Z.J. Haas // *IEEE Networks Special Issue on Network Security*. – November/December 2000.

8. Brinkley, J. *Authenticated Ad Hoc Routing at the Link Layer for Mobile Systems* / J. Brinkley, W. Trost // *Wireless Networks* 7. – 2001. – P. 139–145.

9. Карманов, М.Л. *Mesh networks: проблемы безопасности [Текст]* / М.Л. Карманов // *Снежинск и наука – 2006: сб. науч. тр. междунар. науч.-практ. конф.* – Снежинск: СГФТА, 2006. – С. 158–159.

10. Карманов, М.Л. *Системы дистанционно-*

*го обучения: вопросы безопасности [Текст]* / П.В. Збицкий, М.Л. Карманов // *Математика. Механика. Информатика: сб. тез. Всерос. науч. конф.* – Челябинск: ЧелГУ, 2006. – С. 58.

11. Карманов, М.Л. *Маршрутизация в ad-hoc сетях: вопросы безопасности [Текст]* / М.Л. Карманов // *Безопасность информационного пространства: материалы междунар. науч.-практ. конф.* – Екатеринбург: УрГУПС, 2006. – С. 62–66.

12. Карманов, М.Л. *Протоколы маршрутизации для ad-hoc сетей [Текст]* / М.Л. Карманов // *Безопасность информационного пространства: материалы межвузовской науч.-практ. конф.* – Тюмень: ТюмГУ, 2007. – С. 19–26.

13. Карманов, М.Л. «Защита беспроводных сетей» [Электронный ресурс] / М.Л. Карманов // *Цифровые радиоэлектронные системы*. – Режим доступа: <http://www.drts.susu.ac.ru/niires/>. Зарегистрирован в Комитете РФ по печати № 0146611 от 20.03.1996. Твердая копия № 7, 2007–2008. – Челябинск: ЮУрГУ, 2008. – С. 108–115.

14. Карманов, М.Л. *Самоорганизующиеся беспроводные сети: алгоритмы маршрутизации [текст]* / М.Л. Карманов // *Наука ЮУрГУ: материалы 60-й юбилейной науч. конф. Секция технических наук*. – Челябинск: Изд-во ЮУрГУ, 2008. – Т. 2. – С. 64–68.

*Поступила в редакцию 10 апреля 2009 г.*

## ТЕРМИНАЛЬНАЯ СИСТЕМА WTPRO: АРХИТЕКТУРА, ФУНКЦИОНАЛЬНОСТЬ, СХЕМА ПРИМЕНЕНИЯ

С.А. Рожков

## TERMINAL SYSTEM WTPRO: ARCHITECTURE, FUNCTIONALITY, APPLICATION CIRCUIT

S.A. Rozkhov

В последнее время возрос интерес к терминальным системам. Во многих компаниях накоплен обширный парк компьютерной техники, как правило, морально устаревшей, но технически исправной. Чрезмерная требовательность к аппаратным ресурсам со стороны, например, Windows Vista, требует замены техники. Однако, для офисной работы, имеющиеся средства пригодны. С другой стороны, применение терминалов позволит централизовать управление сетью, что позволит проводить единую политику в области информационной безопасности и не допустить установку вредоносного кода, который может быть внесен недобросовестными или беспечными пользователями рабочих станций. В-третьих, терминальные системы позволяют снизить лицензионные отчисления за использование программного обеспечения. В статье описывается компактная система для работы в тонких клиентах, созданная автором статьи.

*Ключевые слова: тонкий клиент, сетевая загрузка, терминал, безопасность, Linux, Windows.*

An interest to terminal system increases at the last time. There are a fleets of obsolete computers in many firms. But they are technicaly in a good repara. The unreasonable demand of Windows Vista to apparatus resourse demands to change this technic. But it is enough for work at the office. On the other hand using of terminal systems allows to centralize net-administration. It allows to conduct united politic of information security and doesn't allow a putting of harmful code, that can be brought in by the careless users At thirds the using of terminat systems allows to pay not so high assignments for programs. Text runs about using of think clients, created by the auther of article.

*Keywords: thin client, netboot, terminal, security, Linux, Windows.*

### Введение

WTPRO – терминальный клиент для подключения к Windows и Unix серверам. Он не требует операционной системы и жесткого диска. В настоящее время эта разработка используется в более чем ста организациях на территории бывшего СССР. Получены свидетельства авторские и об отраслевой регистрации разработки [1, 2–5]. Создан и поддерживается сайт <http://www.WTPRO.ru>. Разработка отражена в монографической литературе [6].

Терминалом называется устройство ввода и отображения информации. По сути дела – это бездисковая (без HDD-дисковая) рабочая станция, загрузка которой осуществляется по сети. На сервер с клиента передаются нажатия клавиш и движения мыши. Обратно приходят снимки экрана. Все прикладные программы выполняются на терминальном сервере, но для пользователя все выглядит так, как если бы компьютер с такой же вычислительной мощностью, как сервер стоял у него на столе.

### 1. Схема применения

Выделяется сервер. Сервером может быть обычная рабочая станция с большим объемом оперативной памяти. На сервер устанавливается операционная система и все необходимые пользователю программы. Клиентскими машинами могут быть как специализированные тонкие клиенты, так и обычные морально устаревшие компьютеры. На клиентах отсутствуют жесткие диски, а загрузка операционной системы происходит по сети.

Настройка WTPRO производится через веб-интерфейс. Возможно также использование Windows-утилиты для редактирования файла конфигурации.

Чаще всего WTPRO применяют:

- для создания мобильного офиса. В этом случае сервер находится на ноутбуке, а в качестве терминалов используются доступные в данный момент компьютеры. Развернуть подобный учебный класс или мобильный офис можно за считанные минуты;

- организации учебных классов. Используется один мощный сервер, который обновляется, в случае необходимости. Терминальные клиенты, в силу нетребовательности к их ресурсам, обновлять не нужно. Преподаватель имеет возможность со своего рабочего места общаться с учеником;

- организации презентаций. Не всегда есть возможность демонстрировать презентации на большом экране. Но если есть компьютерный класс, то можно вывести изображение на все компьютеры в классе;

- развертывания операторных залов. Организация рабочих мест для большого числа пользователей, работающих с одним набором ПО;

- организации работы бухгалтерии или отдела кадров и т.д.

Минимальные системные требования для работы терминальной системы WTPRO:

- процессор: i486;
- Оперативная память:
  - 6 Мб – без графического режима,
  - 8 Мб – без перенаправления устройств,
  - от 12 Мб – с перенаправлением устройств,
  - от 16 Мб – с библиотекой Xwindow;
- Видеокарта VGA, монитор;
- Клавиатура, мышь;
- Сетевая карта.

## 2. Архитектура

WTPRO поддерживает четыре терминальных протокола – два текстовых:

- telnet – небезопасный протокол, позволяет перехватить имя пользователя и пароль,

- ssh – защищенный аналог telnet

и два графических:

- RDP – удаленный рабочий стол Windows,

- VNC – открытый протокол для доступа как к Windows, так и к UNIX-серверам [8].

Основной сложностью при разработке программного комплекса был компромисс для выполнения двух противоречивых требований:

- Необходимо поддерживать большое количество оборудования.

- Размер комплекса должен быть небольшим. Чтобы поместиться в оперативную память даже морально устаревших машин.

Для экономии оперативной памяти основная программа RDP, которая служит для подключения к терминальным серверам, была разбита на несколько модулей:

- RDP – содержит графическую подсистему, которая обеспечивает видеовывод, обработку операций мыши и клавиатуры;

- librdesktop.so – библиотека, основанная на коде проекта rdesktop, для доступа к функциям протокола rdp;

- libvnc.so – библиотека, основанная на открытых реализациях протокола vnc, предоставляет доступ к функциям протокола vnc;

- libscard.so – библиотека, основанная на патче Алексея Волкова [16] к программе rdesktop для поддержки перенаправления на сервер смарт-карт;

- xrdp – аналог программы RDP, использующей библиотеку видеовывода xorg.

В WTPRO использует три вида драйверов для видеовывода.

Svgalib – облегченная библиотека, рекомендуется для использования на слабых машинах. Поддерживает не все видеокарты, а некоторые современные работают некорректно.

FrameBuffer – данные видеодрайвера включены в ядро Linux. Для своей работы требуют больший объем оперативной памяти, рекомендуется использовать в том случае, если драйвера svgalib работают некорректно.

Xorg – наиболее ресурсоемкая библиотека, хорошо работает с современными видеокартами на машинах с объемом оперативной памяти от 16 Мб.

Так как терминал должен эмулировать полноценную рабочую станцию, то необходимо передавать на сервер не только нажатия клавиш и движения мыши, но и действия локальных устройств: CD-ROM, дисководов, принтеров и т.д.

Рассмотрим, какие возможности поддерживает WTPRO для перенаправления устройств.

CDROM. Перенаправление локально установленного привода для чтения компакт дисков осуществляется средствами протокола RDP, который, начиная с Windows 2003, поддерживает перенаправление дисков. Проблема перенаправления данного привода заключается в следующем. В Windows при вставке компакт-диска в привод CD-ROM происходит автоматическое монтирование (подключение) данного диска, в результате чего он становится доступным пользователю. В UNIX системах принято, что пользователь вручную монтирует необходимые ему устройства. Вторая сложность заключается в том, что пользователь Windows может в любое время извлечь компакт диск, а пользователь UNIX не имеет такой возможности, поскольку при монтировании компакт диска кнопка для извлечения дисков блокируется программно. Для решения этих задач был написан демон (сервис в терминологии Windows), который каждые 5 секунд проверяет статус привода компакт дисков. Если статус изменился на «CDS\_DISC\_OK» и тип диска определен как диск с данными, то пробуем примонтировать сначала его как CD-диск, а в случае неудачи как DVD-диск. Для того чтобы можно было извлечь примонтированный компакт-диск вызываем функцию ядра ioctl (cd, CDROM\_LOCKDOOR, 0), которая запрещает блокировку клавиши извлечения диска. Если статус привода изменился на CDS\_NO\_INFO, CDS\_NO\_DISC, CDS\_TRAY\_OPEN или CDS\_DRIVE\_NOT\_READY, то отмонтируем подключенную файловую систему.

Дисковод. Перенаправление дисковода реализовано аналогично перенаправлению CD-ROM. Различия состоят в том, что дискета может быть

доступна не только для чтения, но и для записи. В силу этого нужен способ для записи данных без стандартного ручного монтирования. Второй сложностью является то, что опрашивать дисковод каждые N секунд нерационально, так как дисковод не имеет состояний как CD-ROM. У дисковода доступно только два состояния – данные читаются и данные недоступны. В любом случае при доступе к дисководу производится шум, а непрерывное гудение дисковода не понравится пользователю. Поэтому было решено использовать патч к ядру `supermount`. Этот патч позволяет один раз примонтировать устройство и в дальнейшем, только при обращении к каталогу, в который примонтировано устройство, обращаться к этому устройству. В отличие от стандартного монтирования данные, которые пишутся на устройство, не кэшируются, а сразу записываются, что исключает потерю информации при извлечении дискеты. Пример для монтирования дисковода

```
mount -t supermount -o fs=vfat,
dev=/dev/floppy/0,-,nosuid,nodev,noexec,
noatime none /ram/mnt/floppy.
```

**Локальные диски.** Если на терминале находится локальный жесткий диск, то можно использовать и его. Для этого используется стандартное монтирование при старте системы. Bash-скрипт просматривает доступные диски и пытается их примонтировать стандартным способом.

**Flash диски.** При монтировании flash-дисков основной проблемой является сохранность данных при извлечении диска из системы. Эта проблема решена аналогично задаче с дисководом. Следующей сложностью является автоматическое монтирование flash-дисков. В данном случае нельзя использовать алгоритм как в случае с локальными дисками, потому что внутренне имя flash-диска в ОС LINUX постоянно меняется при подключении и отключении устройства. Для решения этой задачи был написан основанный на `devfsd` демон [8]. `Devfs` – виртуальная файловая система, которая динамически создает файлы устройств при их подключении и удаляет их после отключения. В ядрах Linux, начиная с 2.6.13 поддержка `devfs` удалена, вместо этого рекомендуется использовать файловую систему `udev`, которая выполняет аналогичные функции, но не на уровне ядра, а на уровне пользовательского приложения. Эта файловая система довольно медлительна (на компьютере Pentium 100 создание устройств занимает около трех минут). Поэтому мною ведется поддержка `devfs` в современных ядрах. Написанный `devfsd`-демон реагирует на создание нового устройства определенного типа и монтирует данные с него.

**Звук.** Перенаправление звука реализовано средствами `gdp`-протокола.

**Смарт-карты.** Перенаправление осуществляется средствами протокола RDP. Так как поддержка смарт-карт ресурсоемкая задача, то библиотека для доступа к функциям смарт-карт была вынесена

из библиотеки `librdesktop.so`. Если в файле конфигурации указано, что необходимо использовать перенаправление смарт-карт, то с сервера загружаются необходимые приложения и библиотеки. Запускается поддержка смарт-карт, и библиотека `librdesktop.so` динамически подключает библиотеку `libscard.so`.

**COM-порты.** В настоящее время перенаправление COM-портов реализовано двумя способами: средствами протокола RDP, средствами дополнительного сервиса. В протоколе RDP перенаправление портов реализовано не полностью, например подключиться к интернету, используя модем, подключенный к COM-порту терминала не получится. Дополнительный сервис прослушивает TCP порт на терминале? При поступлении команд на этот порт они перенаправляются в com-порт, в результате большинство устройств на этом порту работают корректно.

**Принтеры.** В первых версиях WTPRO использовалось перенаправление принтера средствами RDP протокола, но оно имело ряд минусов. Принтер был доступен только при подключении к терминальному серверу. Для подключения принтера необходимо обладать правами администратора. При каждом новом подключении к терминальному серверу необходимо было его переподключать. Поэтому было решено использовать эмуляцию аппаратного принт-сервера от HP.

Аппаратный принт-сервер работает так:

- прослушивается tcp-порт;
- в этот порт поступают «сырые» данные для принтера;
- данные перенаправляются напрямую в устройство.

Для корректной работы принтера достаточно установить стандартные Windows-драйвера на сервер. При использовании этого способа не будут работать «win-принтеры». То есть принтеры, которые не имеют драйверов в стандартной поставке и не работающие под DOS.

### 3. Этапы загрузки терминальной системы

На первом этапе запускается программа, прошенная в загрузочный образ сетевой карты. Эта программа отправляет запрос по протоколу BOOTP/DHCP и узнает у сервера IP-адрес tftp-сервера [9]. С этого сервера скачивается загрузчик `pxelinux`. `Pxelinux` загружает с tftp-сервера свой файл конфигурации, в котором указано, какое ядро необходимо загрузить.

На следующем шаге загружается ядро Linux и образ `initrd` [10]. В WTPRO файл `Initrd.gz` был заменен образом с файловой системой `SquashFS`. В этом образе содержится микро-версия терминала, пригодная для работы на компьютерах с 8 Мб оперативной памяти. При необходимости, по протоколу tftp подгружаются необходимые образы.

Например, так реализована поддержка звука на терминале. Рассмотрим ее подробнее. После

запуска системы, перед подключением к терминальному серверу, проверяется какие расширения необходимо загрузить [11] с tftp сервера. Далее скачивается файл содержащий образ файловой системы SquashFS, в котором находятся модули к ядру Linux с драйверами звуковых карт. После этого запускается программа, определяющая какой модуль необходимо загрузить (используется алгоритм, определяющий по PCI ID нужный модуль) [12]. Загружается необходимый модуль. Отмонтируется образ с драйверами и удаляется файл, в котором этот образ находится. В результате оперативная память расходуется экономно.

#### **4. Терминальный протокол RFB**

Одним из терминальных протоколов является протокол VNC (RFB). Рассмотрим некоторые его особенности. Virtual Network Computing (VNC) система удаленного доступа к компьютеру, использующая RFB-протокол (Remote FrameBuffer). Она передает нажатия клавиш на клавиатуре и движения мыши с одного компьютера на другой, транслируя обновления экрана в обратном направлении по сети.

По протоколу VNC нельзя перенаправлять устройства. VNC очень простой протокол, основанный на графических примитивах: «Положить прямоугольник пикселей на заданную позицию x, y».

Сервер посылает небольшие прямоугольники framebuffer'a клиенту. Такая схема в своей примитивной форме потребляет большую часть пропускной возможности канала. Для снижения нагрузки на канал используются различные методы. Существует вариант использования сжатия – метод определения наиболее эффективного способа передачи прямоугольников пикселей.

VNC-протокол позволяет клиенту и серверу договориться об используемом методе кодирования. Самый простой метод кодирования, поддерживаемый всеми клиентами и серверами – «raw encoding». В этом случае пиксели передаются в порядке слева-направо, сверху-вниз. А после передачи первоначального состояния экрана, передаются только изменившиеся пиксели.

Этот метод работает очень хорошо при незначительных изменениях изображения на экране. Например, при движении указателя мыши по рабочему столу, в случае набора текста под курсором и т.д. Однако нагрузка канала становится очень высокой при одновременном изменении большого количества пикселей, например при просмотре видео.

В отличие от RDP-протокола VNC-сервер для авторизации не использует системные имена пользователей и пароли. Он вообще не использует имена пользователей. Вместо имени пользователя используется номер порта [13]. На Unix-машинах можно запустить несколько VNC-серверов. На Windows-машинах мы ограничены только одним сеансом. То есть при работе в Unix наш сервер

поддерживает столько клиентов, сколько VNC-серверов запущено. В Windows все пользователи работают с одним сеансом, а именно с локальной консолью сервера. Все пользователи используют одну и ту же клавиатуру и одну и ту же мышь. Поэтому вероятны конфликты, когда пользователи на различных компьютерах будут тянуть в разные стороны одну и ту же мышь. В связи с этим в VNC-сеансе существует два вида паролей: пароль на чтение и запись (зная этот пароль, пользователь полностью управляет сеансом); пароль только на чтение (пользователь видит, что происходит в сеансе, но не может вмешиваться – данный режим удобен для презентаций).

Еще одной особенностью протокола является возможность «разделить» сессию – разрешить нескольким пользователям параллельно работать с одним портом и использовать общий сеанс. В этом режиме целесообразно использовать пароль «только на чтение». В режиме презентации мы делаем общедоступной одну сессию, а все остальные пользователи подключаются, используя пароль только на чтение. В результате, все, что происходит на центральном компьютере транслируется на все остальные машины. В ряде случаев это удобней использования проектора, так как изображение находится в непосредственной близости от пользователя.

#### **5. Оптимизация**

Для того чтобы WTPRO мог работать на системах с очень ограниченными ресурсами, нами было проведено множество оптимизаций. В частности стандартная, но «тяжелая» библиотека glibc была заменена на легкую, но урезанную функционально библиотеку uclibc. Многие программы были перекомпилированы со статическим включением библиотек [14]. При статической линковке приложений в полученную программу включается только используемый код, а при динамической линковке в библиотеке остается даже тот код, который не используется в программе. Некоторые программы, например, librdesktop.so, требуют установки дополнительных библиотек, таких как libiconv.so, которые занимают много места. Эти библиотеки были заменены эквивалентным кодом, достаточным для выполнения программ. Таким образом, удалось на 70 % сократить размер бинарного кода. Многие стандартные утилиты были нами переписаны таким образом, что теперь используют только вызовы ядра и не требуется ни одной библиотеки. Например, программа перезагружающая компьютер, которая идет с дистрибутивом Fedora Core 4 занимает 11 Kb и требует две библиотеки libc (1,5 Mb) и ld (124 Kb). После компиляции наша версия программы занимает 624 байта и не требует ни одной библиотеки.

Также сильной модификации подверглось ядро Linux. Как уже было сказано выше, в нашем ядре присутствует поддержка devfs. В ядро вклю-

чены патчи для нестандартных файловых систем SquashFS, supermount. Также в ядро мною был включен патч Linux-tiny, который позволяет запускать операционную систему на машинах с весьма ограниченными системными ресурсами [15]. Данный патч уже не поддерживается разработчиками Linux, начиная с версии ядра 2.6.14. Теперь он поддерживается нами.

Этот патч реализует следующие функции: удаляет неиспользуемые системные вызовы, в нем отсутствует вывод сообщений ядра, «тяжелый» код заменен на менее функциональный, но существенно меньший по размеру.

Так же большим ограничением является то, что на терминальном клиенте нельзя хранить настройки (отсутствует жесткий диск). Поэтому используется несколько способов конфигурации клиента. Часть параметров, о конфигурации сети и IP-адресах терминальных серверов, передаются по DHCP. Остальные настройки, в случае необходимости, загружаются посредством протокола TFTP. Некоторые параметры – видеокарта, разрешение, тип мыши т. д. определяются автоматически [16].

#### Заключение

Результатом данной работы является разработка, внедрение и поддержка терминального клиента WTPRO. Тонкий клиент имеет модульную архитектуру. Это позволяет запускать его как на устаревших, так и на современных компьютерах. Нами была произведена оптимизация кода и реализованы алгоритмы, для автоматического определения оборудования и загрузки его на терминал. Получившийся продукт может подключаться как к Windows, так и UNIX подобным серверам.

