



CONTROL SCIENCES

Научно-технический  
журнал

6 номеров в год

ISSN 1819-3161

УЧРЕДИТЕЛЬ

Институт проблем управления  
им. В.А. Трапезникова РАН

Главный редактор

Д.А. Новиков

Заместители главного  
редактора

Л.П. Боровских, Ф.Ф. Пащенко

Редактор

Т.А. Гладкова

Выпускающий редактор

Л.В. Петракова

Издатель

ООО «СенСиДат-Контрол»

Адрес редакции  
117997, ГСП-7, Москва,  
ул. Профсоюзная, д. 65, к. 272.  
Тел./факс (495) 334-92-00

E-mail: pu@ipu.ru  
www.ipu.ru/period/pu

Оригинал-макет  
и электронная версия  
подготовлены  
ООО «ЭЛЕКТРОНИНФОРМ»

Отпечатано с готовых диапозитивов  
в типографии ГКС

Подписано в печать  
3.07.2007 г.

Заказ № РВ407

Журнал зарегистрирован  
в Министерстве  
Российской Федерации  
по делам печати,  
телерадиовещания  
и средств массовых  
коммуникаций

Свидетельство о регистрации  
ПИ №77-11963  
от 06 марта 2002 г.

Журнал входит в Перечень ведущих  
рецензируемых журналов и изданий,  
в которых должны быть опубликованы  
основные научные результаты  
диссертаций на соискание ученой  
степени доктора и кандидата наук

Подписные индексы:  
80508 и 81708 в каталоге Роспечати  
38006 в объединенном каталоге  
«Пресса России»

# ПРОБЛЕМЫ УПРАВЛЕНИЯ

## 4.2007

### СОДЕРЖАНИЕ

#### Анализ и синтез систем управления

- Баранов В.В., Матросов В.М. Модель динамики в задачах управления деградирующими системами . . . . . 2
- Рывкин С.Е. Оценивание компонент вектора состояния сингулярной нелинейной динамической системы . . . . . 8
- Анисимов Д.Н., Хрипков А.В. Законы распределения оценок параметров динамических объектов при идентификации методом экспоненциальной модуляции . . . . . 14

#### Информационные технологии в управлении

- Алексеевский М.В. Интеграция и глобализация программного обеспечения обработки данных . . . . . 19
- Алгулиев Р.М., Алыгулиев Р.М. Ранжирование web-страниц с использованием взаимной информации между гиперссылками . . . . . 24
- Затуливетер Ю.С., Фищенко Е.А. Многопроцессорная архитектура ПС-2000 на кристалле СБИС . . . . . 30
- Васильев И.Л., Сидоров Д.Н. Приложение кластерного анализа к автоматическому распознаванию дефектов. . . . . 36

#### Управление в социально-экономических системах

- Чеботарев П.Ю., Митькин А.Н, Шмерлинг Д.С. Об оценивании вклада ведомственных целевых программ в достижение целей Правительства . . . 43
- Никитин В.В., Мальцева С.В., Дорофеюк А.А. и др. Классификация объектов профессиональной деятельности специалиста при проектировании профессиональных и образовательных стандартов . . . . . 51

- Винокур В.М.**, Мыльников Л.А., Перминова Н.В. Подход к прогнозированию успешности инновационного проекта . . . . . 56
- Волик Б.Г. Экономическая эффективность управляющих систем. . . . . 60
- Угольницкий Г.А., Усов А.Б. Математическая формализация методов иерархического управления эколого-экономическими системами . . . . . 64
- Михеев О.В., Самохина А.С. Алгоритм оперативной верификации опасного эпидемического заболевания вероятностным методом и его реализация . . . . . 70
- Беляков А.Ю., Пуйсанс С.Г., Столбова И.Д., Харитонов В.А. Модели системы управления качеством подготовки специалистов в инновационных технологиях обучения . . . . . 74

#### Управление подвижными объектами

- Борисов В.Г., Данилова С.К., Чинакал В.О. Создание и применение компьютерной технологии повышения безопасности управления морскими подвижными объектами . . . . . 79
- Евстигнеева Л.В., Тарасов Н.Н., Тахтамышев М.Г. Синтез алгоритма для одной задачи управления морскими подвижными объектами. . . . . 84

\* \* \*

- Contents & abstracts . . . . . 88

# МОДЕЛЬ ДИНАМИКИ В ЗАДАЧАХ УПРАВЛЕНИЯ ДЕГРАДИРУЮЩИМИ СИСТЕМАМИ<sup>1</sup>

В.В. Баранов, В.М. Матросов

Центр исследований устойчивости и нелинейной динамики  
при Институте машиноведения РАН, г. Москва

Построена модель переходной функции марковского процесса, описывающего управляемый процесс деградации. Модель ориентирована на применения к решению проблемы управления эффективностью и безопасностью деградирующих систем.

## ВВЕДЕНИЕ

В работе [1] рассмотрена проблема управления эффективностью и безопасностью деградирующих систем. Деградация определяется внутренними процессами износа, старения, усталости и т. д., которые ухудшают качество функционирования системы и завершаются отказом либо гибелью системы. Динамика таких процессов описывается марковским процессом с монотонными, обрывающимися траекториями. Применение управляющих воздействий по предупреждению отказов порождает управляемый марковский процесс, описывающий процесс использования, деградации и восстановления работоспособности системы. Стохастическая закономерность такого процесса определяется переходной функцией  $q^g(S|S \times Y)$ , задающей вероятности переходов на множестве состояний  $S$  под воздействием управлений из множества  $Y$  и зависящей от структурной альтернативы  $g \in G$  как от параметра. Развиваемые в работе [1] методология и методы используют переходную функцию  $q^g(S|S \times Y)$  в неявном виде. Однако для возможности практического решения задач в соответствии с методологией [1] переходная функция должна быть задана в явном виде. Поэтому задача настоящей статьи состоит в построении переходной функции  $q^g(S|S \times Y)$  в явном виде.

<sup>1</sup> Работа выполнена при поддержке РФФИ, проект № 05-0833574 а.

## 1. МОДЕЛЬ ДИНАМИКИ ПРОЦЕССА ДЕГРАДАЦИИ

Процесс деградации описывается марковским процессом с монотонными, обрывающимися траекториями. Без ограничения общности можно полагать, что траектории *не убывают* и принимают значения в компактном борелевском пространстве состояний  $(S, \mathcal{B})$ , где множество состояний является единичным отрезком  $S = [0, 1] \subset R^1$ ,  $\mathcal{B}$  –  $\sigma$ -алгебра его подмножеств. Кроме того, будем полагать, что процесс является *строго* марковским. Для задания его переходной функции необходимо указать вероятностную меру на переходах за время  $t > 0$  из произвольной точки  $z \in S$  в любое подмножество  $\Gamma = (a, b] \subset S$ , где  $0 \leq a < b \leq 1$ . Требуемую переходную функцию условимся обозначать  $q(t, \Gamma|z)$ . Для ее построения рассмотрим теоретико-вероятностные события, связанные с указанными переходами. Введем событие вида:

$$A(t, \Gamma|z) = \{\text{за время } t > 0 \text{ совершился переход из точки } z \in S \text{ в множество } \Gamma = (a, b]\}.$$

Поскольку траектории случайны и являются обрывающимися, то в случайный момент времени  $\xi > 0$  может произойти обрыв траектории, который условимся называть «отказом» системы. Очевидно, что событие  $A(t, \Gamma|z)$  происходит лишь в том случае, если за время  $t > 0$  отказ не произошел, т. е. при условии события  $B(t, \Gamma|z)$  вида:

$$B(t, \Gamma|z) = \{\text{за время } t > 0 \text{ перехода из точки } z \in S \text{ в множество } \Gamma \text{ отказ не произошел}\}.$$



С другой стороны, если событие  $A \equiv A(t, \Gamma|z)$  произошло, то произошло и событие  $B \equiv B(t, \Gamma|z)$ . Отсюда следует справедливость представления:

$$A = (A \cap B). \quad (1)$$

Дополнением  $\bar{A}$  к событию  $A$  является событие вида:

$\bar{A} \equiv \bar{A}(t, \Gamma|z) = \{ \text{за время } t > 0 \text{ совершился переход из точки } z \in S \text{ в множество } S \setminus \Gamma, \text{ включающее поглощающее состояние } \hat{s} = 1 \}.$

Поскольку события  $A$  и  $\bar{A}$  несовместны, то  $P(A) + P(\bar{A}) = 1$ . При этом вероятности событий  $A$  и  $\bar{A}$  являются вероятностями соответствующих переходов. С учетом этого наша задача состоит в указании таких вероятностей в явном виде.