#### Литература

1. А.с. РФ. Препроцессор «Elinux»: свидетельство об отраслевой регистрации разработки № 5399 / С.А. Рожков. – № 50200501646; заявл. 08.11.2005.; опублик. 05.12.2005; Компьютерные учебные программы и инновации № 9(10). – 1 с.
2. А.с. РФ. Терминальная система ElinuxT: свидетельство об отраслевой регистрации разработки № 5491 / С.А. Рожков. – № 50200501788; заявл. 08.11.2005.; опублик. 20.12.2005; Компьютерные учебные программы и инновации № 9(10). – 1 с.
3. А.с. РФ. Терминальная система «ElinuxT»: свидетельство об официальной регистрации программы для ЭВМ № 2006611848. Зарегистрировано в Реестре программ 30.05.2006.

4. Рожков, С.А. Защищенная терминальная система WTPRO: свидетельство о государственной регистрации программы для ЭВМ / С.А. Рожков. – № 2009611320. Зарегистрировано в Реестре программ 04.03.2009.

5. А.с. РФ. Защищенная терминальная система WTPRO: свидетельство об отраслевой регистрации разработки № 12153 / С.А. Рожков. – № 50200900159; заявл. 08.12.2008.; опублик. 16.01.2009.

6. Стахов, А. Linux-сервер в Windows-окружении / А. Стахов. – СПб.: БХВ, 2007. – 656 с.

7. [http://sourceforge.net/tracker/index.php?func=detail&aid=1314556&group\\_id=24366&atid=381349](http://sourceforge.net/tracker/index.php?func=detail&aid=1314556&group_id=24366&atid=381349)

8. Вахалия, Ю. UNIX изнутри / Ю. Вахалия. – СПб.: Питер, 2003. – 844 с.

9. Таненбаум, Э. Современные операционные системы / Э. Таненбаум. – 2-е изд. – СПб.: Питер, 2002. – 1040 с.

10. Таненбаум, Э. Распределенные системы. Принципы и парадигмы / Э. Таненбаум. – СПб.: Питер, 2003. – 877 с.

11. Хелд, Г. Технологии передачи данных / Г. Хелд. – СПб.: Питер, 2003. – 720 с.

12. Рожков, С.А. Дистрибутив Elinux [Текст] / С.А. Рожков // Безопасность информационного пространства: материалы Всерос. науч.-практ. конф. – Екатеринбург: УГТУ-УПИ, 2005. – С. 83–84.

13. Рожков, С.А. Терминальные системы для предприятий / С.А. Рожков // Снежинск и наука – 2006: сб. науч. тр. междунар. науч.-практ. конф. – Снежинск: СГФТА, 2006. – С. 184–186.

14. Рожков, С.А. Разработка терминальной системы ElinuxT / С.А. Рожков // Каталог тезисов проектов научно-технических коллективов молодых ученых, аспирантов и студентов, представленных на всероссийский конкурсный отбор по приоритетным направлениям науки и высоких технологий. – М.: РГУТИ, 2006. – С. 109–111.

15. Рожков, С.А. Терминальные системы для предприятий / С.А. Рожков // Безопасность информационного пространства: материалы междунар. науч.-практ. конф. – Екатеринбург: УрГУПС, 2006. – С. 96–97.

16. Рожков, С.А. Сравнение терминальных систем / С.А. Рожков // Безопасность информационного пространства VI: сб. тр. межвуз. науч.-практ. конф. студентов, аспирантов и молодых ученых. – Тюмень: Изд-во ТюмГУ, 2007. – С. 183–188.

Поступила в редакцию 10 апреля 2009 г.



# МОДЕЛЬ МЕТАМОРФНОГО ПРЕОБРАЗОВАНИЯ ИСПОЛНЯЕМОГО КОДА

*П.В. Збицкий*

## EXECUTABLE CODE METAMORPHIC TRANSFORMATION MODEL

*P.V. Zbitskiy*

В работе рассмотрена формализация основных техник изменения исполняемого кода с использованием элементов теории формальных грамматик и автоматов. Предложена новая модель метаморфных вирусов, которая учитывает методы расширения и сжатия кода. Обосновано утверждение об недетектируемости вирусов, построенных с учетом данной модели.

*Ключевые слова: компьютерный вирус, метаморфизм, полиморфизм, математическая модель, формальная грамматика, конечный автомат.*

In the paper built model of metamorphic transformation of executable code by using finite automaton and formal grammars. This is a new model which considers code-growing and code-shrinking techniques. The statement about static undetectable of viruses built by described method is proved

*Keywords: computer virus, polymorphism, metamorphism, mathematical model, formal grammar, finite automaton.*

### Введение

Целью работы является рассмотрение существующих моделей полиморфных и метаморфных вирусов, а также усовершенствование моделей и приведение их в соответствие с реально существующими вариациями вирусов данных типов. Обсуждается вопрос обнаружимости полиморфных и метаморфных вирусов. Также нельзя обойти вниманием применимость вирусных техник в «мирном» программном обеспечении: *практически все* современные системы навесной защиты ПО (упаковщики и протекторы) применяют приемы, которые впервые были использованы при написании вирусов.

### 1. Полиморфные вирусы

Следуя [6], определим *полиморфный вирус* как вирус, состоящий из зашифрованной полезной нагрузки, полиморфного генератора и дешифратора – изменяемой процедуры расшифрования.

Полиморфный генератор – процедура, обеспечивающая изменение вируса при распространении. Обычно состоит из модуля шифрования полезной нагрузки и генератора соответствующих различных дешифраторов.

Впервые модель полиморфного генератора на основе формальных грамматик предложена в 1999

году [1]. (Основные понятия формальных языков и грамматик описаны, например, в [9–11].)

Пусть  $G = (N, T, P, S)$  – формальная грамматика,  $T = \{a, b, c, d, x, z\}$  – терминальный алфавит, состоящий из инструкций процессора *x86*;  $a, b, c, d$  – мусорные инструкции;  $x, z$  – инструкции дешифратора;  $N = \{A, B, S\}$  – нетерминальный алфавит,  $S$  – начальный символ.

$$P = \begin{cases} S \rightarrow aS \mid bS \mid cS \mid dS \mid xA \\ A \rightarrow aA \mid bA \mid cA \mid dA \mid zB \\ B \rightarrow aB \mid bB \mid cB \mid dB \mid \varepsilon \end{cases} \text{ – правила вывода.}$$

Полученный, согласно данным правилам, дешифратор может быть, например, таким: *abbcxaaazd*.

Также в [1] рассмотрен вопрос обнаружения данного полиморфного генератора путем построения соответствующего конечного автомата (рис. 1).

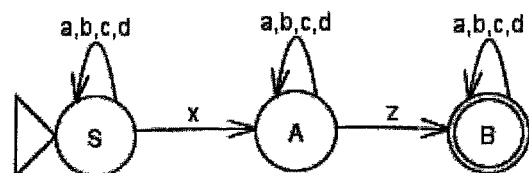


Рис. 1. Автомат, распознающий простой вирусный дешифратор

Действительно, для регулярных грамматик (которой является рассмотренная), построение автомата, распознающего грамматику, возможно [8, 9, 10]. Однако, данное утверждение [1] некорректно распространено и на другие типы грамматик (в частности, типа 0 по Хомскому). Публичное опровержение данного факта дано лишь в 2007 году [2]. Анализ показывает [7], что успешность сигнатурного подхода детектирования полиморфных вирусов обеспечена использованием полиморфных движков, которые могут быть представлены соответствующей регулярной грамматикой. Данный факт еще раз подтверждает недостаток теоретических работ в области компьютерной вирусологии.

Таким образом, использование при проектировании полиморфных генераторов грамматик общего предоставляет возможность создания действительно необнаружимых вирусов.

В 1999–2000-х годах появились первые метаморфные вирусы. Самые известные представители: metaPHOR (Simile), Evol, Zmorph, Zmist...

Согласно [6], **метаморфный вирус** – это вирус, полностью изменяющий свой код при распространении (в отличие от полиморфных вирусов, которые изменяют только декриптор, а остальную часть просто шифруют).

## 2. Модель метаморфного генератора

Рассмотрим подход к формализации метаморфных вирусов на основе грамматик, подсказанный анализом метаморфного вируса Zmist [7]. Данный вирус при распространении использует сохраненный «скелет», который и изменяется при производстве новых копий. Вирус дизассемблирует инструкции скелета и последовательно мутирует в новую версию, в которой также присутствует сохраненный скелет.

Данный тип метаморфных вирусов можно смоделировать с помощью формальных грамматик: полезная нагрузка объединяется с метаморфным генератором в виде правил некоторой формальной грамматики. То есть, каждая новая копия вируса должна иметь зашифрованный алгоритм, при выполнении которого будут получаться новые различные копии метаморфного вируса. Назовем вирусы такого типа **полуметаморфными вирусами**, так как они попадают под определение метаморфного вируса (изменение основного тела программы), но вынуждены иметь некую постоянную составляющую, подобно полиморфным вирусам. Для наглядности рассмотрим простой пример.

Допустим, нам нужно создать простую полуметаморфную программу, завершающую работу ОС Windows:

```
push 0
push 4
call ExitWindowsEx
push 0
call ExitProcess
```

(1)

Составим грамматику, описывающую метаморфные преобразования данного кода. Предположим, что адреса используемых функций уже разрешены. Также для простоты исключим генерацию кода декриптора зашифрованного/сжатого метаморфного движка.

Пусть  $G = (N, T, P, S)$  – формальная грамматика;  $T$  – терминальный алфавит, состоящий из инструкций процессора x86, и  $a, b, c, d$  – мусорные инструкции;  $N$  – нетерминальный алфавит;  $S$  – начальный символ.

Обозначим  $x \oplus y$  конкатенацию инструкций  $x$  и  $y$ ,  $EW$  – адрес функции ExitWindowsEx,  $EP$  – ExitProcess, и определим правила вывода грамматики:

1.  $S \rightarrow aS | bS | cS | dS | (push\ 0)A | (xor\ ebx, ebx \oplus push\ ebx)A | (sub\ esp, 4 \oplus mov\ [esp], 0)A$
2.  $A \rightarrow aA | bA | cA | dA | (push\ 4)B | (mov\ eax, N \oplus xor\ eax, <N\ xor\ 4> \oplus push\ eax)B$
3.  $B \rightarrow aB | bB | cB | dB | (call\ EW)C | (push\ \$+10 \oplus jmp\ EW)C | (mov\ esi, EW \oplus call\ esi)C$
4.  $C \rightarrow aC | bC | cC | dC | (push\ 0)D | (xor\ ebx, ebx \oplus push\ ebx)D | (sub\ esp, 4 \oplus mov\ [esp], 0)D$
5.  $D \rightarrow aD | bD | cD | dD | (call\ EP)E | (push\ \$+11 \oplus push\ EP \oplus ret)E | (mov\ esi, EP \oplus call\ esi)E$
6.  $E \rightarrow aE | bE | cE | dE | \varepsilon$

После разбора примера сразу видны недостатки генераторов такого вида: во-первых, огромный размер генератора, ведь для каждой инструкции полезной нагрузки соответствует целый набор правил, а во-вторых, «простота» грамматики.

Перепишем правила вывода в виде (остальные правила остались без изменения):

1.  $S \rightarrow XA$
4.  $C \rightarrow XD$
7.  $X \rightarrow aX | bX | cX | dX | (push\ 0) | (xor\ ebx, ebx \oplus push\ ebx) | (sub\ esp, 4 \oplus mov\ [esp], 0)$

Полученная грамматика не является регулярной, это неукорачивающаяся грамматика, для которой проблема выводимости ( $w \in L(G)$  или  $w \notin L(G)$ ) также алгоритмически разрешима [8–10]. Немного упростим модель.

Положим  $G = (N, T, P, S)$  – грамматика, в правилах которой описан алгоритм программы, правила вывода имеют вид

$$P = \begin{cases} S \rightarrow A_1 A \\ A \rightarrow B_1 B \\ \dots \\ X \rightarrow X_1 X \end{cases}$$

В этом случае программа составлена в виде последовательности  $A_1 B_1 \dots X_1$ , где  $A_1, B_1, \dots, X_1$  можно рассматривать как внутренний язык, на котором написана программа. Таким образом, правила дан-

ной грамматики задают скелет программы. Назовем данные правила метаправилами, а грамматику – метаграмматикой. Метаправила – это «заготовки» для конкретных правил. Далее, рассмотрим грамматику  $G_1 = (N_1, T_1, P_1, S_1)$ , правила которой будут задавать правила трансляции «внутреннего языка» в конкретные команды или блоки кода.

Например,

$$P_1 = \begin{cases} A_1 \rightarrow \text{push } 0 | (\text{xor } ebx, ebx \oplus \text{push } ebx) \\ \dots \\ X_1 \rightarrow (\text{mov } eax, EP \oplus \text{call } eax) | (\text{mov } ebx, EP \oplus \text{call } ebx) \end{cases}$$

Грамматика  $G_1$  описывает мутацию 1-го поколения,  $G_2$  – второго и т.д. Тут возникает самый главный вопрос: как из  $G_1$  получить  $G_2$  таким образом, чтобы правила трансляции при каждой мутации изменялись? Можно сделать  $G_1$  максимально объемной, и на каждом шаге выбрасывать из нее некоторые правила. Минус такого подхода очевиден – через конечное число шагов, вирус перестанет мутировать. Как вариант решения – всегда иметь копию максимальной грамматики  $G_1$ , и при достижении критического уменьшения количества правил в  $G_i$ , грамматику  $G_{i+1}$  генерировать снова из  $G_1$ .

С другой стороны, программа (1) написана на собственном языке.

1. *Завершить работу Windows*
2. *Выйти из программы*

Это открывает более широкие возможности для изменения кода, чем на уровне команд процессора: каждую высокоуровневую команду можно проинтерпретировать разными наборами команд процессора, а каждую команду процессора также можно проинтерпретировать ее эквивалентами.

Назовем вирусы, которые строят новую версию потомка на основании дизассемблирования и мутации собственного кода, **метаморфными вирусами** общего типа.

Ниже рассмотрены некоторые проблемы, связанные с реализацией метаморфных генераторов этого типа, а также математическая модель подобного преобразования.

Классический недетерминированный автомат, моделирующий подобного типа вирусы, выглядит весьма громоздко, так как для каждого состояния предусмотрены все возможные входные символы (рис. 2):

Опишем автомат формально: автомат

$$A = (Q, \Sigma, \delta, q_0),$$

множество состояний

$$Q = \{q_0\} \cup \{\text{инструкции x86}\},$$

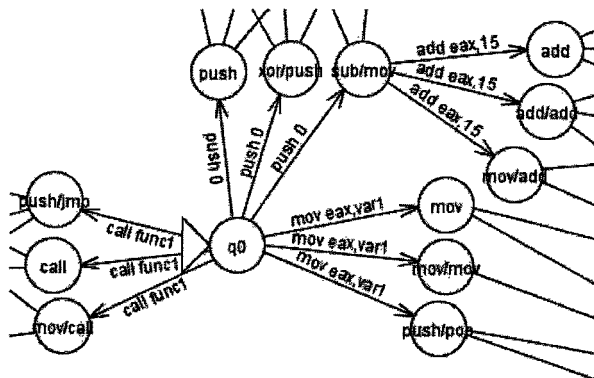
входной алфавит

$$\Sigma = \{\text{инструкции x86}\},$$

функция перехода

$$\delta : Q \times \Sigma \rightarrow Q.$$

Входная программа – некоторое слово (цепочка) из  $\Sigma^*$ . Мутацией в таком случае будет путь автомата  $q_1 q_2 \dots q_n$ , то есть набор состояний, которые он посетил при обработке входного слова. Это модель *универсального метаморфного генератора*.



**Рис. 2. Автомат, моделирующий метаморфный генератор**

Однако стоит заметить, что функция перехода  $\delta$  фактически задает грамматику, согласно правилам которой и работает автомат. Проблема в том, что грамматика, как и автомат, должна быть «связной». Проще говоря, из начального символа нужно уметь вывести произвольную последовательность инструкций.

Опишем данную грамматику формально:

$$G = (N, T, P, q_0) \text{ – грамматика,}$$

$N = \{q_0, A, B, C, \dots\}$  – ее нетерминальный алфавит,

$T = \{a_1, a_2, a_3, \dots, z_n\} = \{\text{инструкции x86}\}$  – терминальный алфавит,

$q_0$  – начальный символ.

Правила вывода имеют вид

$$P = \begin{cases} q_0 \rightarrow AA | BB | \dots | ZZ \\ A \rightarrow BB | CC | \dots | ZZ \\ \dots \\ A \rightarrow a_1 | a_2 | \dots | a_n \\ Z \rightarrow z_1 | \dots | z_m \end{cases}$$

Данный набор правил реализует следующие соображения:

1. Каждый нетерминал представляет собой все варианты трансляции некоторой команды.
2. Двойное вхождение нетерминала в правую часть правил обеспечивает «связность» – выводимость произвольных последовательностей.

Работа описанного генератора выглядит следующим образом: получаем цепочку нетерминалов, которую затем транслируем в терминальные символы – инструкции процессора. Например:

$$q_0 \rightarrow BB \rightarrow BDD \rightarrow BDKK \rightarrow BDKLL \rightarrow \dots \\ \rightarrow BDKLMN \rightarrow b_1 d_5 k_2 l_4 m_8 n_{13}.$$

Таким образом, получена неукорачивающаяся

(бесконтекстная) грамматика, что, в общем-то, неудивительно, так как она была построена по автомату. Но теперь можно применить элементы теории формальных грамматик к оценке сложности обнаружения мутации такого вируса, а также попробовать изменить грамматику.

Проблема рассмотренного автомата и грамматики (как и многих метаморфных преобразователей вообще) в том, что мутировавший код получается большего размера, чем исходный.

Заманчивой выглядит идея переписать правила сжатия следующим образом. Пусть  $X, Y, Z$  – команды *x86*, причем  $XY \equiv Z$ . Тогда правило вывода  $XY \rightarrow Z$  будет соответствовать тому, что последовательность инструкций  $XY$  сжимается в инструкцию  $Z$ .

Рассмотрим пример. Пусть  $X \equiv \text{xor eax, eax}$ ,  $Y \equiv \text{push eax}$  и  $z \equiv \text{push 0}$ . Имеется следующий набор правил вывода:

- $M \rightarrow XX$
- $X \rightarrow YY$
- $Y \rightarrow NN$
- $XY \rightarrow z$

В этом случае, вывод (работа метаморфного генератора) имеет вид

$$M \rightarrow XX \rightarrow XYY \rightarrow \begin{cases} zY \rightarrow zNN \rightarrow \dots \\ XYNN \rightarrow \dots \end{cases}$$

Отметим, что обе ветки семантически эквивалентны – поместить на вершину стека нуль.

Возникает вопрос, как в этом случае изменить автомат? Пусть автомат находится в состоянии  $M$ , и ему на вход подается символ *xor eax, eax*. Автомат при этом попадает в состояние  $X$ , соответствующее одной из трансляций команды *xor eax, eax*. Далее автомату приходит символ *push eax*. Наряду с переходами в состояние  $Y$ , имеется переход в состояние  $z$ , соответствующее команде *push 0*. Состояние *push 0* особое. При попадании в него, автомат должен выкинуть предыдущее состояние  $X$  из своего пути, как будто бы существовала дуга  $(M, z)$  и  $z \equiv \text{push 0}$ . Рис. 3 демонстрирует описанный принцип сжатия кода.

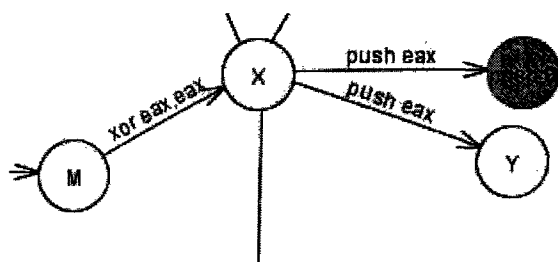


Рис. 3. Часть автомата, поясняющая сжатие кода

Отметим, что ни о какой формальной эквивалентности первоначальных и полученных грамма-

тик и автоматов речи нет. Приведенные рассуждения показывают гибкость грамматико-автоматного подхода при моделировании метаморфных генераторов.

Итак, в грамматике появились правила вывода типа  $XY \rightarrow z$ , следовательно, грамматика перестала быть неукорачивающейся. Более того, это грамматика общего вида. Для грамматик этого типа проблема принадлежности слова языку, порожденного грамматикой, неразрешима [8–10]. Данный факт еще раз доказывает возможность создания недетектируемых метаморфных вирусов.

### 3. Обсуждение

Предложена общая модель метаморфного вируса [2].

Пусть  $G_1 = (N, T, P, S)$ ,  $G_2 = (N', T', P', S')$  – грамматики,  $T'$  – множество формальных грамматик,  $S' = G_1$  – стартовая грамматика,  $R'$  – правила вывода относительно  $(N' \cup T')$ . Тогда метаморфный вирус описывается  $G_2$  и каждая из его мутаций есть слово в  $L(L(G_2))$ , то есть слово в  $L(x)$  для некоторого  $x \in L(G_2)$ , где  $x$  – грамматика.

Данное определение описывает факт, что при мутации от одной формы до другой, ядро вируса изменяется, а также изменяются сами правила мутации. На основе данной модели [2] построен теоретический прототип необнаруживаемого метаморфного вируса *PB\_MOT*. Стоит отметить, что на практике данная модель крайне трудно реализуема.

Этого недостатка лишены рассмотренные выше модели полуметаморфного вируса и метаморфного вируса на основе грамматико-автоматного подхода.

Основным методом детектирования метаморфных вирусов является поведенческий анализ, основная проблема которого – необходимость исполнения вирусного кода. В противовес данному подходу развиваются идеи статического исследования метаморфных программ [3–5]. Эти подходы построены на следующем факте: чтобы вирус смог создать новую мутацию, он должен уметь дизассемблировать себя. А если вирус может себя дизассемблировать, то сможет это сделать и антивирусное ПО. В [2, 7] показывается несостоятельность данного подхода при использовании методов обфускации. Но можно пойти еще дальше и лишить создателей систем статического анализа их главного тезиса о возможности дизассемблирования: рассмотренным выше полуметаморфным вирусам *не требуется* проводить дизассемблирование, чтобы создать новую полноценную метаморфную копию.

### Выводы и развитие результатов

Основное назначение рассмотренных техник мутаций кода – создание абсолютно недетекти-

руемого антивирусами кода. Грамматико-автоматный подход позволяет данный факт доказать формально. Стоит отметить, что на сегодняшний день существуют алгоритмы, способные частично детектировать метаморфные вирусы. Но все они содержат весьма жесткие ограничения, следовательно, необходимы новые исследования в данной области. Но помимо самих вирусов и их эффективного обнаружения, отдельно стоит применимость метаморфизма в индустрии защиты ПО. И здесь открываются широкие перспективы: от внедрения водяных знаков в программы на уровне исполняемого кода до генерации новой копии программы при распространении ПО через Интернет с целью исключить саму возможность применения патчей как таковую.

#### **Литература**

1. Qozah. *Polymorphism and grammars* / Qozah // *29A E-zine*. – 1999. – № 4.
2. Filiol, E. *Metamorphism, Formal Grammars and Undecidable Code Mutation* / E. Filiol // *Proc. of PWASET*. – 2007. – V. 20.
3. Bruschi, D. *Using Code Normalization for Fighting Self-Mutating Malware* / D. Bruschi, L. Martignoni, M. Monga // *Security & Privacy, IEEE*. – 2007. – V. 5. – P. 46–54.
4. Lakhota, A. *Are metamorphic viruses really invincible?* / A. Lakhota, A. Kapoor // *Virus Bulletin*. – Dec. 2004. – P. 5–7.
5. Zhang, Q. *MetaAware: Identifying Metamorphic Malware* / Q. Zhang, D. Reeves // *Proc. of the 23rd Annual Computer Security Applications Conference*. – Dec, 2007.
6. Szor, P. *Hunting for Metamorphic* / P. Szor, P. Ferrie // *Virus Bulletin*. – Sept, 2001. – P. 123–144.
7. Filiol, E. *Computer viruses: from theory to applications* / E. Filiol. – France: Springer-Verlag, 2005. – 405 p.
8. Szor, P. *The Art of Computer: Virus Research and Defense* / P. Szor. – USA: Symantec Press, 2005. – 744 p.
9. Гросс, М. *Теория формальных грамматик* / М. Гросс, А. Лантен. – М.: Мир, 1971. – 296 с.
10. Гладкий, А.В. *Формальные грамматики и языки* / А.В. Гладкий. – М.: Наука, 1973. – 368 с.
11. Пентус, А.Е. *Теория формальных языков: учебное пособие* / А.Е. Пентус, М.Р. Пентус. – М.: изд-во ЦПИ при МГУ, 2004. – 80 с.

**Поступила в редакцию 10 апреля 2009 г.**

# НАБЛЮДЕНИЕ АНОМАЛИЙ В СЕТИ

**А.А. Сапожников**

## NETWORK ANOMALIES OBSERVATION

**A.A. Sapozhnikov**

Основная цель данной работы – создание системы мониторинга ИТ-инфраструктуры предприятия на основе различных продуктов с открытыми исходными кодами. Элементы данной системы должны требовать минимум усилий по конфигурированию, и динамически адаптироваться к изменяющимся параметрам наблюдаемых систем.

*Ключевые слова: обнаружение вторжений, аномалии, сети, мониторинг.*

That paper discusses one of network monitoring problem decision based on various open source software. Monitoring system is easily configurable and dynamically adopt to variations of monitored network.

*Keywords: intrusion detection, anomaly, networks, monitoring.*

### Введение

Локальные вычислительные сети современных предприятий наполнены различными приложениями и решают множество задач. Поведение сети формируется пользователями, сервисами программно-аппаратных средств, другими сетевыми устройствами. Под нормальным функционированием сети мы имеем в виду следующее: гарантированное предоставление сервисов и устойчивость к различным стрессовым воздействиям, которые в самом общем рассмотрении можно разделить на непроизвольные (отказ оборудования, ошибки в программах) и произвольные (целенаправленные атаки).

### 1. Наблюдаемые события

Существуют три уровня обороны ИТ-инфраструктуры. Уровень первый – межсетевой экран. На данном уровне происходит защита от ненужного или злонамеренного трафика. Второй уровень – сетевые системы обнаружения вторжений (NIDS). Данные системы детектируют нежелательный и злонамеренный трафик внутри ИТ-инфраструктуры. Третий уровень – верификация целостности данных – отслеживание изменений в системе, например, путем сравнения криптографических контрольных сумм файлов с данными в базе.

Для достижения максимального уровня защищенности объекта ИТ-инфраструктуры, необходимо реализовывать на нем все три уровня защиты, причем каждый уровень должен быть независимым, т.е. в том случае, если злоумышленник проникнет за один из уровней защиты, то все остальные продолжают полноценно функционировать. Таким образом, реализуется добавочная защита [1].

Количество возможных векторов удаленной атаки, весьма ограничено по сравнению с локальными атаками.

Как правило, предотвратить атаку невозможно. Если известно, что какая-то система уязвима, то значительно эффективнее будет ликвидировать данную уязвимость, либо удалить систему из ИТ-инфраструктуры до полного восстановления статуса защищенной системы.

Именно поэтому есть смысл сконцентрировать все усилия на мониторинге последствий атаки.

Атаки можно разделить на три уровня.

- Транспортный уровень.

Под атаками на транспортный уровень будем понимать, атаки на протоколы канального, сетевого и транспортного уровней стека TCP/IP. К этому уровню относятся, например, различные виды сканирования, arp-spoofing, ip-spoofing, ping of death [4].

- Прикладной уровень.

Под этим уровнем будем понимать атаки направленные на ошибки в реализации различных протоколов прикладного уровня стека TCP/IP.

Примерами подобных атак могут быть dns cache-poisoning, smb-die, apache http-chunked encoding.

- Уровень сервиса.

К атакам данного уровня отнесем, всевозможные атаки вызванные ошибками в некорректной обработке пользовательских данных. Примерами подобных атак будут XSS, SQL-injection, различные переполнения, двойное освобождение памяти и т.д. [5, 6]

Обнаружить атаку можно двумя методами.

- Сигнатурный. Данный метод сводится к поиску признаков уже известных атак.

• Аномальный. Заранее известно какими функциональными параметрами обладает то или иное приложение или сервис в нормальном состоянии, и любое отклонение от него считается атакой.

Преимущество сигнатурного метода в том, что он практически не подвержен ложным срабатываниям. Минусом данного метода, является невозможность обнаруживать не заложенные в систему атаки.

Метод поиска аномалий позволяет реагировать на ранее неизвестные атаки, но подвержен ложным срабатываниям и требует точной настройки для каждого наблюдаемого объекта.

Оба метода обнаружения атак могут работать на всех трех уровнях.

Результатом любой атаки, в случае успеха, является утечка информации, например, атака ether-leak, либо изменение каких-либо параметров атакующей системы, например открытие порта, прекращение работы некоторой программы, значительное изменение объемов использования некоторого ресурса системы программой (например, памяти, процессорного времени и т.п.), запуск новой программы на выполнение. Подобные события, происходящие в системе, можно отслеживать с помощью программных средств [2].

Так как значительное изменение объемов использования ресурсов может происходить и при нормальном режиме работы системы (например, увеличение обращений к почтовому серверу в начале рабочего дня), то возникает проблема определения поведения программы: нормального или аномального. Предлагается решение данной проблемы на основе статистических методов анализа.

## **2. Обнаружение аномалий**

Идея заключается в следующем. В основном, все реальные автоматизированные информационные системы имеют циклический характер функционирования, который определяется рабочей неделей или производственным технологическим процессом. Предположим, что существует отдельный период на начальном этапе работы системы, в течение которого мы можем утверждать, что система работает в нормальном режиме. Если такой интервал времени существует, то назовем его периодом обучения [7].

В течение периода обучения будем отслеживать использование различных ресурсов программами и на базе накопленной информации сможем построить функцию прогнозирования дальнейшего поведения программ.

В период рабочей эксплуатации системы будем постоянно производить мониторинг использования различных ресурсов программами. Полученные данные будем сравнивать с прогнозируемыми. Если различие между прогнозируемым и фактическим использованием ресурсов превышает некоторое допустимое значение, то поведение

программы считается аномальным. При этом принимается решение об изменении режима функционирования программы с целью предотвращения нарушения стабильности работы всей системы в целом.

В некоторых случаях возможно применять более простой способ обнаружения аномалий, для этого у наблюдаемого параметра системы определяются верхняя и нижняя границы его некоторой числовой характеристики в период нормальной работы. В случае же выхода за их пределы, детектируется аномальное поведение.

Вырожденным случаем определения границ, является полное отсутствие, или близкое к нулю, какого-либо параметра. Например, о наблюдаемой системе известно, что на ней функционирует только веб-сервер, т.е. в период обучения наблюдался только трафик на порт 80, но в результате атаки, обнаруживается трафик на порт 6667, которого раньше не было.

## **3. Структура системы**

Разработанная система представляет собой совокупность следующих модулей:

- сбора статистики;
- обучения;
- прогнозирования дальнейшего поведения программ;
- реагирования на аномальное поведение программ.

Модульная структура позволяет создать приложение легко переносимое на различные платформы, причем в дальнейшем возможно будет реализовать клиент-серверную архитектуру без внесения больших изменений, поскольку интерфейс обмена данными между модулями строго унифицирован.

Одной из проблем мониторинга больших сетей является огромное количество информации, которое необходимо проанализировать, поэтому с целью снижения требований к аппаратной реализации модулей сбора статистики, необходимо увеличить их число с целью реализации распределенной системы мониторинга, что повысит отказоустойчивость системы в целом [9].

Модули сбора статистики получают информацию об использовании различных ресурсов автоматизированной системы программами и передают ее модулю анализа через стандартизированный интерфейс. Это позволяет сделать наиболее сложные модули обучения, анализа и прогнозирования поведения программ платформенно-независимыми.

Модуль обучения анализирует данные, собранные модулем статистики в период обучения, и строит функцию, прогнозирующую дальнейшее поведение программы. Функция прогнозирования строится на базе функциональной зависимости использования ресурса программой от времени в

течение периода обучения (далее исходная функция) следующим образом.

Вначале производится сглаживание исходной функции для отфильтровывания случайных шумов. Сглаживание выполняется вычислением скользящего взвешенного пятиточечного среднего с оптимально подобранными весовыми коэффициентами:

$$WMA_t = \frac{\sum_{i=t-5}^t h_i \cdot W_{t-i}}{\sum_{i=1}^5 W_i}, \quad (1)$$

где  $t$  – время измерения параметра,  $W$  – весовой коэффициент,  $h$  – наблюдаемый параметр.

Далее производится разделение исходной функции на две составляющие. Первая – тренд – определяет тенденцию в использовании ресурса. Вторая – сезонная компонента – определяет периодическую составляющую исходной функции. Выделение тренда осуществляется вычислением математического ожидания значений исходной функции на достаточно большом временном отрезке. Сезонная компонента находится как разность сглаженной исходной функции и тренда. При дальнейшем анализе сезонной компоненты производится выделение основных гармонических составляющих. Для этого используется дискретное преобразование Хартли (аналог преобразования Фурье) с последующим отбрасыванием малозначимых коэффициентов.

Прямое дискретное преобразование Хартли

$$H_k = \sum_{i=0}^{N-1} h_i \cdot \cos\left(\frac{2\pi}{N} ki\right), \quad k = 0, \dots, N-1. \quad (2)$$

Обратное дискретное преобразование Хартли:

$$h_i = \frac{1}{N} \sum_{k=0}^{N-1} H_k \cdot \cos\left(\frac{2\pi}{N} ik\right), \quad i = 0, \dots, N-1, \quad (3)$$

где

$$\cos \alpha = \cos \alpha + \sin \alpha. \quad (4)$$

Функция прогнозирования строится как сумма тренда и основных гармонических составляющих сезонной компоненты с соответствующими коэффициентами. Затем производится расчет среднеквадратичного отклонения сглаженной исходной функции от функции прогнозирования, на основе которого определяется качество прогноза и вычисляется максимально допустимое отклонение фактического поведения от ожидаемого.

Модуль реагирования на аномальное поведение программ анализирует поступающие сообщения и может выполнять ряд функций, таких как оповещение администратора системы о произошедшем событии, отключение источника аномалии от остальной сети, с целью снизить возможное негативное влияние на другие компоненты наблюдаемой автоматизированной системы (АС), запись события в журнал регистрации для дальнейшего анализа. Между тем, сами факты обнаружения

аномалий являются событиями в наблюдаемой АС, которые должны быть подвергнуты анализу, как сигнатурному, так и статистическому. Примером подобного события может быть массовое распространение вируса внутри сетевого сегмента АС, реакцией на подобное происшествие может быть полное отключение сегмента до выяснения и устранения причин, вызвавших распространение зловредного кода, с целью ограничить распространение зловредного кода на другие сегменты сети наблюдаемой автоматизированной системы.

#### 4. Перспективы

В рамках данного проекта в перспективе планируется:

- создание адаптивной модели поведения процесса, позволяющей учитывать новые формы поведения, не являющиеся аномальными для данного процесса и типа ресурса; которые могут возникать в результате переконфигурирования процесса или какого-либо изменения в режиме функционирования;
- снижение необходимости человеческого вмешательства в работу системы, для принятия решений о предпринимаемых действиях при обнаружении аномального поведения; вплоть до полной автоматизации процесса реагирования на какие-либо аномальные отклонения показателей в работе системы [8].

#### Выводы

Данная разработка является инструментом, позволяющим повысить отказоустойчивость и облегчить администрирование автоматизированных систем. Следствием этого является снижение затрат на обслуживание АС в целом. Кроме того, кратковременный сбой в работе большинства существующих информационных систем может привести к потере важной информации, значительному экономическому ущербу, уменьшению количества клиентов и т.п. Поэтому, любое увеличение отказоустойчивости системы способно снизить потери, возникающие в результате сбоев АС [3].

#### Литература

1. Щеглов, А.Ю. *Защита компьютерной информации от несанкционированного доступа* / А.Ю. Щеглов. – СПб.: Наука и Техника, 2004. – 384 с.
2. Lucas, Michael W. *Absolute OpenBSD. UNIX for the Practical Paranoid* / Michael W. Lucas. – No Starch Press. ISBN 1-886411-99-9, 2003. – 528 p.
3. Garfinkel, Schwartz, Spafford. *Practical Unix & Internet Security* / Garfinkel, Schwartz, Spafford. – O'Reilly. ISBN 0-596-00323-4, 2003. – 984 p.
4. Russell. *Stealing the Network: How to Own the Box* / Russell, Dubrawsky, FX, Grand, Mullen. – Syngress Publishing. ISBN 1-931836-87-6, 2003. – 330 p.



5. Foster, James C. *Programmer's Ultimate Security DeskRef* / James C. Foster, Steven C. Foster. – Syngress. ISBN 1-932266-72-0, 2004. – 700 p.

6. Hoglund, Greg. *Exploiting Software. How to Break Code* / Greg Hoglund, Gary McGraw. – Addison Wesley. ISBN 0-201-78695-8, 2004. – 512 p.

7. Мониторинг состояния автоматизированной системы и обеспечение стабильности / А.А. Сапожников, А.В. Жуков, М.Л. Карманов, П.В. Збицкий // Сборник материалов Всероссийского конкурса инновационных проектов аспирантов и студентов по приоритетному направлению развития науки и техники «Информационно-

телекоммуникационные системы». – М.: ГНИИ ИТТ «Информика», 2005. – С. 123–124.

8. Сапожников, А.А. Нейронные сети для анализа журналов регистрации / А.А. Сапожников // Безопасность информационного пространства: материалы всерос. науч.-практ. конф. – Екатеринбург: УГТУ-УПИ, 2005. – С. 21–22.

9. Сапожников, А.А. Практика централизованного мониторинга сетей / А.А. Сапожников // Международная конференция «Проблемы функционирования информационных сетей»: материалы конф. – Новосибирск: ЗАО РИЦ Прайс Курьер, 2006. – С. 257–260.

*Поступила в редакцию 10 апреля 2009 г.*

## ИМИТАЦИОННОЕ МОДЕЛИРОВАНИЕ ТЕХНИЧЕСКОГО ОБСЛУЖИВАНИЯ ОСНОВНЫХ ПРОИЗВОДСТВЕННЫХ ФОНДОВ

*А.Н. Шурыгин*

## SIMULATION OF MAINTENANCE OF FIXED ASSETS

*A.N. Shurygin*

Представлена математическая модель, оптимизирующая составление графика технического обслуживания основных производственных фондов предприятия в период перехода с методологии планирования технического обслуживания «по регламенту» на методологию «по фактическому состоянию».

*Ключевые слова: обслуживание производственных фондов, имитационное моделирование.*

A mathematical model for optimizing the scheduling of maintenance of fixed assets of the company during the transition from planning methodologies maintenance «on Rules» to methodology «of the actual state».

*Keywords: maintenance of productive assets, simulation.*

### **Введение**

Одной из наиболее актуальных задач в области управления основными производственными фондами является планирование технического обслуживания и ремонтов (далее – ТОиР) оборудования исходя из его реального технического состояния.

Методология ТОиР по фактическому техническому состоянию направлена на предотвращение отказов оборудования и предполагает постоянный мониторинг его технического состояния, выявление имеющихся или развивающихся дефектов и определение оптимальных сроков проведения ремонтных работ. Эта методология основана на том, что существует взаимосвязь между возможными неисправностями конструктивных элементов оборудования и соответствующими техническими показателями. Следовательно, проводя мониторинг различных показателей, характеризующих работу оборудования, можно вовремя обнаружить изменение его технического состояния и проводить техническое обслуживание только тогда, когда существует реальная угроза выхода этих показателей за допустимые пределы.

Практическое применение данной методологии предполагает выполнение следующих условий:

1. Глубокое понимание технического устройства оборудования: для каждого экземпляра оборудования определены:

– набор показателей, каждый из которых опре-

деляет техническое состояние конструктивного элемента экземпляра оборудования, а весь набор полностью отражает состояние этого экземпляра. При этом для каждого показателя определен некоторый критический уровень, достижение которого означает высокую вероятность поломки оборудования;

– перечень неисправностей (поломок). Вид неисправности определяется значениями набора технических показателей для этого экземпляра оборудования. Трудозатраты и расходные материалы для устранения неисправностей известны.

2. Для каждого экземпляра оборудования производится мониторинг его технических показателей в реальном времени (как правило, решается использованием АСУТП).

3. Ведется статистический учет значений технических показателей и история ТОиР каждого экземпляра оборудования. Данная информация служит основанием для пересмотра критических значений показателей и нормативов технического обслуживания и ремонтов.

Для большинства российских промышленных предприятий характерны следующие проблемы:

1. Отсутствие полного перечня оборудования и единой структуры производственных фондов.

2. Отсутствие в полном объеме необходимой технической документации на оборудование.

3. Отсутствие нормативов работ по ремонту оборудования в части трудозатрат и расхода материалов.

4. Отсутствие исторических данных по эксплуатации и ремонту оборудования, в том числе и о наработке на отказ каждого экземпляра оборудования.

### **1. Постановка задачи планирования ТОиР оборудования**

Таким образом, переход к планированию ТОиР оборудования по фактическому состоянию осложняется тем, что на предприятии зачастую отсутствует необходимая для этой методологии информация. Необходимо построить математическую модель планирования ТОиР оборудования в условиях, когда нет достоверной информации по наработке на отказ каждого экземпляра оборудования.