По теореме умножения вероятностей из условия (1) получаем:

$$P(A) = P(A \cap B) = P(A|B)P(B). \quad (2)$$

Отсюда следует, что для вычисления вероятности  $P(A)$  события  $A$  необходимо задать условную вероятность  $P(A|B)$  и безусловную вероятность  $P(B)$ .

По определению события  $A$  вероятность  $P(A)$  является переходной вероятностью  $q(t, \Gamma|z)$ . Условная же вероятность  $P(A|B)$  является вероятностью перехода из точки  $z \in S$  в множество  $\Gamma \subset S$ , но при условии, что за время  $t > 0$  отказ не произошел. С учетом этого условную вероятность  $P(A|B)$  можно обозначить  $q(t, \Gamma|z, \xi > t)$ . Будем называть ее *условной* переходной функцией. Такая функция задает вероятностную меру на подмножествах  $\Gamma = (a, b) \subset [0, 1]$ , в которые может попасть процесс  $\{s_t\}$  за время  $t > 0$  из точки  $z \in S$  при условии, что за время  $t > 0$  отказ не произошел. Поскольку траектории  $s_t$  не убывают по  $t$ , то указанная мера сосредоточена на полуинтервале  $[z, 1) \subset [0, 1]$ .

Рассмотрим приращения  $r(t|z) = (s_t - z) \geq 0$  траектории  $s_t$  за время  $t \geq 0$  из точки  $z \in [0, 1)$ . Величина  $r(t|z)$  случайна и принимает значения из отрезка  $[0, 1 - z]$ . Поэтому и вся вероятностная мера, рассматриваемая на приращениях, сконцентрирована на отрезке  $[0, 1 - z]$ . При этом в силу однородности процесса она зависит лишь от длительности интервала времени  $(0, t]$ , на котором рассматриваются приращения, но не от момента времени  $t \geq 0$ . Очевидно, что вероятностная мера на приращениях однозначно определяет вероятности переходов  $q(t, \Gamma|z, \xi > t)$ . Способ построения требуемых вероятностей переходов основывается на существовании взаимно однозначного соответствия между вероятностными мерами и функциями распределения [2]. В частности, пусть задана

функция распределения  $F(x|z, t)$  приращений  $z(t|z) = s_t - z$ . Тогда вероятностная мера  $P$  на подмножествах  $\Gamma = (a, b) \subset [0, 1]$  определяется единственным способом в соответствии с условием:

$$P((a, b)|t, z) = F(b|t, z) - F(a|t, z), \quad \forall (a, b) \subset [0, 1].$$

С использование этого соотношения переходные вероятности  $q(t, \Gamma|z, \xi > t)$ ,  $\Gamma = (a, b) \subset [z, 1]$ ,  $b > a$ ,  $a \geq z$  определяются условиями вида:

$$q(t, (a, b)|z, \xi > t) = \begin{cases} F(b|t, z) - F(a|t, z), & (a, b) \subset [0, 1 - z], \\ 0, & (a, b) \not\subset [0, 1 - z]. \end{cases} \quad (3)$$

Заметим, что до момента обрыва траектория  $\{s_t\}$  непрерывна. Тогда естественно полагать, что функция распределения  $F(x|t, z)$  абсолютно непрерывна по переменной  $x$  на отрезке  $[0, 1 - z]$ . По переменной же времени  $t$  требует уточнений асимптотическое поведение функции  $F(x|t, z)$  при  $t \rightarrow 0$  и  $t \rightarrow \infty$ . С этой целью для определенности и наглядности рассуждений положим, что деградация определяется износом объекта. При этом естественно полагать, что за «малое время» объект «изнашивается мало». Отсюда следует, что на малом промежутке времени приращения траектории также малы и при  $t \rightarrow 0$  они будут стремиться к 0. В этих предположениях функция распределения  $F(x|t, z)$  должна удовлетворять следующему условию.

*Условие 1.* При  $t \rightarrow 0$  функция распределения  $F(x|t, z)$  стремится к вырожденной функции

$$F^0(x|z) = \begin{cases} 1, & x \geq 0, \\ 0, & x < 0. \end{cases}, \quad x \in R^1. \quad \blacklozenge$$

С другой стороны, ясно, что в силу деградации при неограниченном возрастании времени объект полностью изнашивается. В таком случае для заданного начального состояния  $z \in [0, 1)$  приращения траектории при  $t \rightarrow \infty$  стремятся к значению  $(1 - z)$ . В этих условиях функция распределения  $F(x|t, z)$  должна удовлетворять следующему условию.

*Условие 2.* При  $t \rightarrow \infty$  функция распределения  $F(x|t, z)$  стремится к вырожденной функции

$$F^\infty(x|z) = \begin{cases} 1, & x > 1 - z, \\ 0, & x \leq 1 - z, \end{cases}, \quad x \in R^1. \quad \blacklozenge \quad (4)$$

С учетом этих условий наша ближайшая задача состоит в построении функции распределения  $F(x|t, z)$  в явном виде.

Поскольку функция распределения  $F(x|t, z)$  абсолютно непрерывна по переменной  $x \in [0, 1]$ , то существует плотность распределения  $f(x|t, z)$  на

отрезке  $[0, 1]$ . Заметим, что среди непрерывных на единичном отрезке распределений, для которых плотность может быть задана в *явном виде*, известно лишь одно, так называемое  $\beta$ -распределение. Его плотность имеет вид:

$$f(x) = \begin{cases} \frac{1}{B(\alpha, \beta)} x^{\alpha-1} (1-x)^{\beta-1}, & x \in [0, 1], \\ 0, & x \notin [0, 1], \end{cases}$$

$\alpha > 0, \beta > 0$ , где  $\alpha$  и  $\beta$  — параметры,  $B(\alpha, \beta)$  — бета-функция. С учетом этого нужную функцию распределения  $F(x|t, z)$  будем строить с использованием плотности  $\beta$ -распределения.

Поскольку функция  $F(x|t, z)$  зависит от переменной времени  $t$ , то с использованием  $\beta$ -распределения его параметры также должны быть некоторыми функциями времени. Однако заметим, что при значениях параметров  $\alpha > 1$  и  $\beta > 1$  плотность  $\beta$ -распределения унимодальна с модой в точке

$$x = \frac{\alpha - 1}{\alpha + \beta - 2} \quad [3].$$

Но согласно условию 2 функция распределения  $F(x|t, z)$  при  $t \rightarrow \infty$  стремится к вырожденной функции распределения  $F^\infty(x|z)$  вида (4). Отсюда следует, что при  $t \rightarrow \infty$  требование унимодальности плотности  $\beta$ -распределения становится избыточным. Тогда естественно строить функцию распределения  $F(x|t, z)$ , фиксируя один параметр  $\beta$ -распределения, а другой рассматривать в качестве функции времени. Для определенности зафиксируем параметр  $\beta$ , положив  $\beta \equiv 1$ , параметр же  $\alpha$  будем определять некоторой функцией времени  $\alpha(t), t \in [0, \infty)$ . В рассматриваемых условиях плотность  $\beta$ -распределения, определенная на приращениях  $r(t|z) = (s_t - z)$ , записывается в виде

$$f(x|t, z) = \begin{cases} \alpha(t) \frac{x^{\alpha(t)-1}}{(1-z)^{\alpha(t)}}, & x \in [0, 1-z], \\ 0, & x \notin [0, 1-z], \end{cases}$$

$$\alpha(t) > 0.$$

Соответствующая функция распределения  $F(x|t, z)$  выражается в виде:

$$F(x|t, z) = \begin{cases} \left(\frac{x}{1-z}\right)^{\alpha(t)}, & x \in [0, 1-z], \\ 0, & x < 0, \\ 1, & x \geq 1-z. \end{cases} \quad (5)$$

Для однозначного задания функции распределения  $F(x|t, z)$  необходимо указать явный вид параметр-функции  $\alpha(t)$ . Для этого вначале сформулируем требования к ее асимптотическому поведению при  $t \rightarrow 0$  и  $t \rightarrow \infty$ .

Из условия 1 следует, что при  $t \rightarrow 0$  функция распределения вида (5) должна удовлетворять условию:

$$\lim_{t \rightarrow +0} \left(\frac{x}{1-z}\right)^{\alpha(t)} = 1, \quad x \in (0, 1-z].$$

Отсюда в свою очередь следует, что  $\alpha(t) \downarrow 0$  при  $t \rightarrow +0$ .

Аналогичным образом, из условия 2 следует, что  $\alpha(t) \uparrow \infty$  при  $t \rightarrow \infty$ .