Рассмотрим промышленное предприятие, производящее несколько наименований продукции, для каждого из которых определены технологические процессы, имеющие несколько стадий передела, на которых, в свою очередь, выполняется определенный вид элементарных технологических операций. Каждый технологический процесс характеризуется пропускной способностью, которая равна минимальной пропускной способности входящих в технологический процесс переделов.

На предприятии имеется несколько групп оборудования, объединенных по функциональному признаку. Каждый из экземпляров оборудования, входящих в определенную группу, выполняет одинаковый набор элементарных технологических операций, при этом производительность экземпляров оборудования может быть неодинаковой. Для каждого экземпляра оборудования определен набор технических показателей, отражающих состояние этого экземпляра. Сотрудниками ремонтного подразделения предприятия на основе данных технической документации производителей оборудования и частично документированного опыта эксплуатации и ремонта подготовлены для каждого экземпляра оборудования: 1) перечень неисправностей и нормы трудозатрат и расходных материалов на их устранение; 2) зависимости состояния каждого конструктивного элемента оборудования от его состояния в прошлом и от состояния других конструктивных элементов. Мониторинг технических показателей производится АСУТП в реальном времени. Исторические данные мониторинга технических показателей, а также достоверная информация о наработке на отказ каждого экземпляра оборудования отсутствуют.

Предполагается, что минимальный спрос на каждое наименование продукции предприятия точно установлен. В таком случае можно определить минимальную пропускную способность каждого передела для всех технологических процессов. Пропускная способность любого передела равна либо превышает установленную минимальную пропускную способность. Однако любой экземпляр оборудования может выйти из строя с некоторой вероятностью (зависит от текущего

технического состояния, наработки, даты последних ремонтных работ), что может повлечь за собой недостаточность пропускной способности передела.

Установлено, что реальный спрос на некоторые наименования продукции предприятия превышает минимальный. Однако возможный их выпуск всегда ограничен максимальной пропускной способностью технического процесса.

Для предотвращения выхода оборудования из строя проводится его техническое обслуживание, стоимость и время которого установлены с достаточной небольшой погрешностью. Предполагается, что затраты на ТОиР всех экземпляров оборудования, вероятность отказа которых достаточна, превышают бюджет на эти цели. В таком случае сначала проводится ТОиР оборудования, вероятность отказа которого достигает критической.

Требуется определить, на обслуживание каких экземпляров оборудования выделить средства, при условии минимизации возможной упущенной выгоды и обеспечения минимального требуемого выпуска каждого наименования продукции.

### **2. Построение математической модели**

Целевая функция имеет следующий вид:

$$\begin{cases} \sum_{n=1}^G \pi_n(Y_n) \rightarrow \max_{Y_n, n=1, G}; \\ Y_n \geq Y_n^*, n = \overline{1, G}; \\ \sum_{n=1}^G W_n(Y_n) \leq b, \end{cases} \quad (1)$$

где  $\pi_n(Y_n)$  – известная функция прибыли от выпуска  $n$ -го наименования продукции,  $n = \overline{1, G}$ ;  $Y_n$  – выпуск  $n$ -го наименования продукции,  $n = \overline{1, G}$ ;  $Y_n^*$  – минимальный необходимый выпуск  $n$ -го наименования продукции,  $n = \overline{1, G}$ ;  $G$  – количество наименований продукции (количество технологических процессов);  $W_n(Y_n)$  – минимальные средние затраты на ремонт оборудования за один период для  $n$ -го процесса,  $n = \overline{1, G}$ ;  $B$  – бюджет на ремонт оборудования;  $b$  – средний бюджет на ремонт оборудования за один период (под периодом понимается некоторый неделимый в рамках модели квант времени).

Целевая функция имеет вид задачи нелинейного программирования, и задача (1) решается методом множителей Лагранжа. Однако, вначале необходимо для каждого технологического процесса определить зависимость минимальных средних затрат на ремонт оборудования от выпуска продукции  $W_n(Y_n)$ . Эта зависимость определяется с помощью имитационной модели.

Максимизация прибыли будет достигаться за счет максимальных продаж наиболее прибыльных наименований продукции, спрос на которые пре-

выпадает минимальный необходимый. Следовательно, необходима такая пропускная способность технологических процессов, выпускающих наиболее прибыльные товары, которая соответствует реальному спросу. Это требует определенных затрат на ремонт соответствующего оборудования. Учитывая ограниченность бюджета на ТОиР, минимальная необходимая пропускная способность каждого передела должна достигаться с минимальными затратами на техническое обслуживание.

Рассмотрим  $n$ -й технологический процесс производства продукции. Обозначим:

- $K$  – количество переделов технологического процесса;
- $V_j$  – количество экземпляров оборудования (станков) на  $j$ -м переделе,  $j = \overline{1, K}$ ;
- $T_j^*(Y^*)$  – минимальная необходимая пропускная способность  $j$ -го передела (определяется исходя из минимального необходимого выпуска продукции  $Y^*$ ),  $j = \overline{1, K}$ .

Для  $i$ -го станка  $j$ -го передела существует  $q_{ij}$  технических показателей, полностью определяющих его состояние, т.е. состояние станка определяется набором:

$${}^{ij}L = \left\{ {}^{ij}L_1, {}^{ij}L_2, \dots, {}^{ij}L_{q_{ij}} \right\}, \quad (2)$$

где  $i = \overline{1, V_j}$ ,  $j = \overline{1, K}$ .

Состояние конструктивного элемента станка определяется исходя из динамики изменения его состояния и состояний смежных конструктивных элементов в прошлые периоды. Пусть количество рассматриваемых периодов  $I$ . Это значит, что  $L_i^t$ ,  $i = \overline{1, q}$  зависит от  $L_j^{t-1}$ ,  $j = \overline{1, q}$ , и не зависит от  $L_j^{t-I-1}$ ,  $j = \overline{1, q}$ .

Зависимость значения технического показателя от его значений и значений смежных показателей в прошлые периоды запишем в следующем виде:

$${}^{ij}L_k^t = f^{ij} \left( L_1^t, {}^{ij}L_1^{t-1}, \dots, {}^{ij}L_1^{t-I}, \dots, {}^{ij}L_{q_{ij}}^t, {}^{ij}L_{q_{ij}}^{t-1}, \dots, {}^{ij}L_{q_{ij}}^{t-I} \right). \quad (3)$$

Найти данную функцию для каждого экземпляра оборудования можно с использованием:

- 1) регрессионных моделей – при наличии достаточных статистических данных;
- 2) экспертных оценок – в противном случае.

Рассмотрим набор технических показателей как многомерную случайную величину:

$${}^{ij}L = \left\{ {}^{ij}L_1, \dots, {}^{ij}L_{q_{ij}} \right\}, \quad (4)$$

где  $i = \overline{1, V_j}$ ,  $j = \overline{1, K}$ .

Для каждого технического показателя определен критический уровень, сигнализирующий о неисправности соответствующего конструктивно-

го элемента станка. Набор критических значений технических показателей выглядит следующим образом:

$${}^{ij}L^{kp} = \left\{ {}^{ij}L_1^{kp}, \dots, {}^{ij}L_{q_{ij}}^{kp} \right\}, \quad (5)$$

где  $i = \overline{1, V_j}$ ,  $j = \overline{1, K}$ .

Для  $L$  определим  $q$ -мерную функцию распределения:

$${}^{ij}F_L \left( {}^{ij}L_1^{kp}, \dots, {}^{ij}L_{q_{ij}}^{kp} \right) = P \left\{ {}^{ij}L_1 < {}^{ij}L_1^{kp}, \dots, {}^{ij}L_{q_{ij}} < {}^{ij}L_{q_{ij}}^{kp} \right\}, \quad (6)$$

где  $i = \overline{1, V_j}$ ,  $j = \overline{1, K}$ .

Эта величина отражает вероятность того, что значения всех технических показателей не выйдут за пределы критического уровня.

Необходимо определить зависимость

$${}^{ij}F_L \left( {}^{ij}L_1^{kp}, \dots, {}^{ij}L_q^{kp} \right), \quad i = \overline{1, V_j}, \quad j = \overline{1, K}. \quad (7)$$

При наличии достаточной и достоверной статистики это можно сделать обычными регрессионными моделями бинарного выбора, такими как probit- и logit-модели.

Так как нам известен характер изменения значений каждого технического показателя, можно найти функцию распределения каждого показателя:

$${}^{ij}F_{L_k} \left( {}^{ij}L_k^{kp}, \forall k, k = \overline{1, q_{ij}} \right), \quad (8)$$

где  $i = \overline{1, V_j}$ ,  $j = \overline{1, K}$ .

Поэтому, при отсутствии достаточных статистических данных можно воспользоваться экспертными оценками. К примеру, можно представить  $F_L(L^{kp})$  в следующем виде:

$${}^{ij}F_L(ijL^{kp}) = \sum_{m=1}^{q_{ij}} {}^{ij}l_m \times {}^{ij}F_{L_m}({}^{ij}L_m^{kp}), \quad (9)$$

где  $l_m$  – вес  $m$ -го технического показателя в общем

состоянии станка,  $\sum_{m=1}^{q_{ij}} {}^{ij}l_m = 1$ ;  $i = \overline{1, V_j}$ ;  $j = \overline{1, K}$ .

Введем обозначения:

– стоимость ремонта (определена для каждой неисправности каждого экземпляра оборудования):

$${}^{ij}Z = {}^{ij}Z({}^{ij}L), \quad i = \overline{1, V_j}, \quad j = \overline{1, K}; \quad (10)$$

– время ремонта (определено для каждой неисправности каждого экземпляра оборудования):

$${}^{ij}T = {}^{ij}T({}^{ij}L), \quad i = \overline{1, V_j}, \quad j = \overline{1, K}; \quad (11)$$

–  $P_{ij}^{kp}$  – критическая вероятность поломки станка – такая вероятность, при которой немедленно начинаем ремонт станка;

–  $M_{ij}$  – производительность  $i$ -го станка  $j$ -го передела согласно технической документации;

–  $F_{ij}$  – реальная производительность  $i$ -го станка  $j$ -го передела;

–  $S_j$  – количество полуфабрикатов  $j$ -го передела, находящееся на складе;

–  $A_j$  – размер склада полуфабрикатов для  $j$ -го передела.

Согласно постановке задачи требуется обеспечить все  $T_j^*(Y^*)$  с минимальными затратами на ремонт оборудования:

$$\sum_{i=1}^{V_j} F_{ij} + S_j \geq T_j^*, j = \overline{1, K}. \quad (12)$$

### 3. Описание имитационной модели

#### 3.1. Входные данные модели

1. Данные по каждому станку в соответствии с формулами (2)–(10).

2. Начальные  $P_{ij}^{kp}$  для каждого станка. Если нет практических соображений о начальном уровне  $P_{ij}^{kp}$ , устанавливается 0. Кроме того, указывается шаг, на который будет увеличиваться  $P_{ij}^{kp}$  (например, 0,01).

3. Минимальный необходимый объем выпуска продукции  $Y^*$ . Исходя из него, определяется минимальная необходимая пропускная способность  $T_j^*(Y^*)$ .

4. Размер склада полуфабрикатов для  $j$ -го передела  $A_j$ . Если возможность складирования отсутствует, устанавливается 0.

5. Количество периодов моделирования  $N$  (при увеличении количества периодов достоверность расчетов повышается).

#### 3.2. Алгоритм выполнения

Описанный ниже алгоритм выполняется отдельно для каждого этапа.

1. Рассчитываются показатели таблицы при заданных минимальных вероятностях  $P_{ij}^{kp}$  для каждого станка. Если неравенство (12) выполняется, то для обеспечения минимального необходимого выпуска продукции ремонтов оборудования не потребуется. Иначе переходим к п. 2. При этом,

если  $\sum_{i=1}^{V_j} F_{ij}^t > T_j^*$ , то  $\sum_{i=1}^{V_j} F_{ij}^t - T_j^*$  полуфабрикатов

поступает на склад.

2. Так как на каждом этапе обработка производится на нескольких станках, часть из которых требует ремонта в данном периоде, в общем слу-

чае существует несколько вариантов распределения усилий и бюджетных средств на ремонт оборудования для обеспечения минимального необходимого выпуска продукции  $Y^*$ . Необходимо выбрать наименее затратный из них. Для этого для каждого станка  $i = \overline{1, V_j}$  поочередно выполняются следующие действия:

2.1.  $P_{ij}^{kp}$  увеличивается на 1 шаг.

2.2. Рассчитываются все показатели таблицы для  $t = \overline{0, N}$ . Если в каком-либо периоде условие неравенства (10) не выполняется, то такой вариант распределения исключается из дальнейших расчетов.

2.3.  $P_{ij}^{kp}$  уменьшается на 1 шаг.

3. Если по результатам п. 2 все варианты были исключены, то  $P_{ij}^{kp}$  одного или нескольких станков увеличивается определенным образом (существует несколько вариантов увеличения  $P_{ij}^{kp}$ : одновременно для всех станков на один шаг, только у самого производительного станка на 2 шага и т.д. Нужный вариант выбирается, исходя из требуемой точности расчетов и объемов вычислений). После этого п. 2 повторяется.

4. Если по результатам п. 2 осталось более одного варианта, то выбирается тот, по которому стоимость ремонта оборудования наименьшая. Таким образом, мы получим минимальные затраты на  $j$ -й этап. Разделив их на количество периодов ( $N$ ), получим минимальные средние затраты этапа за период, позволяющие обеспечить  $T_j^*(Y^*)$ . Обозначим их  $E_j^*(T_j^*)$ .

#### 3.3. Результат применения модели

1. Суммарные минимальные затраты на один период по технологическому процессу

$$W^*(Y^*) = \sum_{j=1}^k E_j(T_j^*). \quad (13)$$

2. Все  $P_{ij}^{kp}$  для конкретного технологического процесса.

#### Заключение

Представленная математическая модель может применяться для планирования ТОиР на про-

Показатели для расчета

Станок	1	2	i	$V_i$
Вероятность поломки в следующем периоде			$P_{ij}^t$	
Стоимость требуемого ремонта			$Z_{ij}^t$	
Время требуемого ремонта			$T_{ij}^t$	
Производительность по паспорту			$M_{ij}$	
Запас на складе на начало периода			$S_j^t$	
Состояние (логический)			Работает / Не работает	
Реальная производительность			$F_{ij}^{t=}$ $\begin{cases} M_{ij}, & \text{не в ремонте} \\ 0, & \text{в ремонте} \end{cases}$	

мышленных предприятиях с учетом следующих ограничений:

1. При недостаточности пропускной способности передела какого-либо технологического процесса не учитывается возможность использования простаивающего оборудования аналогичного назначения, относящегося к другому технологическому процессу.

2. Данная модель предназначена для планирования ТОиР оборудования на малых и средних предприятиях с относительно небольшим парком оборудования. Данное ограничение связано с большим количеством производимых вычислений.

3. Использование данной модели имеет смысл только в период перехода предприятия на планирование ТОиР оборудования по состоянию. По мере адаптации к данной методологии и накопления статистических данных по эксплуатации оборудования целесообразно на постоянной основе

применять более производительные (по скорости вычислений) модели.

#### *Литература*

1. Вентцель, Е.С. *Теория вероятностей и ее инженерные приложения: учеб. пособие для вузов* / Е.С. Вентцель, Л.А. Овчаров. – 2-е изд., стер. – М.: Высш. шк., 2000. – 480 с.

2. Волков, И.К. *Случайные процессы: учеб. для вузов* / И.К. Волков, С.М. Зуев, Г.М. Цветкова; под ред. В.С. Зарубина, А.П. Крищенко. – М.: Изд-во МГТУ им. Н.Э. Баумана, 1999. – 448 с.

3. Кобелев, Н.Б. *Основы имитационного моделирования сложных экономических систем: учеб. пособие* / Н.Б. Кобелев. – М.: Дело, 2003. – 336 с.

4. Самарский, А.А. *Математическое моделирование: Идеи. Методы. Примеры* / А.А. Самарский, А.П. Михайлов. – 2-е изд., испр. – М.: ФИЗМАТЛИТ, 2005. – 320 с.

*Поступила в редакцию 16 марта 2009 г.*

## ОПЕРАТИВНОЕ УПРАВЛЕНИЕ ЭКОНОМИЧНОСТЬЮ ВОДЯНЫХ ТЕПЛОВЫХ СЕТЕЙ НА ОСНОВЕ МАКРОМОДЕЛИРОВАНИЯ

*Е.В. Гойтина, Д.А. Шнайдер*

## OPERATIONAL MANAGEMENT OF EFFICIENT WATER HEATING NETWORKS BASED ON MACROMODELLING

*E.V. Goytina, D.A. Shnayder*

В работе рассматривается актуальная задача разработки методов оперативного анализа экономичности систем теплоснабжения, основанных на построении упрощенных макромоделей тепловых сетей по данным эксплуатации, позволяющих производить расчеты с достаточной для практического использования точностью в режиме реального времени.

*Ключевые слова: системы теплоснабжения, экономичность, модели тепловых сетей.*

We consider the urgent task of developing of the methods for rapid analysis of the cost of heat supply systems based on building a simplified macro model of district heating networks for real data, allowing to perform calculations with sufficient accuracy for practical use in a real time.

*Keywords: heating systems, economy, models of heating networks.*

### Введение

Базовым подходом к теплоснабжению на крупных промышленных предприятиях и в жилищно-коммунальном хозяйстве Российской Федерации является централизованное теплоснабжение, позволяющее снизить затраты на производство тепла за счет совместной выработки электрической и тепловой энергии. Важную роль в эффективном функционировании централизованных систем теплоснабжения играет согласованная работа источников и потребителей при регулировании подачи тепла. Однако одной из основных проблем систем централизованного теплоснабжения в настоящее время является низкая надежность и недостаточная экономичность водяных тепловых сетей. В этой связи актуальным является решение задач оперативного управления, позволяющего осуществлять контроль и управление процессом теплоснабжения в различных режимах, а также своевременно выявлять потери и нерациональное использование тепла, проводить оценку экономичности теплоснабжения.

В настоящее время управление режимами теплоснабжения потребителей промышленных предприятий осуществляется, как правило, с помощью автоматизированных систем диспетчерского управления (АСДУ), которые выполняют функции сбора, накопления и представления на ЭВМ информации о параметрах теплоносителя на источниках, в магистральных трубопроводах и у крупных потребителей. Для повышения надежности и экономичности теплоснабжения целесообразным является также проведение оперативного анализа текущих режимов функционирования тепловых сетей на основе математического моделирования процесса теплоснабжения. Целью такого анализа является выявление потребителей и участков сети, характеризующихся повышенными потерями (утечками) теплоносителя и нерациональным использованием тепла (пониженным теплосъемом).

Современные системы теплоснабжения крупных промышленных предприятий являются сложными системами и содержат большое количество разнородных потребителей, распределенных на

---

Гойтина Екатерина Владимировна – младший научный сотрудник кафедры автоматизации и управления ЮУрГУ; goytina@list.ru

Шнайдер Дмитрий Александрович – канд. техн. наук, доцент кафедры автоматизации и управления ЮУрГУ; shnayder@ait.susu.ac.ru

---

Goytina Ekaterina Vladimirovna – associate research fellow of automation and control department of SUSU; goytina@list.ru

Shnayder Dmitry Alexandrovich – PhD, associate professor of automation and control department of SUSU; shnayder@ait.susu.ac.ru

большой территории. Известные подходы к анализу режимов подобных сложных систем основаны на разработке математической модели реальной системы с помощью специализированного программного обеспечения. Сложность построения такой модели на практике для крупных сетей состоит в необходимости получения большого объема данных, включающих детальные характеристики трубопроводов тепловых сетей (длины, диаметры, коэффициенты шероховатости и т.д.) и потребителей, что требует значительных затрат времени и средств, а в ряде случаев практически не реализуемо.

Для оперативного анализа модель должна учитывать текущие фактические параметры теплоносителя в различных точках системы теплоснабжения и позволять оперативно рассчитывать режимы в случае переключений в тепловых сетях, что накладывает дополнительные требования к сходимости используемых математических методов, объему данных и скорости вычислений, и может быть достигнуто за счет сокращения объема информации, используемой при расчетах, а также повышения эффективности вычислительных алгоритмов.

С учетом сказанного актуальной является задача разработки методов оперативного анализа экономичности системы теплоснабжения, основанных на построении упрощенных макромоделей тепловых сетей по данным эксплуатации, позволяющих производить расчеты с достаточной для

практического использования точностью в режиме реального времени.

При этом с учетом масштабы систем теплоснабжения крупных промышленных предприятий, является актуальной задача разработки подсистемы автоматизированного анализа режимов теплоснабжения (САР-Т).

### 1. Подсистема автоматизированного анализа режимов теплоснабжения

Обобщенная структурная схема САР-Т приведена на рис. 1.

Система теплоснабжения крупного промышленного предприятия, как правило, состоит из нескольких источников теплоснабжения (ТЭЦ, котельные), потребителей и соединяющих их трубопроводов тепловых сетей. Как видно из рис. 1, на нижнем (полевом) уровне САР-Т находятся приборы учета (ПУ) тепла на источниках и у потребителей системы теплоснабжения, включающие в общем случае датчики расхода, температуры и давления теплоносителя. Данные о параметрах источников и крупных потребителей с приборов учета нижнего уровня поступают на сервер базы данных существующей АСДУ. Далее необходимые данные из базы данных АСДУ поступают на сервер базы данных САР-Т.

Данные с теплосчетчиков мелких потребителей, не подключенных к АСДУ, поступают на рабочую станцию оператора, в функции которого

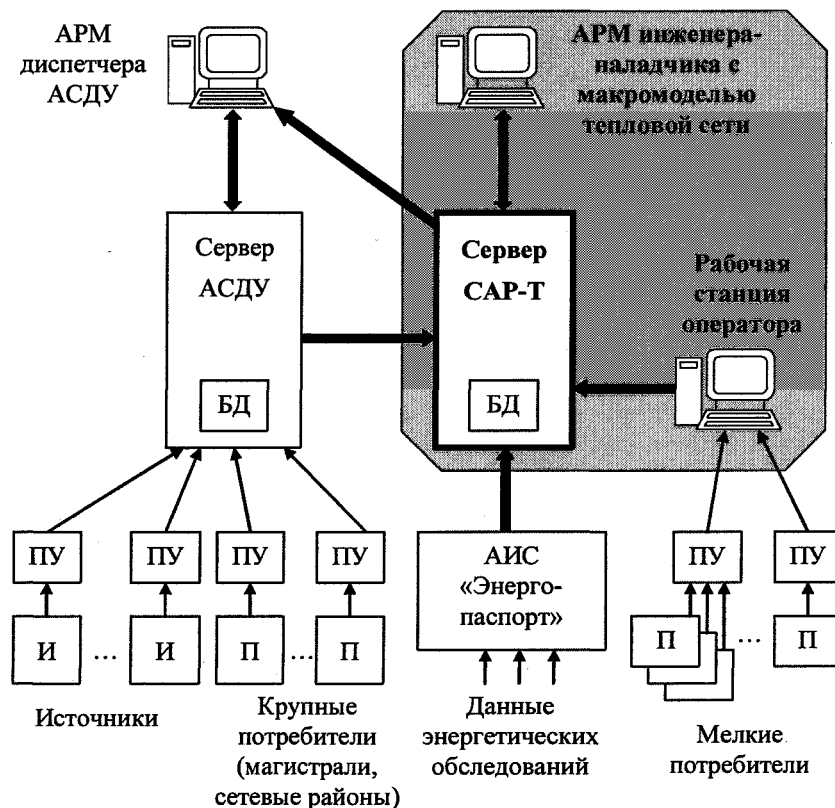


Рис. 1. Структура автоматизированной системы анализа режимов теплоснабжения



входит автоматизированный ввод в базу данных САР-Т суточных архивов теплопотребления, считываемых с теплосчетчиков один раз за отчетный период (обычно месяц).

Кроме того, в базу данных сервера САР-Т поступают данные из автоматизированной информационной системы (АИС) «Энергопаспорт» [1], представляющей собой систему ведения энергопаспортов промышленных потребителей, содержащих технико-эксплуатационные показатели и проектные нагрузки.

Далее все необходимые данные из базы данных САР-Т поступают на автоматизированное рабочее место (АРМ) инженера-наладчика и диспетчера АСДУ. На АРМ инженера-наладчика функционирует специализированное программное обеспечение для оперативного моделирования и анализа режимов функционирования тепловых сетей. Программное обеспечение включает в себя математическую макромодел тепловых сетей, позволяющую проводить анализ режимов при различных вариантах структурных переключений в тепловых сетях в период подготовки проведения плановых ремонтных работ и ликвидации аварийных ситуаций. На основе предварительного моделирования режимов теплоснабжения с учетом фактических параметров теплоносителя и структуры тепловых сетей анализируются различные режимы теплоснабжения с точки зрения безопасности и экономичности, после чего инженер-наладчик принимает решение о реальном использовании того или иного режима. Для настройки параметров модели по запросу инженера-наладчика периодически повторяется процедура идентификации макромоделей на основе текущих и архивных данных эксплуатации, а также экспертных оценок. Необходимость проведения идентификации параметров математической модели (а именно, фактических гидравлических сопротивлений участков тепловой сети по данным о расходах и напорах теплоносителя на абонентских вводах потребителей тепла) определяется тем, что рассчитываемые на основании паспортных характеристик значения гидравлических сопротивлений участков тепловых сетей в процессе эксплуатации претерпевают изменения вследствие зарастания внутренней поверхности трубопроводов, существенно отклоняются от проектных данных и фактически являются неизвестными величинами. В связи с этим необходимо периодически повторять процедуру идентификации параметров математических моделей тепловых сетей по данным эксплуатации с целью уточнения значений гидравлических сопротивлений участков трубопроводов, используемых далее для теплогидравлических расчетов тепловых сетей. Результаты идентификации и моделирования режимов хранятся в базе данных САР-Т.

Программное обеспечение инженера-наладчика также позволяет проводить оценку экономичности использования теплоносителя у потребителей.

Экономичность режимов теплоснабжения оценивается с помощью показателей, которые характеризуют использование теплового потенциала сетевой воды.

В качестве показателей экономичности целесообразно использовать следующие характеристики:

1) температурный напор, представляющий собой разность температур теплофикационной воды между подающим ( $T_1$ ) и обратным ( $T_2$ ) трубопроводами:

$$dT = (T_1 - T_2), \text{ } ^\circ\text{C};$$

2) процент утечек (разность расходов в подающем ( $G_1$ ) и обратном ( $G_2$ ) трубопроводах), который с учетом погрешности измерительных приборов  $\delta$  может быть рассчитан по следующей формуле:

$$\Delta = (G_1(1-\delta) - G_2)/G_1 \cdot 100, \%$$

Анализ режимов тепловых сетей может быть проведен на основе приведенных выше показателей экономичности по следующим критериям:

#### 1. Температурный напор.

Если система отопления обеспечивает перепад температур на тепловом вводе потребителей, предусмотренный температурным графиком регулирования теплоснабжения, то это свидетельствует об эффективном использовании тепловой энергии, т.е.

$$dT_{\text{факт}} \geq dT_{\text{норм.}}$$

#### 2. Процент утечек теплоносителя.

Если фактические значения утечек теплоносителя на тепловых вводах потребителей не превышают нормативные утечки, то это свидетельствует о нормальном состоянии тепловых систем потребителей, т.е.

$$\Delta_{\text{факт}} \leq \Delta_{\text{норм.}}$$

Если приведенные выше критерии не выполняются, то необходимо проведение детального энергетического обследования тепловых систем потребителей.

Таким образом, сопоставление фактических значений показателей функционирования тепловых сетей с их нормативными значениями на основе выбранного критерия эффективности позволяет проводить оценку эффективности функционирования тепловых сетей, определять наиболее проблемные участки сетей и потребителей, требующих проведения наладочных работ, и тем самым оперативно влиять на улучшение режимов теплоснабжения.

## 2. Методика анализа режимов тепловых сетей на основе макромоделирования

Решение задачи оперативного анализа экономичности режимов функционирования тепловых сетей осуществляется на основе математического моделирования процесса теплоснабжения. Для эффективного моделирования системы теплоснабжения в реальном времени требуется разработка упрощенной макромоделей тепловой сети, не требующей чрезмерной детализации характери-

стик сети и позволяющей проводить расчет и анализ режимов теплоснабжения с достаточной для практического применения точностью.

Суть предлагаемого подхода состоит в представлении реальной тепловой сети в виде многоуровневой структуры с выделенными сетевыми районами, отдельными крупными потребителями и соединяющими их магистральными тепловыми сетями (рис. 2). В отличие от полной модели, отражающей состояния всех имеющихся межэлементных связей, в макромоделе отображаются состояния значительно меньшего числа межэлементных связей, что соответствует описанию объекта при укрупненном выделении элементов.

В основу построения макромоделе для расчета тепловых сетей могут быть положены следующие принципы:

1. Сетевые районы ( $CP_i$ ) со сложной схемой сетевых соединений, содержащие большое количество относительно маломощных потребителей, рассматриваются как единый эквивалентный потребитель.

2. Потребители ( $\Pi_j$ ), подключенные непосредственно к магистральной сети, рассматриваются как таковые.

3. При необходимости более подробного рассмотрения параметров гидравлических режимов внутри отдельных сетевых районов может быть осуществлен переход на следующий уровень макромоделе, отражающий состояние потребителей ( $\Pi_{i,k}$ ) и вновь выделенных сетевых районов ( $CP_{i,j}$ ) внутри сетевого района  $CP_i$ . При этом объект  $CP_i$ , являющийся потребителем на верхнем уровне

макромоделе, выступает в роли источника для сетевого района  $CP_{i,j}$ .

На основе указанных выше принципов реализуется возможность использования при анализе одной и той же тепловой сети нескольких моделей, различающихся сложностью, точностью и полнотой отображения свойств. Таким образом, в зависимости от целей моделирования может изменяться степень детализации представления и описания модели тепловой сети. При этом следует отметить, что результаты однократно выполненного трудоемкого моделирования отдельного сетевого района могут многократно применяться далее в укрупненной общей макромоделе, что обуславливает общее снижение объема вычислений.

Применение на практике математической макромоделе тепловой сети, описанной выше, в составе автоматизированной системы диспетчерского управления позволяет оперативно анализировать функционирование систем теплоснабжения в различных режимах и повысить надежность и экономичность процесса теплоснабжения.

### 3. Программное обеспечение анализа режимов тепловых сетей на основе макромоделе

Обобщенная структура разработанного программного обеспечения подсистемы автоматизированного анализа режимов теплоснабжения (ПО САР-Т), отражающая потоки информации между программными модулями и базой данных сервера САР-Т, представлена на рис. 3.

ПО САР-Т включает в себя автоматизирован-

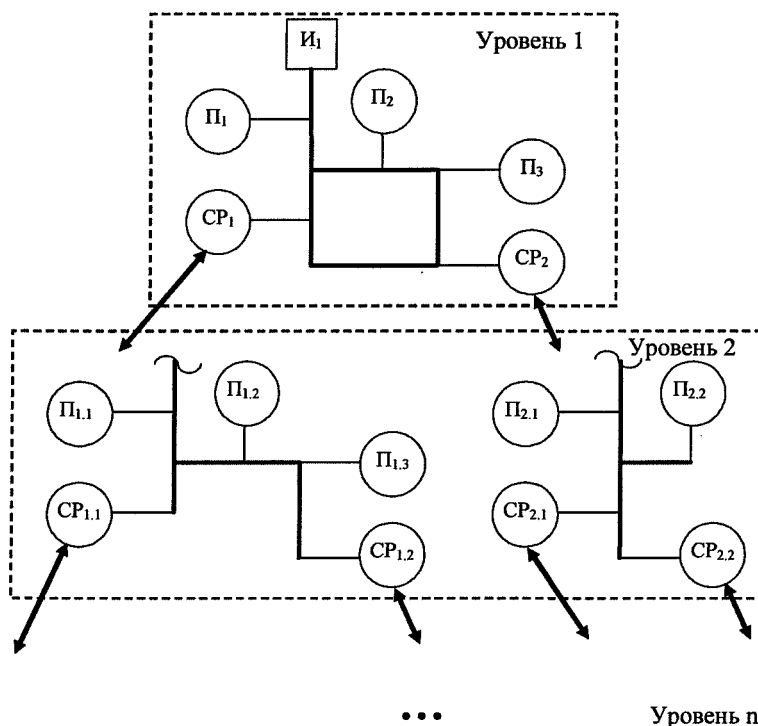


Рис. 2. Многоуровневая структура тепловой сети



**Рис. 3. Структура программного обеспечения SCADA**

ную базу данных, и следующие программные модули:

1. Модуль взаимодействия сервера базы данных SCADA с АСДУ теплоснабжения, позволяющий получать данные о параметрах источников и крупных потребителей с приборов учета тепла.

2. Модуль взаимодействия с АИС «Энергопаспорт» для автоматического ввода данных о технико-эксплуатационных показателях, проектных нагрузках потребителей, параметрах и структуре тепловых сетей.

3. Модуль ввода данных экспертных оценок параметров теплоносителя у потребителей, не оборудованных приборами учета.

4. Модуль автоматизированного ввода данных оператором рабочей станции SCADA суточных архивов теплопотребления с теплосчетчиков мелких потребителей, не подключенных к АСДУ.

5. Модуль идентификации параметров макромоделей тепловых сетей по данным эксплуатации в различных режимах работы тепловой сети [2].

6. Модуль теплогидравлических расчетов тепловых сетей на основе макро моделирования [3]. ПО SCADA включает в себя макро модель системы теплоснабжения и предусматривает возможность моделирования различных режимов теплоснабжения путем изменения параметров модели (напоры источников, сопротивления потребителей), а также изменением структуры модели тепловой сети при помощи открытия/закрытия задвижек на магистральных трубопроводах и внесением дополнительных элементов в структуру тепловой сети.

7. Модуль анализа данных и расчета характеристик энергоэффективности потребления тепловой энергии. На основании параметров, которые хранятся в базе данных на сервере SCADA, проводится расчет характеристик энергоэффективности

теплоснабжения, результаты которого сохраняются модулем в базе данных SCADA.

8. Модуль визуализации результатов моделирования и анализа, формирования отчетов отображает данные эксплуатации, технологические параметры, результаты моделирования и показатели эффективности функционирования, т.е. позволяет осуществлять визуальный контроль качества теплопотребления, отклонений технологических параметров от требуемых значений.

На основе использования разработанного программного обеспечения SCADA был проведен автоматизированный анализ режимов теплоснабжения потребителей промплощадки ОАО «ММК».

В качестве исходных данных для анализа использовались фактические данные приборного учета тепла ЦЭСТ ОАО «ММК» за январь 2007 г. Согласно данным диспетчерской УТЭ ОАО «ММК» средняя температура наружного воздуха в январе 2007 г. составляла  $-3,6^{\circ}\text{C}$ . Согласно утвержденному температурному графику на промплощадке разность температур между подающим и обратным теплоносителем при указанной температуре наружного воздуха должна составлять  $15^{\circ}\text{C}$ . В таблице приведен фрагмент данных по потребителям с температурным напором в январе 2007 г. менее  $15^{\circ}\text{C}$ , полученные автоматизированным способом с помощью разработанного программного обеспечения. Общее число проанализированных потребителей (с установленными приборами учета тепла) составляло 229. Было выявлено 109 потребителей с недостаточным температурным напором.

В результате анализа с помощью разработанного программного обеспечения было выявлено 80 потребителей, у которых фактическая утечка теплоносителя превышает нормативную. В качестве фактических данных в расчетах использовались

значения разности расходов подаваемого и обратного теплоносителя с приборов учета тепла. Нормативные значения утечек рассчитывались по методике, принятой на ОАО «ММК». Было выявлено, что для отдельных потребителей фактическая утечка теплоносителя превышает нормативные значения более чем на 100 %.

Перечень потребителей с недостаточным температурным напором

№ п/п	Потребители	$dT, ^\circ\text{C}$
1	ЦЭСиП. Подстанция № 63	1,2
2	ЖДТ. Ст. «Гранитная». ЦТП	3,0
3	Сортовой цех. Стан 450	3,3
4	ПСЦ. Котельная № 2	4,0
5	ЦЭСиП. Подстанция № 80	4,3
6	ЦЭС. УМИС	4,3
...	...	...

Полученные данные свидетельствуют о существующих резервах энергосбережения в системе теплоснабжения промплощадки ОАО «ММК» и необходимости проведения наладочных и регулировочных работ с использованием разработанного ПО САР-Т. Расчетный экономический эффект от проведения наладочных работ на объектах промплощадки ОАО «ММК» с повышенной циркуляцией и утечками теплоносителя, выявленных с использованием разработанной программы анализа режимов тепловых сетей на основе макро моделирования по данным за январь 2007 г., составляет 470 000 руб. в месяц.

#### Заключение

Таким образом, подсистема автоматизированного анализа режимов теплоснабжения САР-Т обеспечивает выполнение следующих функций:

- ведение информационной базы данных технологических параметров и показателей функционирования системы теплоснабжения;
- идентификация и уточнение параметров макромоделей тепловых сетей;
- макро моделирование режимов теплоснабжения с учетом структурных переключений;
- расчет и графическое отображение результатов моделирования режимов и показателей экономичности теплоснабжения.

Практическое использование подсистемы автоматизированного анализа режимов теплоснабжения, построенной на основе макро моделирования, позволяет оперативно выявлять и устранять нерациональное использование теплоносителя и тем самым повысить надежность и экономичность теплоснабжения потребителей промышленных предприятий.

#### Литература

1. Казаринова, В.Л. Проведение энергетической паспортизации объектов ОАО «ММК» с использованием средств автоматизации / В.Л. Казаринова, Д.А. Шнайдер // Информационные технологии в управлении промышленностью и экономикой субъектов РФ: сб. науч. тр. / под ред. О.В. Логиновского. – Челябинск: Изд-во ЮУрГУ: ЦНТИ, 2003. – С. 50–54.
2. Гойтина, Е.В. Методика идентификации параметров модели тепловой сети по данным эксплуатации / Е.В. Гойтина, Д.А. Шнайдер // Автоматизация и современные технологии. – 2007. – № 9. – С. 20–22.
3. Гойтина, Е.В. Автоматизация расчета гидравлических режимов многоконтурных тепловых сетей / Е.В. Гойтина // XXVI Российская школа по проблемам науки и технологий: краткие сообщения. – Екатеринбург: УрО РАН, 2006. – С. 124–126.

Поступила в редакцию 19 мая 2008 г.

# МОДЕЛИРОВАНИЕ РАБОТЫ ЭЛЕКТРОННОГО ИЗМЕРИТЕЛЯ МЕХАНИЧЕСКОГО МОМЕНТА ЭЛЕКТРИЧЕСКИХ МАШИН В СИСТЕМАХ АВТОМАТИЧЕСКОГО УПРАВЛЕНИЯ

*Г.И. Волович, И.Г. Топольская, Д.В. Топольский*

## SIMULATION OF THE ELECTRONIC MEASURING DEVICE OF THE MECHANICAL MOMENT OF ELECTRIC DRIVES IN AUTOMATIC CONTROL SYSTEMS

*G.I. Volovich, I.I. Topolskaya, D.V. Topolskiy*

В работе рассмотрена возможность определения механического момента с погрешностью 0,5–3 % путем разработки и использования оптимальных алгоритмов учета энергетических преобразований в электрической машине, реализация которых возможна при использовании аналого-цифровой и вычислительной техники.

*Ключевые слова: определение механического момента, измерение механического момента, имитационное моделирование.*

In this paper we consider the possibility of determining the mechanical moment to an accuracy of 0,5–3 % through the development and use of optimal algorithms of energy transformation in electric drives, which is possible with a usage of an analog-digital and computer technology.

*Keywords: definition of the mechanical moment, the measurement of the mechanical moment, simulation.*

### Введение

Многими исследователями предпринимаются попытки использовать электрические машины, входящие в состав различных электротехнических комплексов и систем не только в качестве электропривода, но и в качестве измерителей различных физических величин, информация о которых необходима при функционировании той или иной системы. В работе рассмотрена возможность определения механического момента с погрешностью 0,5–3 % путем разработки и использования оптимальных алгоритмов учета энергетических преобразований в электрической машине, реализация которых возможна при использовании аналого-цифровой и вычислительной техники. Этот метод базируется на контроле за энергетическими преобразованиями, происходящими в электрической машине [2] и ориентирован на применение в этих целях современных измерительных и вычислительных средств.

### 1. Принцип действия измерителя

Использование того или иного метода определения момента во многом определяется спецификой технологического процесса. Предлагаемый измеритель предназначен для более точного определения механического момента электрических машин по их собственным параметрам и независимым переменным. Такой подход для определения механического момента электрических машин в дальнейшем изложении будем называть энергетическим. Концепцию метода определения механического момента можно описать следующими уравнениями, которые следует рассматривать в контексте с решаемой задачей, как средство изложения решения [3]:

$$P_2 = P_1 \pm \sum \Delta P_i; \quad (1)$$

$$P_2 = \omega M; \quad (2)$$

$$M = \frac{P_1 \pm \sum \Delta P_i}{\omega}, \quad (3)$$

**Волович Георгий Иосифович** – д-р техн. наук, профессор кафедры радиотехнических систем ЮУрГУ; gvolovich@mail.ru

**Топольская Ирина Геннадьевна** – ассистент кафедры электротехники ЮУрГУ; tig@susu.ac.ru

**Топольский Дмитрий Валерианович** – канд. техн. наук, доцент кафедры электротехники ЮУрГУ; tig@susu.ac.ru

**Volovich Georgy Iosifovich** – PhD, professor of radio-technical systems department of SUSU; gvolovich@mail.ru

**Topolskaya Irina Iosifovna** – assistant of electric devices department of SUSU; tig@susu.ac.ru

**Topolskiy Dmitry Valerianovich** – PhD, associate professor of electric devices department of SUSU; tig@susu.ac.ru

где  $\omega$  – угловая частота вращения, знак «+» соответствует генераторному режиму работы, знак «-» соответствует работе электрической машины в режиме двигателя.