Наконец, поскольку при фиксированном  $z \in [0, 1)$  приращения  $r(t|z) = s_t - z$  не убывают по  $t$ , то функция  $\alpha(t)$  должна быть монотонной по  $t$ , возрастая от 0 до  $+\infty$ .

Из выполненных построений следует, что искомая функция  $\alpha(t)$  является параметром переходной вероятностной меры  $q(t, \Gamma|z, \xi > t)$ . Тогда нужную функцию  $\alpha(t)$ , вообще говоря, можно искать из решения уравнения Колмогорова—Чепмена для переходной функции марковского процесса [2, 3]. Однако с использованием плотности  $\beta$ -распределения решение уравнения Колмогорова—Чепмена в явном виде не выражается. Но тогда можно воспользоваться следствием уравнения Колмогорова—Чепмена для моментов, которое для первого момента выражается в явном виде. С использованием переходной плотности уравнение Колмогорова—Чепмена для первого момента запишется в виде:

$$\int_z^1 f(x|t, z) dx = \int_z^1 x \int_z^1 f(x|t - \tau, s) f(s|\tau, z) ds d\tau, \quad \forall \tau: 0 < \tau < t.$$

В условиях используемой здесь плотности  $\beta$ -распределения это уравнение сводится к функциональному уравнению вида:

$$\frac{\alpha(t)}{\alpha(t) + 1} = \frac{\alpha(\tau)}{\alpha(\tau) + 1} + \frac{\alpha(t - \tau)}{(\alpha(\tau) + 1)[\alpha(t - \tau) + 1]}. \quad (6)$$

**Утверждение 1.** Решением уравнения (6) является функция  $\alpha(t)$  вида:

$$\alpha(t) = e^{\mu t} - 1, \quad \mu > 0. \quad \blacklozenge \quad (7)$$

Справедливость этого утверждения может быть проверена непосредственной подстановкой.

**Замечание.** Можно показать, что параметр  $\mu > 0$  полученной в правой части (7) функции имеет смысл средней скорости износа. Отсюда открывается возможность оценки параметра  $\mu$ . Однако способ его оценки здесь мы рассматривать не будем.

С использованием полученной функции  $\alpha(t) = e^{\mu t} - 1$  можно теперь указать явный вид переходной вероятности  $q(t, \Gamma|z, \xi > t)$ . Действительно, пусть задан параметр износа  $\mu > 0$ . Тогда с исполь-



зованием функции распределения  $F(x|t, z)$  вида (5) и представления (3) для переходной вероятности из точки  $z \in S$  в подмножество  $\Gamma = (a, b] \subset [0, 1 - z]$ , получаем следующее выражение для переходной вероятности

$$\begin{aligned}
 q(t, \Gamma|z, \xi > t) &= q(t, (a, b]|z, \xi > t) = \\
 &= F(b|t, z) - F(a|t, z) = \left(\frac{b}{1-z}\right)^{\alpha(t)} - \left(\frac{a}{1-z}\right)^{\alpha(t)} = \\
 &= \left(\frac{b}{1-z}\right)^{e^{\mu t} - 1} - \left(\frac{a}{1-z}\right)^{e^{\mu t} - 1}. \quad (8)
 \end{aligned}$$

Полученное выражение в явном виде определяет условную вероятность  $P(A|B) = q(t, \Gamma|z, \xi > t)$ , используемую для вычисления безусловной вероятности события  $A$  по формуле (2) вида:  $P(A) = P(A|B)P(B)$ . Остается в явном виде выразить безусловную вероятность  $P(B)$ . С этой целью рассмотрим событие вида:

$$B(z) \equiv \{\text{при фиксированном } z \in S \text{ за время } t \text{ отказ не произошел}\}.$$

Тогда событие  $B$  можно представить пересечением  $B = \bigcap_{z \in \Gamma} \{B(z)\}$ , где  $\Gamma \equiv (a, b] \subset [0, 1 - z] \subset S$ .

Если события  $B(z)$  независимы для всех  $z \in \Gamma = (a, b]$ , то вероятность события  $B$  будет определяться произведением:  $P(\bigcap_{z \in (a, b)} \{B(z)\}) = \prod_{z \in (a, b]} P(B(z)) = P(B)$ . Однако в условиях деградации события  $B(z)$ ,  $z \in (a, b]$  на самом деле зависимы в совокупности. С учетом этого имеет место следующее

**Утверждение 2.** В условиях деградации для вероятностей событий  $B = \bigcap_{z \in (a, b)} \{B(z)\}$  и  $B(b)$  справедливо равенство:  $P(B) = P(B(b))$ , где  $b$  — правая граница полуинтервала  $(a, b] \subset S$ . ♦

С учетом определения событий  $B(z)$ ,  $z \in (a, b]$  и полученного утверждения можно использовать обозначение  $P(\xi > t|b) \equiv P(B)$ . Тогда вероятность события  $A$ , определенная в соответствии с формулой (2), запишется в виде:

$$P(A) \equiv q(t, \Gamma|z) = q(t, \Gamma|z, \xi > t)P(\xi > t|b). \quad (9)$$

Поскольку здесь условная вероятность  $q(t, \Gamma|z, \xi > t)$  однозначно определена (формулой (8)), то остается указать способ явного задания вероятности  $P(B) = P(\xi > t|b)$ .

Заметим, что при фиксированном состоянии  $z \in S$  отказ может произойти либо не произойти, и других вариантов событий не существует. Поэтому событие вида

$$\bar{B}(z) \equiv \{\text{при фиксированном состоянии } z \in S \text{ за время } t > 0 \text{ произошел отказ}\}$$

является дополнительным к событию  $B(z)$ . Тем самым  $P(B(z)) + P(\bar{B}(z)) = 1$ . Отсюда имеем

$$P(\bar{B}(z)) = 1 - P(B(z)) = 1 - P(\xi > t|z).$$

Здесь  $\xi$  является случайной величиной времени жизни до отказа. Тогда функция вида

$$1 - P(\xi > t|z) = F(\xi < t|z) \equiv F_\xi(t|z)$$

является функцией распределения случайной величины  $\xi$  при фиксированном состоянии  $z \in S$ .

Из определения события  $B(z)$  и свойств деградации следует, что функция распределения  $F_\xi(t|z)$  должна удовлетворять следующим асимптотическим условиям:

$$\lim_{t \rightarrow \infty} F_\xi(t|z) = 1, \quad \forall z \in S.$$

$$\lim_{t \rightarrow \infty} F_\xi(t|z) = 1, \quad \forall t > 0. \quad (10)$$

$$\lim_{t \rightarrow \infty} F_\xi(t|z) = p(t), \quad 0 < p(t) < 1. \quad (11)$$

Эти условия означают, что для любого  $t > 0$  при любом  $z \in [0, 1]$ , в том числе при  $z = 0$ , отказ может произойти с положительной вероятностью  $p(t) > 0$ , а при  $z = 1$  отказ происходит достоверно при любом  $t > 0$ .

Поскольку эти условия являются очевидными следствиями определения события  $B(z)$  и свойств деградации, то формально будем называть их условиями деградации.

По исходным предположениям процесс деградации является однородным строго марковским процессом с неубывающими и обрывающимися траекториями. В этих условиях моменты  $\xi$  отказа являются марковскими моментами, т. е. не зависящими от будущего [2, 3]. В рассматриваемых предположениях требуется указать явный вид функции распределения  $F_\xi(t|z)$ .

С этой целью введем функцию вида

$$\Lambda(t, z) = -\frac{dF_\xi(t|z)}{1 - F_\xi(t|z)}, \quad t \in [0, \infty), z \in [0, 1].$$

Следуя терминологии теории надежности, такую функцию будем называть опасностью отказа [4]. Но в отличие от предположений теории надежности здесь опасность отказа зависит не только от времени  $t \in [0, \infty)$ , но и от состояния  $z \in [0, 1]$ . Поскольку по предположению процесс деградации является строго марковским, то моменты отказов являются марковскими моментами в том смысле, что вероятность их возникновения на интервале  $(t, t + \tau)$  не зависит от момента времени  $t$ , а зависит лишь от длины интервала времени  $\tau$  и состояния  $z \in S$  [3]. В этих условиях так же, как в

теории надежности, можно показать, что функция опасности отказа  $\Lambda(t, z)$  не зависит от времени  $t$ , и зависит лишь от состояния  $z \in E$ . С учетом этого по аналогии с построениями теории надежности [4] получаем, что функция распределения  $F_\xi(t|z)$  выражается в виде

$$F_\xi(t|z) = 1 - e^{-\Lambda(z)t}, \quad (12)$$

а соответствующая ей плотность вероятности является экспоненциальной функцией вида:

$$f(t|z) = \Lambda(z)e^{-\Lambda(z)t}.$$

Явный вид функции  $\Lambda(z)$  определяется из условий деградации (10) и (11). В частности, из условия (10) следует, что функция  $\Lambda(z)$  при  $z \rightarrow 1$  должна стремиться к  $+\infty$ , а из условия (11) следует, что при  $z \rightarrow 0$  функция  $\Lambda(z)$  должна стремиться к некоторой постоянной  $\lambda > 0$ . Требуемым условиям очевидным образом удовлетворяет функция вида

$$\Lambda(z) = \lambda/(1 - z), \quad z \in [0, 1], \lambda > 0.$$

Отсюда, в частности, следует, что параметр  $\lambda > 0$  имеет смысл интенсивности отказов «нового» объекта (точнее, при значении  $z = 0$ ), и согласно работе [4] может быть вычислен из соотношения  $\lambda = 1/T_0$ , где  $T_0$  — среднее времени жизни при фиксированном значении  $z = 0$ .