Уравнение (3) является наиболее удобной математической моделью для физической реализации измерительного устройства. Это объясняется тем, что две наиболее весомых величины  $P_1$  и  $\omega$ , входящие в уравнение (3), можно достаточно точно контролировать относительно простыми техническими средствами. Третья величина  $\sum \Delta P_i$  составляет лишь малую часть общей мощности. Требования к точности определения данной величины менее жесткие, зависят от КПД электрической машины и существенно снижаются с ростом КПД. Следует учитывать, что при определении вращающего момента величина  $\sum \Delta P_i$  включает только те потери, которые имеют место в основном процессе преобразования энергии.

Мощность  $P_1$  электрических машин целесообразно контролировать путем численного интегрирования мгновенной мощности:

$$P_1 = \frac{1}{N} \sum_{k=1}^N (u_{AB}(k) \cdot i_A(k) - u_{BC}(k) \cdot i_C(k)); \quad (4)$$

$$P_1 = \frac{1}{N} \sum_{k=1}^N (u_{\text{я}}(k) \cdot i_{\text{я}}(k)), \quad (5)$$

где  $u_{AB}(k)$ ,  $u_{BC}(k)$ ,  $i_A(k)$ ,  $i_C(k)$  – мгновенные значения линейных напряжений и токов трехфазной машины, соответствующих  $k$ -му измерению их на периоде  $T$  с постоянным шагом  $T/N$ ;  $N$  – количество измерений за период;  $u_{\text{я}}(k)$  и  $i_{\text{я}}(k)$  – мгновенные значения напряжения и тока якоря машины постоянного тока. Аналогично решается задача определения эффективных и средних значений токов и напряжений электрических машин. Например:

$$I_A = \sqrt{\frac{1}{N} \sum_{k=1}^N i_A^2(k)}; \quad (6)$$

$$U_{\text{я}} = \frac{1}{N} \sum_{k=1}^N u_{\text{я}}(k). \quad (7)$$

Интегральная оценка электрических сигналов позволяет учесть энергетические составляющие высших гармоник. Применение цифровой фильтрации повышает помехоустойчивость и быстродействие измерительного устройства.

Суммарные потери мощности машин переменного и постоянного тока представлены в виде суммы трех составляющих:

$$\sum \Delta P_i = \Delta P_{\text{эл}} + \Delta P_{\text{магн}} + \Delta P_{\text{мех}}, \quad (8)$$

где  $\Delta P_{\text{эл}}$  – электрические потери мощности;  $\Delta P_{\text{магн}}$  – потери мощности в магнитопроводе на гистерезис и вихревые токи;  $\Delta P_{\text{мех}}$  – механические потери мощности на трение и вентиляцию.

Применение программного метода контроля потерь мощности позволяет учесть их зависимость от параметров электрической машины:

$$\Delta P_{\text{эл}} = c_1 I^2 (1 + \alpha (t - t_{\text{ном}})) + \Delta U_{\text{щ}} I; \quad (9)$$

$$\Delta P_{\text{магн}} = c_2 U^2 \omega^2; \quad (10)$$

$$\Delta P_{\text{мех}} = \omega (c_3 + c_4 \omega); \quad (11)$$

где  $I$  и  $U$  – линейные ток и напряжение трехфазной машины или ток якоря и напряжение обмотки возбуждения машины постоянного тока;  $t$  и  $t_{\text{ном}}$  – измеряемая температура и номинальная температура электрической машины;  $\alpha$  – температурный коэффициент сопротивления;  $\Delta U_{\text{щ}}$  – суммарное падение напряжения на переходах «щетка–коллектор»;  $c_1 \dots c_4$  – константы, посредством которых устройство адаптируется к конкретному типу электрической машины. В качестве таких констант приняты величины, определяемые традиционным путем проведения опытов холостого хода и короткого замыкания. Ориентировочные значения констант можно определить по основным данным электрических машин. Квадратичная аппроксимация в уравнениях (10) и (11) является компромиссным решением, упрощающим техническую реализацию устройства. Применение более точного описания магнитно-механических потерь мощности не исключено, но может оказаться малоэффективным вследствие отмеченного ранее фактора случайности.

Для машин постоянного тока формула определения вращающего момента имеет вид

$$M = \left( \frac{1}{N} \sum_{k=1}^N (u_{\text{я}}(k) \cdot i_{\text{я}}(k)) \pm \left( c_1 \frac{1}{N} \sum_{k=1}^N i_{\text{я}}^2(k) \times \right. \right. \\ \times \left( 1 + \alpha \left( \frac{1}{N} \sum_{k=1}^N (t(k) - t_{\text{ном}}) \right) + \Delta U_{\text{щ}} \frac{1}{N} \sum_{k=1}^N i_{\text{я}}(k) \right) + \\ \left. \left. + c_2 \frac{1}{N} \sum_{k=1}^N \omega_{\text{я}}^2(k) + \frac{1}{N} \sum_{k=1}^N \omega_{\text{я}}(k) \times \right. \right. \\ \left. \left. \times (c_3 + c_4 \frac{1}{N} \sum_{k=1}^N \omega_{\text{я}}^2(k)) \right) \right) / \frac{1}{N} \sum_{k=1}^N \omega_{\text{я}}(k), \quad (12)$$

где  $u_{\text{я}}(k)$ ,  $i_{\text{я}}(k)$ ,  $\omega_{\text{я}}(k)$ ,  $t(k)$  – мгновенные значения напряжения, тока, скорости вращения и температуры обмотки якоря машины постоянного тока соответственно, соответствующих  $k$ -му измерению их на периоде  $T$  с постоянным шагом  $T/N$ ;  $N$  – количество измерений за период;  $u_{\text{я}}(k)$  – мгновенное значение напряжения обмотки возбуждения;  $t_{\text{ном}}$  – номинальная температура электрической машины;  $\alpha$  – температурный коэффициент сопротивления;  $\Delta U_{\text{щ}}$  – суммарное падение напряжения на переходах «щетка–коллектор»;  $c_1 \dots c_4$  – константы, посредством которых устройство адаптируется к конкретному типу электрической машины.

Для машин переменного тока вращающий момент определяется по формуле

$$\begin{aligned}
 M = & \left( \frac{1}{N} \sum_{k=1}^N (u_{AB}(k) \cdot i_A(k) - u_{BC}(k) \cdot i_C(k)) \pm \right. \\
 & \pm (c_1 \frac{1}{N} \sum_{k=1}^N i_n^2(k) \cdot (1 + \alpha (\frac{1}{N} \sum_{k=1}^N (t(k) - t_{\text{НОМ}}))) + \\
 & + \Delta U_{\text{щ}} \frac{1}{N} \sum_{k=1}^N i_n(k)) + c_2 \frac{1}{N} \sum_{k=1}^N u_n^2(k) \cdot \frac{1}{N} \sum_{k=1}^N \omega_p^2(k) + \\
 & \left. + \frac{1}{N} \sum_{k=1}^N \omega_p(k) \cdot (c_3 + c_4 \frac{1}{N} \sum_{k=1}^N \omega_p^2(k)) \right) / \frac{1}{N} \sum_{k=1}^N \omega_p(k), \quad (13)
 \end{aligned}$$

где  $u_{AB}(k)$ ,  $u_{BC}(k)$ ,  $i_A(k)$ ,  $i_C(k)$  – мгновенные значения линейных напряжений и токов трехфазной машины соответственно;  $u_n(k)$ ,  $i_n(k)$  – усредненные значения мгновенных значений линейных напряжений и токов;  $\omega_p(k)$  – мгновенное значение скорости вращения ротора. Параметры интегрирования и все прочие обозначения в формуле (12) аналогичны рассмотренным в формуле (11). В формулах (11) и (12) в двигательном режиме работы вместо знака «±» следует подставлять знак «-», в генераторном режиме знак «+».

## 2. Описание модели и результаты моделирования

Учитывая, что механический момент является импульсной величиной, исследование динамических характеристик аналого-цифровых измерительных устройств, использующих при измерении механического момента энергетический подход, является важной задачей. Получение сведений такого характера во многом определяет возможность применения для конкретных целей подобных измерителей. Результаты этого исследования

должны обеспечить возможность разработки систем, в состав которых входит подобное измерительное устройство.

Исследование проведено методом динамического моделирования для различных типовых режимов работы электрических машин. На основе анализа результатов динамического моделирования и известных методов исследования систем автоматического управления рассмотрена эквивалентная структурная схема измерителя (рис. 1).

Основной целью проведенного моделирования было сопоставление момента на валу электрической машины с моментом, измеренным аналого-цифровым устройством по разработанному алгоритму (см. рис. 1). Учитывая, что по сравнению с другими типами электрических машин наилучшими качествами с точки зрения регулирования обладают машины постоянного тока, для проверки работы электронного измерителя момента в динамическом режиме была выбрана математическая модель машины постоянного тока.

Собственная динамика измерительных устройств, работающих на вышеупомянутом принципе, была приведена в следующей постановке задачи. Каждый измерительный канал описан дифференциальным уравнением 2 порядка, что соответствует наличию в каждом канале линейного фильтра, где  $T_{11}$ ,  $T_{12}$ ,  $T_{21}$ ,  $T_{22}$ ,  $T_{31}$ ,  $T_{32}$  – постоянные времени фильтров;  $G_{11}$ ,  $G_{12}$ ,  $G_{21}$ ,  $G_{22}$ ,  $G_{31}$ ,  $G_{32}$  – коэффициенты усиления фильтров;  $i_\phi$ ,  $u_\phi$ ,  $\omega_\phi$  – отфильтрованные ток, напряжение и угловая скорость соответственно;  $y_1$ ,  $y_2$ ,  $y_3$  – переменные состояния фильтров в каналах измерения тока, напряжения и угловой скорости якоря соответственно:

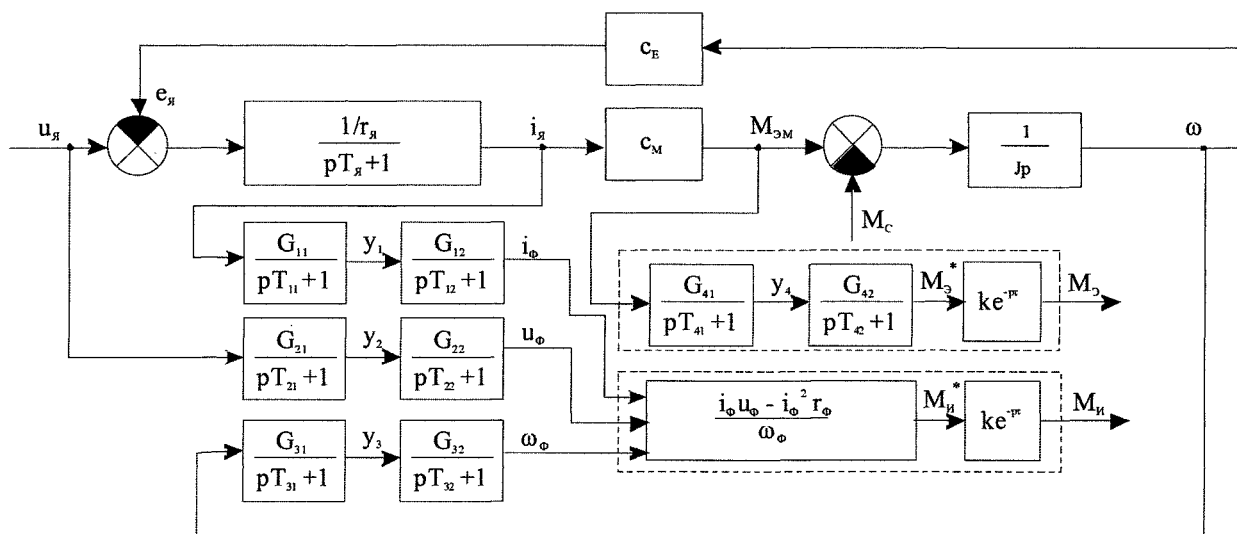


Рис. 1. Математическое описание динамической модели

$$\begin{cases} \frac{dy_1}{dt} = -y_1 \frac{T_{11} + T_{12}}{T_{11}T_{12}} - \frac{i_\phi}{T_{11}T_{12}} + \frac{G_{11}G_{12}}{T_{11}T_{12}} i_a, \\ \frac{di_\phi}{dt} = y_1, \\ \frac{dy_2}{dt} = -y_2 \frac{T_{21} + T_{22}}{T_{21}T_{22}} - \frac{u_\phi}{T_{21}T_{22}} + \frac{G_{21}G_{22}}{T_{21}T_{22}} u_a, \\ \frac{du_\phi}{dt} = y_2, \\ \frac{dy_3}{dt} = -y_3 \frac{T_{31} + T_{32}}{T_{31}T_{32}} - \frac{\omega_\phi}{T_{31}T_{32}} + \frac{G_{31}G_{32}}{T_{31}T_{32}} \omega, \\ \frac{d\omega_\phi}{dt} = y_3. \end{cases}$$

Наличие цифрового вычислительного устройства в измерителе вращающего момента обуславливает дискретный характер его работы. Входными сигналами данного цифрового звена являются ток  $i_\phi$ , напряжение  $u_\phi$  и угловая скорость вращения  $\omega_\phi$  электрической машины. Переменная  $M_n^*$  является результатом вычисления момента по нелинейному алгоритму, без учета дискретности измерителя. На вычисление требуется некоторое время  $\tau$ . Выходным сигналом измерителя является момент  $M_n$ .

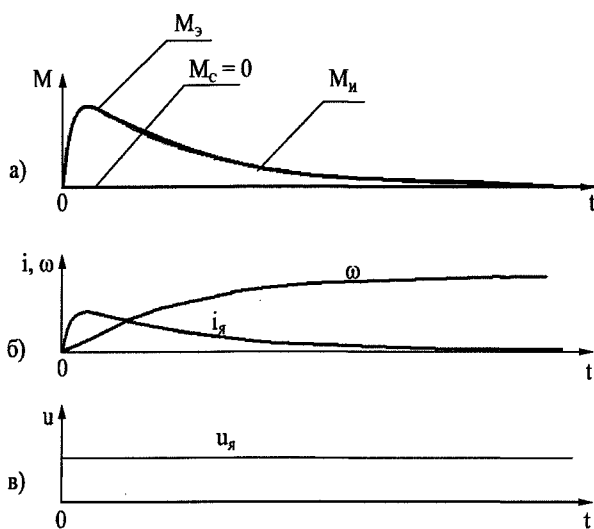


Рис. 2. Реакция системы при пуске двигателя от идеального источника постоянного напряжения

В процессе исследования показания измерителя  $M_n$  сравнивались со значением электромагнитного момента  $M_{эм}$  в различных ситуациях: пуск двигателя от идеального источника постоянного напряжения, управляемого выпрямителя или широтно-импульсного модулятора, а также реакция системы на скачкообразное изменение  $M_c$ . Данные, полученные при пуске двигателя от идеального источника постоянного напряжения, представлены на рис. 2. Рис. 3 демонстрирует измере-

ние  $M_n$  при пуске двигателя при питании от управляемого выпрямителя. На рис. 4 показаны результаты измерения момента при пуске двигателя при питании от широтно-импульсного преобразователя.

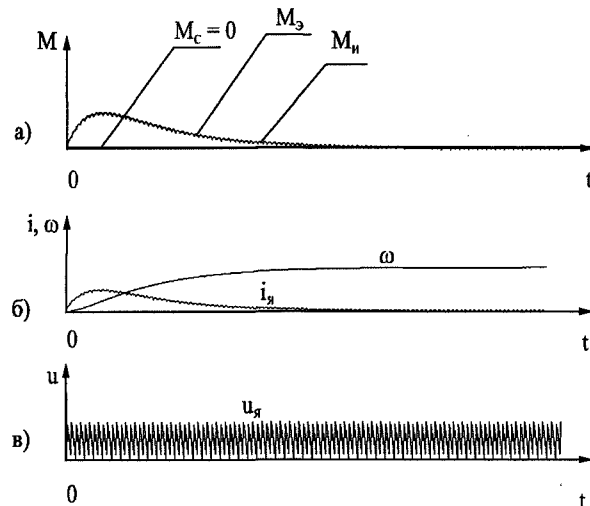


Рис. 3. Реакция системы при пуске двигателя при питании от управляемого выпрямителя

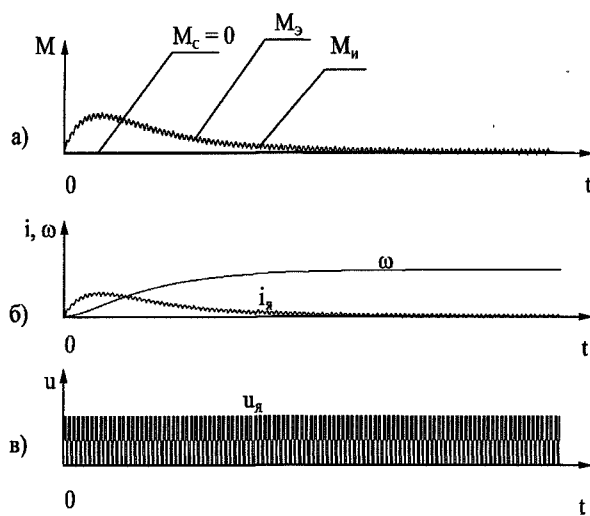


Рис. 4. Реакция системы при пуске двигателя при питании от широтно-импульсного модулятора

Результаты реакции рассматриваемой системы на скачкообразное изменение  $M_c$  приведены на рис. 5. Рис. 5, а соответствует идеальному источнику питания, рис. 5, б демонстрирует случай с питанием от управляемого выпрямителя, а рис. 5, в случай с широтно-импульсным преобразователем.

Эквивалентное значение измеренного момента  $M_n$  получено путем преобразования значения электромагнитного момента  $M_{эм}$ , которое на цифровой модели может быть определено точно.



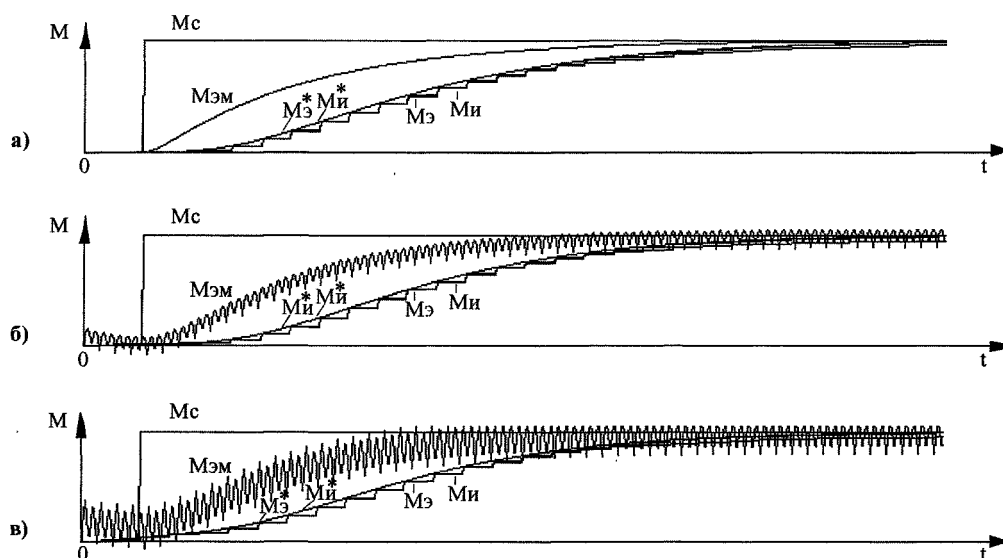


Рис. 5. Сравнение результатов определения вращающего момента по разработанному алгоритму и по эквивалентной функции при скачкообразном изменении нагрузки

### Заключение

В ходе моделирования была исследована динамика цифрового измерителя момента с учетом источника питания. Задача исследования классифицирована как исследование динамики нелинейной цифровой системы, включающей нескольких параллельных каналов для измерения аналоговых сигналов.

Исследование выполнено методом динамического моделирования первого типа [1] на основе численных методов для типовых случаев изменения вращающего момента на валу электрической машины: при изменении момента сопротивления и при пуске. Собственная динамика измерительных устройств, работающих на вышеупомянутом принципе, была приведена в следующей постановке задачи. Каждый измерительный канал описан дифференциальным уравнением 2 порядка, что соответствует наличию в каждом канале линейного фильтра. Динамические сигналы обрабатываются в соответствии с разработанным нелинейным алгоритмом. Значения показаний измерителя сравнивались с электромагнитным моментом. Результаты исследования показали, что измеренный

момент отличаются от электромагнитного в пределах точности, предъявляемой к ряду испытательных стендов (0,5–3 %), а также что скорость изменения измеренного момента соответствует скорости изменения входных параметров измерителя. Кроме того учтена дискретность, характеризующаяся временем, затрачиваемым на вычисления. По результатам проведенного исследования разработана эквивалентная линейная передаточная функция измерителя и разработана методика расчета коэффициентов.

### Литература

1. *Динамическое моделирование и испытание технических систем* / И.Д. Кочубиевский, В.А. Стражмейстер и др.; под ред. И.Д. Кочубиевского. – М.: Энергия, 1978. – 303 с.
2. *Иванов-Смоленский, А.В. Электрические машины* / А.В. Иванов-Смоленский. – М.: Энергия, 1980. – 928 с.
3. *Смолин, В.И. Об одном методе определения вращающего момента электрических машин* / В.И. Смолин, Д.В. Топольский, Н.Н. Гудаев // *Электричество*. – 1999. – № 7.

Поступила в редакцию 25 января 2008 г.

# АЛГОРИТМ ЧИСЛЕННОГО РЕШЕНИЯ ЗАДАЧИ ШОУОЛТЕРА–СИДОРОВА ДЛЯ СИСТЕМ ЛЕОНТЬЕВСКОГО ТИПА

*А.В. Келлер*

## ALGORITHM OF NUMERICAL SOLUTION OF SHOULTER–SIDOROV PROBLEM FOR LEONTIEV-TYPE SYSTEMS

*A.V. Keller*

В статье рассмотрено решение задачи Шоултера–Сидорова для систем леонтьевского типа с необратимым оператором при производной. Рассмотрение начальных данных Шоултера–Сидорова позволяет расширить спектр практического применения модели. Предложен алгоритм ее численного решения, исследована сходимость.

*Ключевые слова:* задача Шоултера–Сидорова, системы леонтьевского типа, исследование сходимости.

The article deals with the solution of Shoulter-Sidorov problem for Leontiev-type systems with irreversible operator at the derivative. Consideration of the initial data of Shoulter-Sidorov enables greater range of practical applications of the model. An algorithm for its numerical solution has been proposed, the convergence has been investigated.

*Keywords:* Shoulter-Sidorov problem, Leontiev-type systems, convergence investigation.

### Введение

В [1, 2] предложен численный алгоритм решения задачи Коши

$$u(0) = u_0 \quad (1)$$

для линейной системы обыкновенных дифференциальных уравнений

$$\dot{u} = Su + g, \quad (2)$$

основанный на идеях теории полугрупп операторов. Здесь  $S$  – квадратная матрица порядка  $n$ . В [3, 4] этот подход был распространен на задачу (1) для вырожденной системы уравнений

$$L\dot{u} = Mu + f \quad (3)$$

с использованием идей теории вырожденных полугрупп операторов [5] (здесь  $L$  и  $M$  – квадратные матрицы, порядка  $n$ , причем  $\det L = 0$ ). Одним из важных случаев системы (3) является хорошо известная система В.В. Леонтьева «затраты-выпуск» с учетом запасов (см. в [6]), поэтому в [3] было предложено такие системы уравнений называть «системами леонтьевского типа».

Простота предложенного в [3, 4] алгоритма обеспечивает высокое качество получаемого программного продукта, что выгодно отличает данный алгоритм от использовавшихся ранее методов Эйлера, Рунге–Кутты, итерационных и других методов [7–9]). Основным недостатком этого алгоритма (как впрочем, и всех остальных)

является принципиальная неразрешимость задачи (1) для системы (3) при произвольных начальных векторах  $u_0 \in \mathbb{R}^n$ . Эта трудность преодолевается, например в [3], где вектор-функция  $f: [0, T] \rightarrow \mathbb{R}^n$  постоянная, введением множества допустимых начальных значений, понимаемых как фазовое пространство системы (3). В [4] при снятии условия постоянства вектор-функции  $f$  налагаются условия согласования  $f$  с начальным значением  $u_0$ . Заметим, что условия согласования в том или ином виде имеют место и во всех других алгоритмах.

Для условий согласования (как и для построения фазового пространства) необходимы проекторы, которые либо выражаются через контурные интегралы от матриц-функций, либо являются пределами матричных последовательностей. Ввиду неустойчивости любого проектора относительно малых возмущений такое вычисление матрицы проектора очень затруднительно. Поэтому, например в [10], при построении системы (3), моделирующей экономику коммунального хозяйства, пришлось ограничиться малыми городами, т.е. такими, где матрицы  $L$  и  $M$  имеют порядок не больше 10. Именно малость порядка матриц  $L$  и  $M$  сделало возможным вычисление проекторов «вручную».

Келлер Алевтина Викторовна – канд. физ.-мат. наук, доцент, зав. кафедрой общеобразовательных дисциплин ЮУрГУ; alevtinak@inbox.ru

Keller Alevtina Viktorovna – PhD, associate professor, head of general disciplines department of SUSU; alevtinak@inbox.ru

Между тем в современной математической литературе существуют попытки теоретического осмысления так называемых «неклассических» задач для системы (3) [11, 12], основным достоинством некоторых является однозначная разрешимость при любых начальных данных  $u_0 \in \mathbb{R}^n$ . Разработка численных алгоритмов решения таких задач позволит избавиться как от трудоемкого построения фазового пространства (и не менее трудоемкой редукции системы (3) к системе (2), заданной на нем), так и от трудоемкой проверки условий согласования. Основная цель данной статьи – построение алгоритма численного решения задачи Шоултера–Сидорова

$$\left[ (\alpha L - M)^{-1} L \right]^p (u(0) - u_0) = 0$$

для системы (3) (числа  $\alpha$  и  $p$  вычисляются на первом шаге алгоритма). Статья кроме введения и списка литературы содержит две части. В первой дается теоретическое обоснование алгоритма, а во второй приведены основные этапы вычислений.

### 1. Задача Шоултера–Сидорова

Пусть  $L$  и  $M$  – квадратные матрицы порядка  $n$ , причем, следуя [13], гл. XII, п. 2, пучок матриц  $\mu L - M$  назовем *регулярным*, если существует число  $\lambda \in \mathbb{C}$  такое, что  $\det(\lambda L - M) \neq 0$ .

Заметим, что условие регулярности пучка матриц эквивалентно условию  $L$ -регулярности матрицы  $M$  [3, 4]. Поэтому, как показано в [5], гл. 4, при условии регулярности пучка существуют единственным образом определяемые матрицы  $H$ ,  $S$ ,

$M_0$ ,  $L_1$ ,  $Q$  порядка  $n$ , такие, что  $L$ -резольвента  $(\mu L - M)^{-1}$  матрицы  $M$  разлагается в ряд Лорана

$$\begin{aligned} (\mu L - M)^{-1} &= \sum_{l=0}^p \mu^l H^l M_0 (I - Q) + \\ &+ \sum_{l=1}^{\infty} \mu^{-l} S^{l-1} L_1 Q \end{aligned} \quad (4)$$

в окрестности бесконечно удаленной точки, причем  $H$  – нильпотентная матрица со степенью нильпотентности  $p$ ,  $Q$  – идемпотентная матрица,  $MM_0$ ,  $M_0M$ ,  $L_1L$  и  $LL_1$  – диагональные матрицы с нулями и единицами на главной диагонали. Поскольку  $\det(\lambda L - M) \neq 0$ , то многочлен  $\det(\lambda L - M) = 0$  имеет не более  $n$  различных нулей, которые расположены в круге радиуса  $a$ , а значит, при  $|\mu| > a$  разложение (4) имеет место. Точка  $\infty$  называется *устранимой особой точкой*  $L$ -резольвенты матрицы  $M$ , если  $p = 0$  в (4); и *полюсом порядка*  $p \in \mathbb{N}$  в противном случае. В дальнейшем, немного отходя от клас-

сического стандарта, будем называть устраняемую особую точку *полюсом порядка нуль*. Итак, пусть пучок  $\mu L - M$  регулярен, и  $\infty$  – полюс порядка  $p \in \{0\} \cup \mathbb{N}$ ; тогда можно выбрать число  $\alpha$  и рассмотреть задачу Шоултера–Сидорова

$$\left[ R_{\alpha}^L(M) \right]^p (u(0) - u_0) = 0, \quad (5)$$

для системы уравнений леонтьевского типа

$$L\dot{u} = Mu + f, \quad (6)$$

где  $R_{\alpha}^L(M) = (\alpha L - M)^{-1} L$  – *правая  $L$ -резольвента матрицы  $M$* , в отличие от ее *левой  $L$ -резольвенты*  $L_{\alpha}^L(M) = L(\alpha L - M)^{-1}$ , а  $f: [0, T] \rightarrow \mathbb{R}^n$  – некоторая вектор-функция.

*Решением системы* (3) называется вектор-функция  $u_0 \in C^1((0; T); \mathbb{R}^n) \cap C([0; T]; \mathbb{R}^n)$ , удовлетворяющая уравнениям системы. Решение системы (6) называется *решением задачи* (5), (6), если оно вдобавок удовлетворяет уравнениям (5). Имеет место [5, гл. 4]

**ТЕОРЕМА 1.1.** Пусть пучок  $\mu L - M$  регулярен,  $p \in \{0\} \cup \mathbb{N}$  – порядок полюса  $L$ -резольвенты матрицы  $M$ , вектор-функция  $f: [0, T] \rightarrow \mathbb{R}^n$  такая, что  $(L_{\alpha}^L(M))^p f \in C([0; T]; \mathbb{R}^n)$ , а

$$I - (L_{\alpha}^L(M))^p f \in C^{p+1}((0; T); \mathbb{R}^n) \cap C^p([0; T]; \mathbb{R}^n).$$

Тогда при любом  $u_0 \in \mathbb{R}^n$  существует единственное решение задачи (5), (6), которое к тому же имеет вид

$$\begin{aligned} u(t) &= - \sum_{q=0}^p H^q M_0^{-1} (I - Q) f^{(q)}(t) + U^t u_0 + \\ &+ \int_0^t R^{t-s} Q f(s) ds. \end{aligned}$$

Здесь

$$U^t = \frac{1}{2\pi i} \int_{\gamma} R_{\mu}^L(M) e^{\mu t} d\mu,$$

$$R^t = \frac{1}{2\pi i} \int_{\gamma} (\mu L - M)^{-1} e^{\mu t} d\mu,$$

$$Q = \frac{1}{2\pi i} \int_{\gamma} L_{\mu}^L(M) d\mu,$$

контур  $\gamma = \{\mu \in \mathbb{C} : |\mu| = r > a\}$ . Контурные интегралы не очень удобны в численных расчетах, поэтому в [3, 4] предложен другой подход, основанный на аппроксимациях типа Уиддера–Поста [5, гл. 2]. Именно справедлива

**ТЕОРЕМА 1.2.** Пусть пучок  $\mu L - M$  регулярен,  $p \in \{0\} \cup \mathbb{N}$  – порядок полюса  $L$ -резольвенты матрицы  $M$ . Тогда

$$\lim_{k \rightarrow \infty} \left[ \left( L - \frac{t}{k(p+1)} M \right)^{-1} L \right]^{k(p+1)} = U^t,$$

$$\lim_{k \rightarrow \infty} \left[ kL_k^L(M) \right]^{(p+1)} = Q,$$

$$\lim_{k \rightarrow \infty} \left[ \left( L - \frac{t}{k(p+1)} M \right)^{-1} L \right]^{k(p+1)-1} \times \left( L - \frac{t}{k(p+1)} M \right)^{-1} = R^t.$$

Теперь пусть пучок  $\mu L - M$  регулярен,  $p \in \{0\} \cup \mathbb{N}$  – порядок полюса  $L$ -резольвенты матрицы  $M$  в точке  $\infty$ . Фиксируем  $T \in \mathbb{R}_+$ ,  $t \in (0, T)$ ,  $k \in \mathbb{N}$  и положим

$$U_k^t = \left[ \left( L - \frac{t}{k(p+1)} M \right)^{-1} L \right]^{k(p+1)},$$

$$Q_k = \left[ kL_k^L(M) \right]^{(p+1)},$$

$$R_k^t = \left[ \left( L - \frac{t}{k(p+1)} M \right)^{-1} L \right]^{k(p+1)-1} \times \left( L - \frac{t}{k(p+1)} M \right)^{-1}.$$

Выберем вектор  $u_0 \in \mathbb{R}^n$ , вектор-функцию  $f \in C^{p+1}((0; T); \mathbb{R}^n) \cap C^p([0; T]; \mathbb{R}^n)$  и построим вектор-функцию

$$u_k(t) = - \sum_{q=0}^p H^q M_0^{-1} (I - Q_k) f^{(q)}(t) + U_k^t u_0 + \int_0^t R_k^{t-s} Q_k f(s) ds.$$

**ТЕОРЕМА 1.3.** Пусть пучок  $\mu L - M$  регулярен,  $p \in \{0\} \cup \mathbb{N}$  – порядок полюса  $L$ -резольвенты матрицы  $M$  в точке  $\infty$ . Тогда существует константа  $C = C(L, M, T) \in \mathbb{R}_+$

такая, что  $\|u(t) - u_k(t)\| \leq \frac{C}{k}$  при всех  $t \in [0; T]$ ,

$u_0 \in \mathbb{R}^n$ ,  $u \in C^{p+1}((0; T); \mathbb{R}^n) \cap C^p([0; T]; \mathbb{R}^n)$ .

Доказательство теоремы основывается на оценках

$$\left\| \left[ kL_k^L(M) \right]^{(p+1)} - Q \right\| \leq \sum_{k=2}^{p+1} \frac{KC_{p+1}^k}{(p+1)\mu^{k-1}\beta^{p+1-k}} \left\| R_k^L(M) \right\|$$

$$\left\| U_k^t - U^t \right\| \leq \frac{(p+1)K^3 t^2}{2\beta^{p-1}k} \left\| \left( (\beta L - M)^{-1} M \right)^2 \right\|$$

взятых из [5, гл. 2], где  $\beta \in \mathbb{R}_+$

## 2. Алгоритм решения

Построение алгоритма начнем с допущения, что  $\det M \neq 0$ . Это допущение не ограничивает общности предыдущих рассуждений. Действительно, при условии регулярности пучка  $\mu L - M$ , можно сделать замену  $u = e^{\lambda t} v$  в уравнении (3) и перейти к уравнению

$$L\dot{v} = (M - \lambda L)v + e^{-\lambda t} f \tag{7}$$

того же вида, что и (3), но  $\det(M - \lambda L) \neq 0$ . Обратный переход от решений системы (7) к решениям системы (3) очевиден.

На первом шаге алгоритма нужно найти числа  $\alpha \in \mathbb{R}$  и  $p \in \{0\} \cup \mathbb{N}$ . Можно разумеется, разложить  $L$ -резольвенту матрицы  $M$  в ряд (4) и тем самым сразу же найти эти числа. Однако, существует другой менее трудоемкий путь. Рассмотрим многочлен

$$\det(\mu L - M) = a_n \mu^n + a_{n-1} \mu^{n-1} + \dots + a_1 \mu + a_0.$$

Поскольку  $a_n = \det L$ , то  $a_n \neq 0$ . Далее, коэффициент  $a_l$  есть сумма слагаемых, каждое из которых есть произведение одного из миноров порядка  $l$  матрицы  $L$  на число,  $l = 1, \dots, n-1$ ,  $a_0 = \det(-M)$ . Поэтому степень многочлена  $\det(\mu L - M)$  не выше  $\text{rank } L$ , т.е. ранга матрицы  $L$ . Итак,

$$\det(\mu L - M) = a_q \mu^q + a_{q-1} \mu^{q-1} + \dots + a_1 \mu + a_0,$$

где  $q = \deg \det(\mu L - M) \leq \text{rank } L$ , поэтому если взять число  $\alpha \in \mathbb{R}$  таким, что

$$|\alpha| > \max \left\{ 1, |a_q|^{-1} \left( \sum_{l=0}^q |a_l| \right) \right\},$$

то  $\det(\alpha L - M) \neq 0$ , и значит, существует матрица  $(\alpha L - M)^{-1}$ . Далее, считая, что матрица  $M$  обратима, представим

$$\det(\mu L - M) = \det M \det(\mu M^{-1} L - I).$$

Зная, что порядок полюса в точке  $\infty$  резольвенты  $(\mu I - M^{-1} L)^{-1}$  равен нулю, легко найти, что порядок полюса  $L$ -резольвенты матрицы  $M$  в точке  $\infty$  равен  $n - q$ . Итак, числа  $\alpha$  и  $p = n - q$  найдены.

Тогда находя значение  $k$ , с которого можно начинать считать приближенные проекторы, получим, что при

$$k > \frac{1}{|a_q|} \sum_{l=q+1}^n |a_l| + 1$$

мы не сможем оказаться даже вблизи точки  $L$ -спектра оператора  $M$ .

Рассмотрим многочлен

$$\det(\mu(p+1)L - tM) = a_q t^q \mu^q (p+1)^q +$$

$$+ a_{q-1} t^{q+1} \mu^{q-1} (p+1)^{q-1} + \dots + a_1 t^{n-1} \mu + t^n a_0,$$

где  $a_q \neq 0, q \leq \text{rank } L$ . Тогда, учитывая  $p = n - q$  при

$$k \left\{ \begin{array}{l} \frac{1}{|a_q|(n-q)^{(n-q)} \sum_{l=q+1}^n |a_l|(n-q+1)^{n-l} + 1, \\ \text{если } |t| < 1 \\ \frac{1}{|a_q||t|^q (n-q)^{(n-q)} \sum_{l=q+1}^n |a_l|(n-q+1)^{n-l} |t|^l + 1, \\ \text{если } |t| \geq 1 \end{array} \right.$$

мы не сможем оказаться даже вблизи точки  $L$ -спектра оператора  $M$ .

Теоретическая оценка сходимости не позволяет сделать вывод о точности предлагаемого алгоритма. Тем не менее, практические результаты показывают, что уже при числе итераций более ста, результаты вычислений дают не менее точные результаты, чем неявная схема Эйлера или метод Рунге–Кутты. Последний факт позволяет надеяться на развитие предлагаемого подхода как в теоретическом, так и в практическом аспектах.

### 3. Пример Леонтьева

Операторы  $L$  и  $M$  зададим матрицами

$$L = \begin{pmatrix} 7/20 & 1/20 & 21/200 \\ 1/100 & 103/200 & 8/25 \\ 0 & 0 & 0 \end{pmatrix},$$

$$M = \begin{pmatrix} 3/4 & -1/5 & -11/20 \\ -7/25 & 10304189 & -70836357 \\ -4/15 & -2/5 & 13/15 \end{pmatrix}.$$

Если переобозначить  $L = B, M = I - A$ , то матрицы  $B$  и  $A$  почти совпадут с матрицами из классического примера [6]. «Почти» означает, что элементы  $m_{22}$  и  $m_{23}$  подобраны специально с целью упростить вычисления и отличаются от приведенных в примере чисел  $22/25$  и  $-3/5$  на величины

$$\begin{aligned} m_{22} &= \frac{22}{25} - \frac{-252291}{11996000}, \\ m_{23} &= \frac{-3}{5} - \frac{1139643}{119960000}. \end{aligned} \quad (8)$$

В.В. Леонтьев рассматривал взаимосвязи между тремя отраслями экономики: сельским хозяйством, промышленностью и домашними хозяйствами. Элемент  $a_{ij}$  матрицы  $A$  означает количество продукции  $i$ -й отрасли, необходимой для производства единицы продукции  $j$ -й отрасли. Элемент  $b_{ij}$  матрицы  $B$  представляет определенный технологический запас особого типа

благ – машин, механических инструментов, промышленных зданий и сооружений, рабочих запасов первичных и промежуточных материалов, производимых отраслью, который используется в отрасли  $j$  для производства единицы ее продукции. Другими словами, каждый столбец матрицы  $B$  описывает потребность некоторой отрасли в физическом капитале (в расчете на единицу ее валового выпуска) таким же образом, как соответствующий столбец матрицы  $A$  описывает ее затраты. Именно поэтому последняя строка матрицы  $B$  содержит только нулевые элементы, так как труд невозможно запастись.

Перейдя к расчету системы (3), найдем  $L$ -спектр оператора  $M$   $\sigma^L(M) = \{0, 2; 2, 7\}$ . Именно для того, чтобы точки  $L$ -спектра оператора  $M$  были рациональными, сделаны поправки (8). Точки  $L$ -спектра оператора  $M$  в исходном примере иррациональны и отличаются от найденных не больше, чем на одну сотую.

Далее по формулам, приведенным в первой части статьи, построим точное и приближенное решение. Приведем точное решение и результаты счета по алгоритму без комментариев, взяв при этом в качестве  $f = (2t, 2t, 2t)$  (табл. 1, 2).

Таким образом, рассмотренный в данной работе алгоритм позволяет численно решать задачу Шоултера–Сидорова с достаточной степенью точности: расхождения в точном и приближенном решении начинаются с тысячных долей.