Теперь напомним, что наша окончательная задача состоит в указании безусловной вероятности события  $A$  в соответствии с формулой (9) вида

$$P(A) \equiv q(t, (a, b]|z) = q(t, (a, b]|z, \xi > t)P(\xi > t|b), \quad (13)$$

где  $b$  — правая граница полуинтервала  $(a, b] \subset S$ , а условная вероятность  $q(t, (a, b]|z, \xi > t)$  определена в явном виде выражениями (8). Вероятность же условия  $P(\xi > t|b) = 1 - F_\xi(t|b) \equiv P(B)$  с учетом выражения (12) определяется в явном виде формулой:

$$P(\xi > t|b) = e^{-\Lambda(z)t} = e^{-\frac{\lambda t}{1-b}}. \quad (14)$$

Подставляя в правую часть (13) соответствующие формулы (8) и (14), получаем явное выражение для переходной вероятности:

$$P(A) \equiv q(t, (a, b]|z) = q(s, t, (a, b]|\xi > t)P(\xi > t|b) =$$

$$= \begin{cases} \left[ \left( \frac{b}{1-z} \right)^{e^{\mu t} - 1} - \left( \frac{a}{1-z} \right)^{e^{\mu t} - 1} \right] e^{-\frac{\lambda t}{1-b}}, & (1 > b > a, a \geq z, \mu > 0, \lambda > 0,) \\ 0, & b = 1, \\ 0, & b = a. \end{cases} \quad (15)$$

Этим в явном виде определена безусловная вероятность события  $A$  вида:

$$A \equiv A(t, \Gamma|z) = \{\text{за время } t > 0 \text{ совершился переход из начальной точки } z \in S \text{ в множество } \Gamma \subset S\}.$$

Остается указать вероятность дополнительного к событию  $A$  события  $\bar{A} \equiv \bar{A}(t, \Gamma|z)$ . Очевидным образом она выражается в виде

$$P(\bar{A}) = (q(t, (E \setminus (a, b)]|z) = 1 - P(A) = 1 - q(t, (a, b]|z),$$

где вероятность  $P(A) \equiv q(t, (a, b]|z)$  определена выражением (15).

Таким образом, функция  $q(t, (a, b]|z)$  вида (15) является вероятностной мерой на  $\sigma$ -алгебре  $\mathcal{B}$ . При этом такая функция очевидным образом измерима по переменной  $z \in S$ . Тем самым она является *переходной функцией* марковского процесса в пространстве состояний  $(S, \mathcal{B})$ , описывающего динамику деградирующего объекта.

Полученная функция  $q(t, \Gamma|z)$  зависит от двух параметров: скорости деградации  $\mu > 0$  и интенсивности отказов  $\lambda > 0$ . С учетом этого для нее будем использовать обозначение  $q^{(\mu, \lambda)}(t, \Gamma|z)$ . Учитывая, что  $\Gamma = (a, b]$  является произвольным множеством из  $S$ ,  $z$  — произвольная точка из  $S$ , для полученной переходной функции можно использовать обозначение  $q^{(\mu, \lambda)}(t, S|S)$ .

## 2. МОДЕЛЬ ДИНАМИКИ УПРАВЛЯЕМОЙ ДЕГРАДАЦИИ

Из явного выражения (15) для переходной функции следует, что имеется две возможности управления процессом деградации: путем изменения параметров  $\mu$  и  $\lambda$  переходной функции  $q^{(\mu, \lambda)}(t, S|S)$  либо путем непосредственного изменения состояния  $z \in S$ . Положим, что изменение параметров  $\mu$  и  $\lambda$  достигается применением альтернатив из заданных множеств альтернатив использования  $u \in U$ , альтернатив безопасности  $v \in V$  и структурных альтернатив  $h \in H$ . С учетом этого будем полагать, что рассматриваемые параметры являются функциями соответственно  $\mu(u, v, h)$  и  $\lambda(u, v, h)$ ,  $(u, v, h) \in U \times V \times H$ .

Конкретный вид таких функций определяется особенностями объекта и предполагается заданным. С учетом этого зависимость переходной функции от выбора соответствующих альтернатив  $(u, v, h) \in U \times V \times H$  при необходимости будем отражать обозначениями:  $q^{(u, v, h)}(\cdot) \equiv q^{(\mu(u, v, h), \lambda(u, v, h))}(\cdot)$ .



Управление же путем непосредственного изменения состояния достигается использованием управляющего воздействия восстановления работоспособности, выбираемого из заданного множества  $P$  в зависимости от ситуации работоспособности  $x \in X$ . Под воздействием  $p \in P$  осуществляется мгновенный детерминированный переход из наблюдаемого состояния  $s \in S$  в новое «улучшенное» состояние  $z < s$  в соответствии с правилом:

$$z(s, p) = (1 - \varepsilon(p)) s,$$

где  $\varepsilon(p) \in [0, 1]$  — мера эффективности восстановления работоспособности, заданная для воздействия  $p \in P$ . Из полученной точки  $z = z(s, p) \in S$  на выбранном шаге времени  $\theta \in \Theta$  происходит эволюция состояний с закономерностью, определяемой переходной функцией  $q^{(\mu, \lambda)}(\theta, S|z)$ , задаваемой выражением (15) в предположении, что  $t = \theta \in \Theta$ ;  $\mu = \mu(u, v, h)$ ;  $\lambda = \lambda(u, v, h)$ ,  $(u, v, h) \in U \times V \times H$ .

Обозначим  $g = (u, v, h, \theta) \in U \times V \times H \times \Theta = G$ . Тогда в соответствии с описанной схемой переходная функция процесса управляемой деградации определяется следующими условиями:

$$\begin{aligned} q^g(\Gamma|s, y) &= q^{(u, v, h)}(\theta, \Gamma|s, (u, p)) = \\ &= \begin{cases} q^{(\mu(u, v, h), \lambda(u, v, h))}(\theta, \Gamma|z), \\ z = (1 - \varepsilon(p))s; \quad s \in S, \end{cases} \end{aligned}$$

где  $s \in S$ ,  $y = (u, p) \in U \times P$ ,  $\Gamma = (a, b) \subset [0, 1]$ ,  $q^{(\mu(u, v, h), \lambda(u, v, h))}(\theta, \Gamma|z)$  — переходная функция, определенная условиями (15).

Этим определена общая структура переходной функции  $q^g(S|S \times Y)$  управляемого процесса деградации. Для ее окончательного однозначного задания требуется уточнение параметр-функций  $\mu(u, v, h)$ ,  $\lambda(u, v, h)$  и меры  $\varepsilon(p)$  эффективности восстановления. Но эти функции могут быть заданы, лишь исходя из условий конкретного объекта. Поэтому без ограничения общности можно пола-

гать их заданными. С учетом этого переходную функцию  $q^g(S|S \times Y)$  управляемого процесса правомерно полагать однозначно построенной.

## ЗАКЛЮЧЕНИЕ

Построена модель *переходной функции*, определяющей стохастическую закономерность динамики управляемой деградации, ориентированной на применения к проблеме управления эффективностью и безопасностью деградирующих систем в соответствии с методологией, развитой в работе [1]. Однако методология формализации проблемы не исчерпывается заданием требуемой переходной функции. Другим важнейшим объектом методологии [1] является *функция полезности*, представляющая априорные предпочтения на управляющих альтернативах. Построение требуемой функции полезности представляет самостоятельный теоретический и практический интерес и является предметом самостоятельного исследования, которое существенно опирается на полученные в настоящей статье результаты и будет представлено в следующей статье.

## ЛИТЕРАТУРА

1. Матросов В.М., Баранов В.В. Проблема превентивной безопасности. Модель и методы принятия решений // Проблемы управления. — 2006. — № 5. — С. 2—11.
2. Ширяев А.Н. Вероятность. — М.: Наука, 1980. — 574 с.
3. Справочник по теории вероятностей и математической статистике. — Киев: Наукова думка, 1978. — 582 с.
4. Гнеденко Б.В., Беляев Ю.К., Соловьев А.Д. Математические методы в теории надежности. — М.: Наука, — 1965. — 524 с.

☎ (495) 132-01-47, e-mail: ciund@orc.ru

Статья представлена к публикации членом редколлегии В.Ю. Рутковским. □

# ФОРСАЙТ

**Вышел в свет первый номер нового информационно-аналитического журнала «Форсайт».**

**Форсайт — это система методов экспертного оценивания для выбора стратегических приоритетов развития сложных систем, таких как государство, регион, крупная компания, отрасль экономики. С помощью этой технологии, в частности, выявляются области перспективных научных исследований.**

**Основная задача журнала — распространение методологических, аналитических и информационных материалов о разработках российских и зарубежных авторов по системе Форсайт.**

**В журнале несколько ключевых разделов: стратегии, инновации и экономика, наука, государство, презентации, мастер-класс.**



# ОЦЕНИВАНИЕ КОМПОНЕНТ ВЕКТОРА СОСТОЯНИЯ СИНГУЛЯРНОЙ НЕЛИНЕЙНОЙ ДИНАМИЧЕСКОЙ СИСТЕМЫ<sup>1</sup>

С.Е. РЫВКИН

*Институт проблем управления им. В.А. Трапезникова РАН, г. Москва*

Для класса нелинейных динамических систем с линейным вхождением вектора оцениваемых компонент полного вектора состояния сформулировано и доказано достаточное условие наблюдаемости. Предложены методы нелинейного оценивания компонент переменных состояния на скользящих режимах. Дан пример синтеза алгоритма оценивания выходных механических переменных синхронного двигателя.

## ВВЕДЕНИЕ

Одна из центральных задач при синтезе высококачественных динамических систем заключается в получении информации о состоянии управляемого процесса и, в частности, о компонентах вектора состояния.

Применение классического подхода, основанного на прямом измерении всех необходимых координат, в настоящее время практически сходит на нет, что обусловлено как сложностью прямого получения информации о некоторых компонентах вектора состояния, так и общим усложнением объектов управления в целом, что неизбежно приводит к ухудшению их эксплуатационных и стоимостных показателей. Возможный путь преодоления этих недостатков состоит в исключении датчиков тех координат, прямое измерение которых является нежелательным, и использование при синтезе управления их оценок, полученных с помощью наблюдателя вектора состояния [1, 2].

Однако большинство объектов управления являются нелинейными, и получение по измерениям входных и выходных переменных необходимых оценок компонент вектора состояния представляет собой сложную задачу и требует разработки новых методов нелинейного оценивания вектора со-

стояния. Для решения рассматриваемой задачи перспективно применение методов теории систем со скользящими режимами [3, 4], предполагающих использование для оценивания адаптивных свойств динамической модели с разрывными управлениями. Это позволяет осуществить эффективную декомпозицию задачи наблюдения благодаря организации многомерного скользящего движения, в котором по значениям компонент эквивалентного управления динамической модели можно получить оценку неизвестных компонент вектора состояния. Данный подход приводит к существенному упрощению схемной реализации, что весьма существенно, особенно с учётом нелинейностей объекта управления и его высокого порядка.

## 1. ПОСТАНОВКА ЗАДАЧИ

Рассмотрим нелинейную динамическую систему следующего вида:

$$dx_1/dt = f_1(x_1, t) + D(x_1, t)x_2(t) + B(x_1, t)u(t), \quad (1)$$

$$dx_2(t) = f_2(x, u, t), \quad (2)$$

$$y(t) = x_1(t), \quad (3)$$

где  $x^T = (x_1^T, x_2^T)$  — вектор состояния,  $x(t) \in R^n$ ,  $x_1(t) \in R^q$ ;  $f_1(x_1, t)$ ,  $f_2(x, u, t)$  — векторы-столбцы,  $f_1(x_1, t) \in R^q$ ,  $f_2(x, u, t) \in R^{n-q}$ ;  $u(t)$  — вектор управления,  $u(t) \in R^m$ ;  $B(x_1, t)$  — матрица размерности

<sup>1</sup> Работа рекомендована к публикации Программным комитетом Третьей международной конференции по проблемам управления (Москва, 20 – 22 июня 2006 г.).





$q \times m$ ;  $D(x_1, t)$  — матрица размерности  $q \times (n - q)$ ;  $y(t)$  — вектор выходных переменных.

В уравнение (1) вектор  $x_2$ , компоненты которого надо определить, входит линейно и может рассматриваться как некоторый параметр. Для получения оценок его компонент можно воспользоваться подходом, основанном на методе «быстрой» идентификации на скользящих режимах [5].

### 1. ДОСТАТОЧНОЕ УСЛОВИЕ НАБЛЮДАЕМОСТИ

**Утверждение 1.** Оценка вектора состояния  $x_2$  нелинейной динамической системы вида (1)–(3) может быть получена, если размерность вектора выходных переменных  $y(t)$ , т. е. измеряемого вектора состояния  $x_1(t)$ , больше или равна размерности оцениваемого вектора состояния  $x_2$ ,  $q \geq (n - q)$ , и в матрице  $D(x_1, t)$  может быть выделена невырожденная квадратная матрица полной размерности  $\hat{D}(x_1, t) \in R^{(n-q) \times (n-q)}$ ,  $\text{rank } \hat{D}(x_1, t) = (n - q)$  при всех  $x_1(t)$ . ♦

Доказательство см. в Приложении.

### 2. НЕПОСРЕДСТВЕННАЯ ОЦЕНКА ПО ВЕКТОРУ ЭКВИВАЛЕНТНОГО УПРАВЛЕНИЯ

Получение информации о векторе состояния  $x_2$  на основе используемой в доказательстве утверждения 1 нелинейной динамической модели (П. 2) — см. Приложение — с единичной матрицей  $\hat{D}^*(t)$ , являющейся нелинейным наблюдателем вектора состояния  $x_2$ , требует дополнения модели блоком линейного преобразования полученных значений компонент вектора эквивалентного управления  $v_{eq}$  согласно выражению (П. 11).

Во избежание этого желательно путем выбора соответствующего вида матрицы  $\hat{D}^*(t)$  и функций переключения обеспечить непосредственное равенство векторов эквивалентного управления  $v_{eq}$  и состояния  $x_2$ .

**Утверждение 2.** В качестве оценки вектора состояния  $x_2$  нелинейной динамической системы вида (1)–(3) может непосредственно использоваться вектор эквивалентного управления, если выполнено достаточное условие наблюдаемости и матрица  $\hat{D}^*(t)$  динамической модели (П. 2) выбрана в виде  $\hat{D}(x_1, t)$ . ♦

Доказательство см. в Приложении.

### 4. НЕПОСРЕДСТВЕННАЯ ОЦЕНКА БЕЗ КООРДИНАТНОГО ПРЕОБРАЗОВАНИЯ

С позиций применения приведенного ранее подхода для построения нелинейного наблюдателя вектора состояния  $x_2$  существенный недостаток заключается в необходимости координатного преобразования функций рассогласования (П. 14), связанного с переходом к новому вектору функций рассогласования  $S^*$ . Этого можно избежать, если формировать этот вектор также с использованием разрывного управления.

**Утверждение 3.** В качестве оценки вектора состояния  $x_2$  нелинейной динамической системы вида (1)–(3) может непосредственно использоваться вектор эквивалентного управления  $v_{eq}$ , а в качестве функций рассогласования — рассогласования модельных и фактических переменных (П. 4), если выполнены условия утверждения 1, матрица  $\hat{D}^*(t)$  выбрана в виде  $\hat{D}(x_1, t)$ , а новый вектор функций рассогласования  $S^{*T} = (S_1^*, \dots, S_{n-q}^*)$  формируется дополнительной динамической системой вид

$$dS^*/dt = -S^* + R_S u(t), \quad (4)$$

где  $u(t) = (u_1, \dots, u_{n-q})$ ,  $u_i = \text{sgn}(\varepsilon_i)$ ,  $i = \overline{1, (n - q)}$ . ♦

Доказательство см. в Приложении.