### Литература

1. Павлов, Б.В. Об одном методе численного интегрирования систем обыкновенных дифференциальных уравнений / Б.В. Павлов, А.Я. Познер // ЖВМиМФ. – 1973. – Т. 13, № 4. – С. 1056–1059.
2. Павлов, Б.В. Численное решение систем линейных обыкновенных дифференциальных уравнений с постоянными коэффициентами / Б.В. Павлов, О.Е. Радионова // ЖВМиМФ. – 1994. – Т. 34, № 4. – С. 622–627.
3. Свиридюк, Г.А. Численное решение систем уравнений леонтьевского типа / Г.А. Свиридюк, С.В. Брычев // Изв. вузов. Математика. – 2003. – № 8. – С. 46–52.
4. Свиридюк, Г.А. Алгоритм решения задачи Коши для вырожденных линейных систем обыкновенных дифференциальных уравнений с постоянными коэффициентами / Г.А. Свиридюк, И.В. Бурлачко // ЖВМиМФ. – 2003. – Т. 43, № 11. – С. 1677–1683.
5. Sviridyuk, G.A. Linear Sobolev Type Equations and Degenerate Semi-groups of Operators / G.A. Sviridyuk, V.E. Fedorov. – Utrecht-Boston-Köln-Tokyo: VSP, 2003.
6. Леонтьев, В.В. Межотраслевая экономика / В.В. Леонтьев. – М.: Экономика, 1997.

Таблица 1

## Точное решение задачи Шоултера–Сидорова

$t$	$x_1$	$x_2$	$x_3$
0	1	1	0,7692307692
1/12	1,079594653	1,115125754	0,6545486955
1/6	1,211221782	1,260474368	0,5698255794
1/4	1,399261210	1,434354380	0,5156284632
1/3	1,649051640	1,634491393	0,4925503470
5/12	1,967129067	1,857877706	0,5012141312
1/2	2,361526596	2,100584518	0,5422779145
7/12	2,842149925	2,357527831	0,6164435987
2/3	3,421246553	2,622172543	0,724463383
3/4	4,113994781	2,886165457	0,867151766
5/6	4,939240310	3,138871569	1,045399350
11/12	5,920418738	3,366797082	1,260189034
1	7,086710966	3,552864395	1,512617818

Таблица 2

## Приближенное решение задачи Шоултера–Сидорова по алгоритму

$t$	$x_1$	$x_2$	$x_3$
0	1	1	0,7692307692
1/12	1,0795957679	1,1151265428	0,6545494099
1/6	1,2112238218	1,260475827	0,5698269423
1/4	1,3992644345	1,4343566045	0,5156305665
1/3	1,6490560881	1,6344943024	0,4925530898
5/12	1,9671347889	1,8578811956	0,5012174099
1/2	2,36153371	2,1005887151	0,5422821007
7/12	2,8421584126	2,3575325354	0,6164483714
2/3	3,4212567008	2,622177848	0,7244687609
3/4	4,1140069697	2,8861709037	0,8671579463
5/6	4,9392549984	3,1388770098	1,0454063117
11/12	5,9204365728	3,3668021207	1,2601968472
1	7,0867328788	3,552868403	1,5126263028

7. Бояринцев, Ю.Е. *Линейные и нелинейные алгебро-дифференциальные системы* / Ю.Е. Бояринцев. – Новосибирск: Наука, 2000.

8. Чистяков, В.Ф. *Избранные главы теории алгебро-дифференциальных систем* / В.Ф. Чистяков, А.А. Щеглова. – Новосибирск: Наука, 2003.

9. Бояринцев, Ю.Е. *Пучки матриц и алгебро-дифференциальные системы* / Ю.Е. Бояринцев, И.В. Орлова. – Новосибирск: Наука, 2006.

10. Брычев, С.В. *Исследование математической модели экономики коммунального хозяйства малых городов: дис. ... канд. физ.-мат. наук* /

С.В. Брычев. – Челябинск: Челябинский гос. ун-т, 2002.

11. Свиридюк, Г.А. *Задача Веригина для линейных уравнений Соболевского типа с относительно  $p$ -секториальными операторами* / Г.А. Свиридюк, С.А. Загребина // *Дифференц. уравнения.* – 2002. – Т. 38, № 2. – С. 1646–1652.

12. Загребина, С.А. *О задаче Шоултера–Сидорова* / С.А. Загребина // *Изв. вузов. Серия «Математика».* – 2007. – № 3. – С. 22–28.

13. Гантмахер, Ф.Р. *Теория матриц* / Ф.Р. Гантмахер. – 4-е изд. – М.: Наука, 1988. – 552 с.

Поступила в редакцию 2 октября 2007 г.

# ПРИМЕНЕНИЕ МЕТОДОВ ИНТЕРВАЛЬНОГО АНАЛИЗА ПРИ КАЛИБРОВКЕ ПРЕОБРАЗОВАТЕЛЕЙ ДАВЛЕНИЯ

А.Е. Попов

## APPLICATION OF INTERVAL ANALYSIS IN CALIBRATION OF PRESSURE TRANSDUCERS

A.E. Popov

Применение интервального анализа при расчете калибровочных коэффициентов обусловлено тем, что в метрологии принята интервальная модель неопределенности. Предполагается, что некоторое измерение получено с помощью неточного прибора с известной абсолютной ошибкой измерения, которая включает как систематическую, так и случайную погрешности. В случае калибровки известен класс точности образцового датчика давления.

*Ключевые слова: интервальный анализ, калибровка, преобразователь давления.*

Application of interval analysis to calculate calibration coefficients is connected to the fact that in the metrology the interval model of uncertainty has been accepted. It is assumed that a measurement is obtained by using inaccurate device with a known absolute error of measurement, which includes both systematic and random errors. In the case of calibration, accuracy class of model pressure drive is known.

*Keywords: interval analysis, calibration, pressure transducer.*

### Введение

Расчет калибровочных коэффициентов измерительных преобразователей давления практически всегда производится с использованием методов регрессионного анализа и вероятностно-статистической модели.

Вероятностно-статистическая модель основана на предположении, что рассматриваемые переменные являются случайными величинами с заданными функциями плотности вероятности. Эта модель относится к классическим вследствие ее глубокой теоретической и методической проработанности для огромного числа приложений, включая регрессионный, корреляционный, факторный и дисперсионный анализ. Вероятностно-статистическая модель до сих пор является наиболее популярной и широко используется в практике калибровки преобразователей давления ввиду её простоты [1].

В рамках вероятностно-статистической модели по данным эксперимента находятся статистические оценки калибровочных коэффициентов прямой модели. Далее производится переход к обратной модели, в ходе которого возникают теоретические сложности при построении доверительного интервала, предсказать который является практически невозможно. Однако в ряде случаев возможность построения доверительного интервала

является необходимой. В этом случае можно гораздо точнее оценить класс точности измерительного преобразователя сразу после калибровочной процедуры и нет необходимости проводить трудоемкую процедуру поверки прибора.

Помимо широко используемого в настоящее время вероятностного подхода, существует также альтернативный подход, основанный на методах интервального анализа, в рамках которого задача построения доверительного интервала решается достаточно просто.

### 1. Введение в интервальный анализ

В интервальной модели неопределенность параметра  $x$  описывается границами его возможных значений в следующем виде:

$$[x] = [x - \Delta; x + \Delta].$$

В отличие от теории вероятности внутри интервала  $[x]$  не задается никакой вероятностной меры, т.е. все значения внутри интервала предполагаются равновероятными.

Применение интервального анализа при расчете калибровочных коэффициентов обусловлено тем, что в метрологии принята интервальная модель неопределенности. Предполагается, что некоторое измерение  $x$  получено с помощью неточного прибора с известной абсолютной ошибкой изме-

рения  $\Delta$ , которая включает как систематическую, так и случайную погрешности. В случае калировки известен класс точности образцового задатчика давления. Именно значение класса точности принимается в качестве известной абсолютной ошибки измерения  $\Delta$  [2].

Предметом интервального анализа является решение задач с интервальными неопределённостями и неоднозначностями в данных, возникающими в постановке задачи или на промежуточных стадиях процесса решения. Методы интервального анализа характеризуются рассмотрением множеств неопределённости как самостоятельных целостных объектов, посредством проведения над ними арифметических, аналитических операций и отношений [3].

Рассмотрим некоторые базовые операции над интервальными данными  $a$  и  $b$ :

$$a + b = [\underline{a} + \underline{b}, \overline{a} + \overline{b}],$$

$$a - b = [\underline{a} - \overline{b}, \overline{a} - \underline{b}],$$

$$a \cdot b = [\min\{\underline{a}\underline{b}, \underline{a}\overline{b}, \overline{a}\underline{b}, \overline{a}\overline{b}\}, \max\{\underline{a}\underline{b}, \underline{a}\overline{b}, \overline{a}\underline{b}, \overline{a}\overline{b}\}],$$

$$a / b = a \cdot [1/\overline{b}, 1/\underline{b}] \text{ для } b \neq 0.$$

Известно, что вероятностно-статистическая модель не позволяет учесть факторы неопределённости, обусловленные систематическими ошибками измерения и ошибками округления. Кроме того, постулируемое в вероятностной модели нормальное распределение, которое задает неограниченный диапазон величины, на практике часто оказывается неадекватным, например, для заведомо положительных переменных. Интервальная модель позволяет учесть любые факторы неопределённости [4].

Попробуем применить методы интервального анализа при расчете калибровочных коэффициентов.

## 2. Метод наименьших квадратов для интервальных данных

В рамках статистического подхода для нахождения оценок параметров обычно используется множественный регрессионный анализ, основанный на методе наименьших квадратов.

Метод наименьших квадратов можно применить при работе с интервальными данными. В этом случае в качестве экспериментальных данных необходимо использовать интервалы. Например, если образцовый задатчик давления показывает величину давления 15 кПа, а класс точности задатчика при работе с диапазоном 160 кПа составляет 0,008 %, то при расчете будет использован следующий интервал:

$$[x] = [15 - \Delta; 15 + \Delta];$$

$$[x] = [15 - 160 \cdot 0,008 / 100; 15 + 160 \cdot 0,008 / 100];$$

$$[x] = [14,9872; 15,0128].$$

Проведение численных расчетов с интервальными данными несколько затруднено в силу неко-

торых особенностей вычислений, однако общий ход вычислений соответствует стандартному методу наименьших квадратов.

Суть метода сводится к следующему.

Пусть полином имеет вид

$$f(x, z) = C_0 + C_1x + C_2x^2 + \dots + C_mx^m + C_{m+1}z + C_{m+2}zx + \dots + C_{m+m+1}zx^m + C_{m+m+2}z^2 + \dots + C_{m+m+3}z^2x + \dots + C_{m+m+m+2}z^2x^m + \dots + C_{m+n+n}z^n + \dots + C_{m+n+m-n}z^n x^m.$$

В более компактном виде

$$F = XZ \cdot C,$$

где  $F$  – матрица значений функции;  $XZ$  – матрица значений аргументов;  $C$  – матрица коэффициентов.

Критерий минимизации:

$$\min\{E^2 = (F - XZ \cdot C)^2\}.$$

Для получения минимума необходимо взять частную производную по вектору  $C$ .

$$\frac{dE}{dC} = -2 \cdot XZ^T \cdot F + 2 \cdot XZ^T \cdot XZ \cdot C = 0,$$

откуда

$$XZ^T \cdot XZ \cdot C = XZ^T \cdot F,$$

получаем

$$C = (XZ^T \cdot XZ)^{-1} XZ^T \cdot F.$$

Все остальные вычисления над матрицами проводятся по приведенному выше алгоритму. Для работы с интервальными матрицами существует специальная библиотека INTLAB, работающая в среде MATLAB.

## 3. Пример расчета калибровочных коэффициентов с использованием интервальных данных

Применяя функции библиотеки INTLAB, рассчитаем калибровочные коэффициенты измерительного преобразователя давления. Экспериментальные данные представлены в табл. 1.

Таблица 1  
Экспериментальные данные

Температура, °C	Код температуры	Давление, кПа	Код давления
1	2	3	4
0	14333207	0	8314658,81
0	14332247	3	8460744,63
0	14331635	5	8531040,45
0	14330439	10	8706786,81
0	14328955	16	8917639,72
0	14327933	20	9058226,81
0	14326759	25	9233991,18
0	14323707	40	9761025,36
0	14319637	60	10463472,8
0	14311885	100	11867937
0	14305841	130	12920312,4
0	14299557	160	13971705,5
20	14029485	0	8362142,09
20	14029111	3	8458444,81



Окончание табл. 1

1	2	3	4
20	14028857	5	8525074,45
20	14028201	10	8691797
20	14027417	16	8891855,72
20	14026909	20	9025206,45
20	14026209	25	9191976,63
20	14024063	40	9692122,27
20	14021113	60	10358671,1
20	14014839	100	11690962,0
20	14009885	130	12689079,7
20	14004717	160	13686233,9
50	13679453	0	8365945,36
50	13679045	3	8458600,81
50	13678767	5	8521816,09
50	13678059	10	8679713,36
50	13677163	16	8869111,72
50	13676577	20	8995417,36
50	13675849	25	9153258,27
50	13673725	40	9626679,90
50	13670861	60	10257539,7
50	13665077	100	11517994,8
50	13660485	130	12462044,4
50	13655805	160	13404765

Рассмотрим некоторые функции библиотеки INTLAB, необходимые для расчета коэффициентов измерительного преобразователя.

Функция «midrad» позволяет осуществить ввод интервального значения, используя центр и радиус интервала. Данную функцию необходимо применить для перевода экспериментальных данных, представленных в табл. 1, в интервальный вид. В качестве центра интервала необходимо принять значение, измеренное прибором. В качестве радиуса интервала примем класс точности образцового задатчика.

Для умножения и транспонирования матриц используются функции «/» и «transpose», которые переопределены для работы с интервальными данными внутри пакета INTLAB.

Для решения системы линейных алгебраических уравнений, заданных в интервальном виде, применяется функция «verifylss», использующая для решения метод Кравчука.

Запрограммируем представленный ранее алгоритм в виде *m*-файла MATLAB.

Итак, после запуска подпрограммы, реализующей метод наименьших квадратов для интервальных данных, получаем набор калибровочных коэффициентов. Данный набор также будет представлен в интервальном виде. То есть в результате работы программы будет произведен расчет нижних и верхних границ значений коэффициентов.

Для проверки результатов расчета подставим исходные коды в полученную математическую модель и вычислим интервалы характеристики полученных значений. Результаты представлены в табл. 2.

Таблица 2

Результаты расчета

Точка, кПа	Нижняя граница, кПа	Верхняя граница, кПа	Радиус интервала, кПа
0	0,060	0,073	0,007
3	2,956	2,971	0,008
5	4,958	4,976	0,009
10	9,966	9,993	0,014
16	15,972	16,012	0,020
20	19,974	20,023	0,025
25	24,976	25,039	0,031
40	39,966	40,079	0,057
60	59,918	60,124	0,103
100	99,739	100,241	0,251
130	129,555	130,394	0,419
160	159,371	160,662	0,646

Таким образом, в результате расчета с использованием методов интервальной статистики были получены значения доверительных интервалов.

Далее необходимо оценить точность полученных результатов. Для этого рассчитаем абсолютную величину отклонения середины полученного интервала от задаваемой точки. Данные расчета представлены в табл. 3.

Таблица 3

Оценка точности вычислений

Точка, кПа	Середина интервала, кПа	Величина абсолютной ошибки, кПа
0	0,067	0,067
3	2,963	0,037
5	4,967	0,033
10	9,980	0,020
16	15,992	0,008
20	19,999	0,001
25	25,008	0,008
40	40,023	0,023
60	60,021	0,021
100	99,990	0,010
130	129,975	0,025
160	160,017	0,017

### Заключение

Приведенный пример показывает, что применение интервального анализа позволяет снять некоторые проблемы и методические сложности, возникающие при решении прикладных задач статистическими методами. Интервал неопределенности позволяет описать широкий класс неопределенных, неоднозначных, вариабельных и неточных исходных данных. Значения ошибок в исходных данных могут колебаться в широких пределах. Результаты, полученные в рамках парадигмы с помощью интервального анализа, имеют ясную и четкую интерпретацию в терминах интервалов и областей неопределенности. В результате расчета

были получены интервальные значения калибровочных коэффициентов измерительного преобразователя, на основе которых затем получены доверительные интервалы.

*Литература*

1. Зоркальцев, В.И. Метод наименьших квадратов / В.И. Зоркальцев. – Новосибирск: ВО «Наука», 1995. – 220 с.
2. Попов, Б.А. Приближение функций для технических приложений / Б.А. Попов, Г.С. Теслер. – Киев: Наукова думка, 1980.
3. Шарый, С.П. Конечномерный интервальный анализ / С.П. Шарый. – М.: XYZ, 2008. – 569 с.
4. Померанцев, А.Л. Построение многомерной градуировки методом простого интервального оценивания / А.Л. Померанцев, О.Е. Родионова. – Журн. аналит. химии. – 2006. – № 61. – С. 1032–1047.

*Поступила в редакцию 3 февраля 2009 г.*

## ИНФОРМАЦИЯ ДЛЯ АВТОРОВ

1. **Тематика.** В журнале публикуются статьи по следующим научным направлениям: управление в различных отраслях техники, а также в административной, коммерческой и финансовой сферах; математическое, алгоритмическое, программное и аппаратное обеспечение компьютерных технологий, в том числе компьютерных комплексов, систем и сетей; измерительные системы, приборостроение, радиоэлектроника и связь.

2. **Предоставляемый материал.** В редакцию предоставляются бумажная и электронная (документ Microsoft Word) версии статьи, экспертное заключение о возможности опубликования статьи в открытой печати, информационный листок авторов, где на русском и английском языке указываются: Ф.И.О., место работы и должность для всех авторов, контактная информация ответственного за подготовку рукописи (рабочий и мобильный телефон, адрес для рассылки авторских экземпляров).

3. **Структура статьи.** Статья содержит УДК, название (не более 12–15 слов), список авторов, аннотацию (не более 300 знаков), список ключевых слов, введение, основной текст (структурированный по разделам), заключение (обсуждение результатов), литературу (в порядке цитирования, по ГОСТ 7.1–2003). В конце статьи следуют элементы на английском языке: название, аннотация, список ключевых слов. Бумажная версия статьи подписывается всеми авторами, в конце статьи ставится пометка, что авторы не возражают против размещения статьи в Интернет.

4. **Параметры набора.** Размеры полей: левое – 3 см, правое – 3 см, верхнее и нижнее – по 3 см. Текст статьи набирать шрифтом Times New Roman размером 10 пт. Выравнивание абзацев – по ширине. Отступ первой строки абзаца – 0,7 см. Междустрочный интервал – одинарный. Включить режим автоматического переноса слов. Все кавычки должны быть угловыми («»). Все символы «тире» должны быть среднего размера («–», а не «-»). Ключевые элементы статьи – шапка, заголовки разделов – следует выделять полужирным. Знак разделения целой и десятичной части числа – запятая. Между числом и единицей измерения должен стоять неразрывный пробел (Ctrl + Shift + Пробел).

5. **Формулы.** Набираются в Microsoft Equation, либо MathType с отступом 0,7 см от левого края. Размер обычных символов – 10 пт, размеры индексов первого порядка – 71 %, индексов второго порядка – 58 %. Номер формулы размещается за пределами формулы, непосредственно после нее, в круглых скобках.

6. **Рисунки и таблицы.** Рисунки имеют разрешение не менее 300 dpi. Рисунки нумеруются и имеют названия (Рис. 1. Здесь следует название рисунка). Таблицы нумеруются и имеют названия (Таблица 1. Здесь следует название таблицы).

7. **Адрес редакции.** 454080 Челябинск, пр. им. В.И. Ленина, 76, корп. 3б, 4 этаж – деканат ПС/КТУР-факультета, зам. отв. ред. д.т.н., проф. Л.С. Казаринову. Адрес электронной почты секретаря журнала: ezfm@yandex.ru

8. **Подробные требования к оформлению.** Полную версию требований к оформлению статей и пример оформления можно загрузить с сайта ЮУрГУ (<http://www.susu.ac.ru>), следуя ссылкой: «Научные исследования», «Издательская деятельность», «Вестник ЮУрГУ», «Серии».

9. Плата с аспирантов за публикацию рукописей не взимается.

**ВЕСТНИК  
ЮЖНО-УРАЛЬСКОГО  
ГОСУДАРСТВЕННОГО УНИВЕРСИТЕТА  
№ 26(159) 2009**

**Серия  
«КОМПЬЮТЕРНЫЕ ТЕХНОЛОГИИ,  
УПРАВЛЕНИЕ, РАДИОЭЛЕКТРОНИКА»  
Выпуск 10**

Редактор М.Н. Атауллина  
Компьютерная верстка С.В. Буновой

Издательский центр Южно-Уральского государственного университета

---

Подписано в печать 13.07.2009. Формат 60×84 1/8. Печать цифровая.  
Усл. печ. л. 10,69. Уч.-изд. л. 11,55. Тираж 500 экз. Заказ 309/322.

---

Отпечатано в типографии Издательского центра ЮУрГУ. 454080, г. Челябинск, пр. им. В.И. Ленина, 76.