### 5. ФИЗИЧЕСКИЙ СМЫСЛ ЭКВИВАЛЕНТНОГО УПРАВЛЕНИЯ И СИНГУЛЯРНОСТЬ

Эквивалентное управление  $u_{eq}$  определяется через среднее значение  $u_{cp}$  разрывного управления  $v(t)$ , полученное при движении системы в реальном скользящем режиме в  $\Delta$ -окрестности пересечения поверхностей разрывов  $S = 0$  с помощью фильтра

$$\tau(du_{cp}/dt) + u_{cp} = v(t),$$

где  $\tau$  — постоянная времени [3].

Оно представляет собой среднее значение разрывного управления, полученное при предельном переходе от реального скользящего режима к идеальному скольжению:

$$\lim_{\substack{\tau \rightarrow 0 \\ \Delta/\tau \rightarrow 0}} u_{cp} = u_{eq}$$

При построении наблюдателя необходимо учитывать эту особенность, т. е. наличие фильтра с собственной динамикой. При анализе требований, накладываемых на наблюдаемую нелинейную систему, необходимо учитывать, что оценка компо-

нент вектора состояния  $x_2$  возможна в том случае, если фильтр является сингулярно возмущенным по отношению к подсистеме (2), т. е. темпы фильтрации должны быть существенно выше (на порядок) темпов изменения оцениваемых компонент вектора состояния.

## 6. СИНТЕЗ НАБЛЮДАТЕЛЯ МЕХАНИЧЕСКИХ КООРДИНАТ СИНХРОННОГО ДВИГАТЕЛЯ

Приведенные ранее достаточные условия оценивания компонент вектора состояния служат основой для синтеза нелинейных наблюдателей, построенных на принципах систем с преднамеренно введенным скользящим режимом. Основная идея синтеза состоит в следующем: по имеющейся информации о векторе выходных переменных строится динамическая модель, наиболее полно описывающая исходную систему, а затем модельное управление выбирается таким образом, чтобы скользящее движение по пересечению поверхностей разрыва было устойчиво, а полученное эквивалентное управление содержало информацию об оцениваемом векторе состояния.

Далее предложенный подход к синтезу нелинейных наблюдателей будет конкретизирован применительно к задаче информационного обеспечения управления автоматизированным синхронным электроприводом с неявнополюсным синхронным двигателем (СД).

Традиционно применяемый подход к синтезу системы управления, основанный на прямом измерении всех необходимых координат, приводит к значительному усложнению конструкции электропривода, ухудшению его эксплуатационных и стоимостных показателей. Для преодоления этих недостатков следует исключить из конструкции системы датчики механических координат и использовать при синтезе управления их оценки, полученные с помощью наблюдателя [6, 7].

Как уже указывалось, наблюдатель, обеспечивающий получение оценок неизменяемых величин, обычно представляет собой имитационную модель исследуемого объекта, степень приближения которой к данному объекту определяется имеющейся в наличии информацией. Очевидно, что его структура, а также алгоритм получения необходимых оценок механических координат, особенно в случае нелинейного объекта, каковым является СД, существенным образом зависят от типа СД.

Приведем процедуру синтеза наблюдателя неявнополюсного СД с возбуждением от постоянных магнитов в неподвижной системе координат  $(\alpha, \beta)$ .

Известной (исходной) информацией служит информация о векторе тока  $I^T = (i_\alpha, i_\beta)$  и векторе

напряжения  $U^T = (u_\alpha, u_\beta)$  статорной обмотки (или, что то же самое, фактические фазные токи и напряжения), измерение которых не вызывает затруднений, и параметры электрических цепей. Необходимо оценить следующие механические координаты: угловое положение  $\Gamma$  ротора и частоту  $\Omega$  его вращения, пропорциональные, соответственно, углу  $\gamma_R$  между неподвижной осью  $R$  трёхфазной системы координат и подвижной осью  $d$  и электрической частоте  $\omega$  ( $\gamma_R = p\Gamma$ ;  $\omega = p\Omega$ , где  $p$  — число пар полюсов двигателя). Для простоты будем полагать, что СД имеет два полюса, т. е. электрические и механические угловые координаты равны между собой.

Электромагнитные процессы в СД в неподвижной системе координат  $(\alpha, \beta)$  описываются следующей системой дифференциальных уравнений:

$$\begin{aligned} di_\alpha/dt &= (-ri_\alpha + \Omega\psi\sin\Gamma + u_\alpha)/L, \\ di_\beta/dt &= (-ri_\beta - \Omega\psi\sin\Gamma + u_\beta)/L, \end{aligned} \quad (5)$$

где  $r$  — активное сопротивление обмотки статора;  $L$  — индуктивности обмоток статора;  $\psi$  — поток возбуждения.

Для определения углового положения ротора  $\Gamma$  и частоты его вращения  $\Omega$  построим динамическую систему, которая по своей структуре максимально приближена к структуре исследуемого объекта (5) и в которой используется вся имеющаяся информация о СД, включая информацию о приложенном к статорной обмотке напряжении и протекающем в ней токе:

$$\begin{aligned} d\hat{i}_\alpha/dt &= (-r\hat{i}_\alpha + u_\alpha)/L + u_1, \\ d\hat{i}_\beta/dt &= (-r\hat{i}_\beta + u_\beta)/L + u_2, \end{aligned} \quad (6)$$

где  $u_1$  и  $u_2$  — корректирующие модельные воздействия.

Для правильного отражения моделью (6) процессов, протекающих в СД, необходимо, чтобы совпадали движения в реальной (5) и модельной (6) электромагнитных системах, т. е. совпадали поведения соответствующих модельных  $\hat{i}_\alpha, \hat{i}_\beta$  и реальных  $i_\alpha, i_\beta$  компонент тока статора. Это можно обеспечить, организовав с использованием корректирующих воздействий  $u_1$  и  $u_2$  скользящее движение по пересечению поверхностей скольжения, представляющих собой нулевые рассогласования по компонентам тока:

$$S_\alpha = i_\alpha - \hat{i}_\alpha, \quad S_\beta = i_\beta - \hat{i}_\beta. \quad (7)$$

В этом случае матрица перед вектором корректирующих воздействий в уравнении проекции движения исходных динамических систем (5) и (6)



на подпространство рассогласований (7) будет единичной:

$$\frac{d}{dt} \begin{vmatrix} S_\alpha \\ S_\beta \end{vmatrix} = -\frac{r}{L} \begin{vmatrix} 1 & 0 \\ 0 & 1 \end{vmatrix} \begin{vmatrix} i_\alpha - \hat{i}_\alpha \\ i_\beta - \hat{i}_\beta \end{vmatrix} + \frac{\Omega\psi}{L} \begin{vmatrix} \sin\Gamma \\ -\cos\Gamma \end{vmatrix} - \begin{vmatrix} 1 & 0 \\ 0 & 1 \end{vmatrix} \begin{vmatrix} u_1 \\ u_2 \end{vmatrix}.$$

Исходная задача синтеза скользящего движения по пересечению поверхностей разрывов  $S_\alpha = 0$  и  $S_\beta = 0$  распадается на две независимые одномерные задачи. Каждая из них решается с учетом условия существования одномерного скользящего движения [3]. Для обеспечения скользящего движения по пересечению указанных поверхностей скользящие корректирующие воздействия  $u_1$  и  $u_2$  должны носить разрывной характер:

$$\begin{aligned} u_1 &= U_1 \operatorname{sgn} S_\alpha, \quad u_2 = U_2 \operatorname{sgn} S_\beta, \\ U_1 &\geq |-r(i_\alpha - \hat{i}_\alpha) + \Omega\psi \sin\Gamma|/L, \\ U_2 &\geq |-r(i_\beta - \hat{i}_\beta) + \Omega\psi \cos\Gamma|/L. \end{aligned}$$

В скользящем режиме рассогласования  $S_\alpha$  и  $S_\beta$  равны нулю, а эквивалентные значения корректирующих воздействий  $u_1$  и  $u_2$

$$u_{1eq} = -\Omega\psi \sin\Gamma, \quad u_{2eq} = \Omega\psi \cos\Gamma \quad (8)$$

несут информацию о частоте вращения и угловом положении ротора.

К сожалению, в данном случае эквивалентные значения каждого из корректирующих воздействий (8) несут информацию и об угловом положении, и о частоте вращения ротора. Поэтому необходимы дополнительные вычислительные операции для получения требуемой информации об угловом положении и частоте вращения ротора:

$$\Gamma = -\operatorname{arctg}(u_{1eq}/u_{2eq}), \quad (9)$$

$$\Omega = \sqrt{u_{1eq}^2 + u_{2eq}^2}/\psi. \quad (10)$$

Полученные оценки механических переменных требуют выделения эквивалентных значений корректирующих воздействий  $u_1$  и  $u_2$  (например, фильтром низких частот) и дальнейшей их обработки согласно выражениям (9) и (10).

Отметим, что требование сингулярной возмущенности системы в данном случае всегда выполнимо.

## ЗАКЛЮЧЕНИЕ

Выделен класс нелинейных динамических систем, для которых сформулировано достаточное условие оценивания неизмеряемых компонент вектора состояния по измеряемым компонентам. Это условие носит конструктивный характер, что открывает возможности для разработки методов синтеза алгоритмов оценивания на скользящих режимах. Предложенные методы оценивания переменных состояния базируются на адаптивных свойствах скользящих режимов и отличаются объемом вычислительных операций, необходимых для получения оценок. Развиваемый подход к решению задачи информационного обеспечения управления особенно эффективен при построении трехфазного электропривода без датчиков механических величин.

## ПРИЛОЖЕНИЕ

Доказательство утверждения 1. Если сформулированные в утверждении условия выполняются, то в исходной динамической системе (1) всегда можно выделить подсистему вида:

$$dx_m/dt = f_m(x_1, t) + \hat{D}(x_1, t)x_2(t) + B_m(x_1, t)u(t) \quad (\text{П. 1})$$

где  $x_m(t)$  — вектор состояния подсистемы,  $x_m(t) \in R^{n-q}$ ,  $x_1^T(t) = (x_m, x_m^*)$ ;  $f_m(x_1, t)$  — вектор-столбец объекта управления,  $f_1^T(x_1, t) = (f_m(x_1, t), f_m^*(x_1, t))$ ,  $f_m(x_1, t) \in R^{n-q}$ ;  $B_m(x_1, t)$  — матрица размерности  $(n-q) \times m$ ,  $B^T(x_1, t) = (B_m(x_1, t), B_m^*(x_1, t))$ .

Построим динамическую модель вида

$$d\hat{x}_m(t)/dt = f_m(x_1, t) + \hat{D}^*(t)v(t) + B_m(x_1, t)u(t), \quad (\text{П. 2})$$

где  $\hat{x}_m(t)$  — вектор состояния модели,  $\hat{x}_m(t) \in R^{n-q}$ ;  $\hat{D}^*(t)$  — квадратная матрица,  $\operatorname{rank} \hat{D}^*(t) = (n-q)$ ,  $v(t)$  — вектор разрывных управлений модели,  $v(t) \in R^{n-q}$ , каждая компонента которого  $v_i(t)$ ,  $i = \overline{1, (n-q)}$ , претерпевает разрыв на своей поверхности разрыва  $S_i(x_m, \hat{x}_m)$  в расширенном пространстве состояний системы и модели и принимает в зависимости от знака функции  $S_i(x_m, \hat{x}_m)$  значения  $v_i^+(x, t)$  и  $v_i^-(x, t)$ :

$$v_i(x, t) = \begin{cases} v_i^+(x, t), & \text{если } S_i(x_m, \hat{x}_m) > 0, \\ v_i^-(x, t), & \text{если } S_i(x_m, \hat{x}_m) < 0, \end{cases} \quad (\text{П. 3})$$

Потребуем, чтобы поведение модели (П. 2) соответствовало поведению подсистемы (П. 1), т. е. совпадали изменения поведения модельного  $\hat{x}_m(t)$  и системного  $x_m(t)$  векторов состояния. Это можно обеспечить

путем организации с помощью вектора разрывных управлений (П. 3) многомерного скользящего движения по пересечению поверхностей скольжения, представляющих собой нулевое рассогласование по компонентам модельного и системного векторов состояния:

$$S(x_m, \hat{x}_m) = \hat{x}_m - x_m = \varepsilon = 0. \quad (\text{П. 4})$$

Для доказательства попадания изображающей точки на пересечения (П. 4) из произвольных начальных условий и устойчивого движения по этому пересечению воспользуемся доказанной в работе [3] возможностью применения второго метода А. М. Ляпунова. Если производная непрерывно дифференцируемой положительной функции будет отрицательна всюду, кроме поверхности разрыва, где она не определена, то изображающая точка из любых начальных условий попадает на пересечение поверхностей разрывов и будет существовать скользящее движение по этому пересечению.

Выберем положительно определенную функцию Ляпунова в виде

$$W = \varepsilon^T \varepsilon / 2, \quad (\text{П. 5})$$

производная которой

$$dW/dt = \sum_{i=1}^{n-q} \varepsilon_i (d\varepsilon_i/dt). \quad (\text{П. 6})$$

Условие существования многомерного скользящего движения будет выполнено, если сумма (П. 6), будет отрицательна везде, кроме пересечения поверхностей разрывов. Производная вектора рассогласования  $\varepsilon$  в силу уравнения движения исходной системы (П.1) и динамической модели (П.2) рассчитывается следующим образом:

$$d\varepsilon/dt = -\hat{D}(x_1, t)x_2(t) + \hat{D}^* v(t). \quad (\text{П. 7})$$

Так как согласно постановке задачи, матрица  $\hat{D}^*(t)$  полного ранга, то для упрощения доказательства выберем ее единичной. В этом случае знак произведения  $i$ -й компоненты вектора рассогласования  $\varepsilon_i$  на ее производную будет всегда отрицательным, если алгоритм изменения разрывного управления

$$v_i = -V_i \text{sgn} S_i, \quad (\text{П. 8})$$

а модуль разрывного управления  $V_i$  выбирается в соответствии с условием

$$V_i = \left| \sum_{j=1}^{n-q} d_{ij}(t)x_2(t) \right|, \quad (\text{П. 9})$$

где  $d_{ij}(t)$  – элемент матрицы  $\hat{D}(x_1, t)$ .

Таким образом, при выполнении условий (П. 8) и (П. 9) для всех компонент вектора модельных управлений  $v(t)$  будет существовать многомерное скользящее движение по пересечению поверхностей разрыва  $S(x_m, \hat{x}_m) = \hat{x}_m - x_m = \varepsilon = 0$ . Само скользящее движение описывается с помощью эквивалентного управления  $v_{eq}$  [3]. Оно вычисляется в силу уравнения  $dS/dt = 0$  и представляет собой непрерывный аналог разрывного

управления  $v(t)$ , т. е. его усреднение в скользящем режиме. В данном случае уравнения  $dS/dt = 0$  имеет вид:

$$-\hat{D}(x_1, t)x_2(t) + v_{eq} = 0. \quad (\text{П. 10})$$

Векторно-матричное уравнение (П. 10) представляет собой систему  $(n - q)$  линейных уравнений с  $(n - q)$  неизвестными. Так как по постановке задачи матрица  $\hat{D}(x_1, t)$  невырожденная при всех  $x_1(t)$ , то система имеет единственное решение относительно вектора  $x_2$ :

$$x_2(t) = [\hat{D}(x_1, t)]^{-1} v_{eq}. \quad (\text{П. 11})$$

Таким образом, утверждение доказано, так как по известному значению эквивалентного управления  $v_{eq}$  и известным элементам квадратной матрицы  $\hat{D}(x_1, t)$ , определяемым с использованием измеренных значений компонент вектора состояния  $x_1$ , согласно выражению (П. 11) может быть вычислено значение вектора состояния  $x_2$ . ♦

Доказательство утверждения 2. Если матрица  $\hat{D}^*(t)$  имеет вид  $\hat{D}(x_1, t)$ , то тогда согласно выражению (П. 7) уравнение изменения ошибки  $\varepsilon$  представляется в виде:

$$d\varepsilon/dt = \hat{D}(x_1, t)(-x_2(t) + v(t)). \quad (\text{П. 12})$$

В скользящем режиме, когда  $de/dt = 0$

$$v_{eq} = x_2(t). \quad (\text{П. 13})$$

Докажем, что из произвольных начальных условий система (П. 1), (П. 2) попадет на пересечение (П. 4) и будет устойчиво двигаться по нему, а эквивалентное управление будет удовлетворять условию (П. 13). Воспользуемся новым вектором функций рассогласования  $S^{*T} = (S_1^*, \dots, S_{n-q}^*)$ , связанным линейным преобразованием с исходным вектором  $S$ :

$$S^* = R_S S, \quad (\text{П. 14})$$

где  $R_S = (\hat{D}(x_1, t))^T$  – матрица линейного преобразования. В силу линейности преобразования  $S = S^* = 0$ .

В этом случае выражение для вычисления производной (П. 6) от выбранной функции Ляпунова (П. 5) с учетом выражения (П. 14) может быть переписана в виде:

$$\begin{aligned} dW/dt &= \varepsilon^T (d\varepsilon/dt) = \\ &= (R_S^{-1} S^*)^T \hat{D}(x_1, t)(-x_2(t) + v(t)). \end{aligned} \quad (\text{П. 15})$$

С учетом выбранной матрицы линейного преобразования  $R_S$  производная функции Ляпунова (П. 15)

$$dW/dt = (S^*)^T (-x_2(t) + v(t)). \quad (\text{П. 16})$$

Если модуль разрывного управления  $V_i$  выбрать из условия

$$V_i \geq |v_{ieq}| = |S_i^* x_2^i|, \quad (\text{П. 17})$$



где  $x_2^i$  —  $i$ -я компонента вектора состояния  $x_2(t)$ , а алгоритм управления имеет вид:

$$v_i = -V_i \operatorname{sgn} S_i^*, \quad (\text{П. 18})$$

то условие существования многомерного скользящего движения (отрицательность производных функции Ляпунова (П. 16) везде, кроме пересечения поверхностей разрывов) будет выполнено, т. е. при любых начальных условиях скользящий режим по пересечению поверхностей разрывов  $S = S^* = 0$  будет иметь место.

В этом случае эквивалентное управление  $v_{eq}$  определяется из уравнения:

$$[\hat{D}(x_1, t)]^T [\hat{D}(x_1, t)] [-x_2(t) + v_{eq}(t)] = 0. \quad (\text{П. 19})$$

Векторное уравнение (П. 19) представляет собой систему  $(n - q)$  линейных уравнений с  $(n - q)$  неизвестными, причем каждое из них содержит только одно неизвестное, т. е.  $v_{eq}(t) = x_2(t)$ .

Утверждение доказано — значение каждой компоненты вектора эквивалентного управления  $v_{eq}$  является непосредственной оценкой значения одной из компонент вектора состояния  $x_2$ . ♦

Доказательство утверждения 3. Требуется доказать, что в системе (4), (П. 1)—(П. 4), (П. 17), (П. 18) порядка  $2(n - q)$  при любых начальных условиях возникает скользящий режим по пересечению поверхностей  $S = S^* = 0$ .

Рассмотрим положительно определенную функцию Ляпунова

$$W = \sum_{i=1}^{n-q} |\varepsilon_i| + \sum_{j=1}^{n-q} |V_j S_j^*| + \sum_{i=1}^{n-q} |x_2^i S_j^*|. \quad (\text{П. 20})$$

Вне поверхностей разрыва производная функции Ляпунова (П. 20) определена и вычисляется в силу уравнений (4), (П. 12), (П. 17), (П. 18):

$$\begin{aligned} dW/dt &= (d\varepsilon/dt)^T \operatorname{sgn} \varepsilon + V^T (dS^*/dt) + x_2^T (dS/dt) = \\ &= -(V + x_2)^T S^*, \end{aligned} \quad (\text{П. 21})$$

где  $V^T = (V_1 \operatorname{sgn} S_1^*, \dots, V_{n-q} \operatorname{sgn} S_{n-q}^*)$ ,  $(\operatorname{sgn} \varepsilon)^T = (\operatorname{sgn} \varepsilon_1, \dots, \operatorname{sgn} \varepsilon_{n-q})$ .

Из выражений (П. 20) и (П. 21) следует, что вне поверхностей разрыва производная функции Ляпунова в силу условия (П. 17) отрицательно определена и если траектория не будет пересекать поверхности разрыва или будет их пересекать в отдельных точках, то функция Ляпунова  $W$  будет монотонно убывать в соответствии с уравнением  $dW/dt + W = 0$ .

В случае возникновения скользящего режима на пересечении части поверхностей, т. е. условие  $S = S^* = 0$  выполняется только для части поверхностей

$$S^p = 0, \quad (\text{П. 22})$$

$$S^{*r} = 0, \quad (\text{П. 23})$$

где  $1 \leq p \leq n - q$  и  $1 \leq r \leq n - q$  — число поверхностей, на которых существует скользящий режим, движение системы описывается с помощью метода эквивалентного управления. Порядок расширенной системы уравнений (4), (П. 12) понижается путем исключения компонент векторов функций рассогласования  $\varepsilon$  и  $S^*$  в силу алгебраических уравнений (П. 22) и (П. 23), а вместо разрывных компонент векторов управлений  $v(t)$  и  $u(t)$  используются их эквивалентные значения, полученные в силу уравнений  $dS^p/dt = 0$ ,  $dS^{*r}/dt = 0$ .

В этом случае из функции Ляпунова (П. 20) могут быть исключены переменные, относительно которых скользящий режим существует. Таким образом, функция становится вновь дифференцируемой, и ее производная вычисляется в соответствии с укороченной системой уравнений (4), (П. 12)

$$\begin{aligned} dW/dt &= \sum_{\substack{i=1 \\ i \neq p}}^{n-q} [(d\varepsilon_i/dt) \operatorname{sgn} \varepsilon_i] + \sum_{\substack{j=1 \\ j \neq r}}^{n-q} (V_j \operatorname{sgn} S_j^* + x_2^j) = \\ &= -(V + x_2)^T S^*, \end{aligned}$$

где вектор  $V^T = (V^1, V^2)$ ,  $(V^1)^T = (V_1 \operatorname{sgn} S_1^*, \dots, V_{n-q-p} \operatorname{sgn} S_{n-q-p}^*)$ ,  $(V^2)^T = (v_{(n-q-p)eq}, \dots, v_{(n-q)eq})$ ,  $S^* \in R^{n-q-p}$ .

Очевидно, что и в этом случае производная функции Ляпунова отрицательна вне поверхностей разрыва и монотонно убывает, а расширенная система (4), (П. 12) асимптотически стремится к многомерному скользящему движению по пересечению поверхностей  $S = S^* = 0$ . ♦

## ЛИТЕРАТУРА

1. *Kvaternak H., Sivan R.* Linear optimal control systems. — N.-Y: John Wiley & Son Inc., 1972.
2. *Поляк Б.Т., Щербаков П.С.* Робастная устойчивость и управление. — М.: Наука, 2002.
3. *Уткин В.И.* Скользящие режимы в задачах оптимизации и управления. — М.: Наука, 1981.
4. *Краснова С.А., Уткин В.А., Михеев Ю.В.* Каскадный синтез наблюдателей состояния нелинейных многомерных систем // Автоматика и телемеханика. — 2001. — № 2. — С. 43—64.
5. *Шубладзе А.М.* Решение задачи «быстрой» идентификации с помощью многомерных скользящих режимов // Там же. — 1980. — № 2. — С. 72—78.
6. *Vitek J., Dodds S.J.* Forced dynamics control of electric drives. — Zilina: EDIS — Publishing Center of Zilina University, 2003.
7. *Рывкин С.Е., Изосимов Д.Б.* Алгоритмы идентификации механических координат электропривода // Электротехника. — 1994. — № 7. — С. 26—30.

☎ (495) 334-23-10, e-mail: rivkin@ipu.rssi.ru

Статья представлена к публикации членом редколлегии В.Ю. Рутковским. □

# ЗАКОНЫ РАСПРЕДЕЛЕНИЯ ОЦЕНОК ПАРАМЕТРОВ ДИНАМИЧЕСКИХ ОБЪЕКТОВ ПРИ ИДЕНТИФИКАЦИИ МЕТОДОМ ЭКСПОНЕНЦИАЛЬНОЙ МОДУЛЯЦИИ

Д.Н. Анисимов, А.В. Хрипков

Московский энергетический институт (технический университет)

Рассмотрен новый метод идентификации динамических объектов, получивший название метода экспоненциальной модуляции. Показано, что он прост в реализации и обладает высокой помехоустойчивостью. Рассмотрены вопросы, связанные с выбором оптимального входного сигнала. Определены законы распределения оценок при наличии помех типа «белый шум» и «синусоида со случайной фазой».

## ВВЕДЕНИЕ

Один из важных этапов проектирования систем автоматического управления заключается в идентификации динамических объектов. Существует немало методов идентификации, например, метод наименьших квадратов в различных модификациях, метод максимального правдоподобия, метод инструментальной переменной, метод модулирующих функций и др. У каждого из них свои достоинства и недостатки. Некоторые методы дают несмещенные оценки параметров объекта, но не эффективны в условиях высоких помех. Другие требуют, чтобы закон распределения помехи имел определенный вид. Иными словами, не существует универсальных методов идентификации, которые полностью удовлетворяли бы разработчика систем автоматического управления.

В статье рассмотрены некоторые аспекты использования метода экспоненциальной модуляции, впервые описанного в работе [1], для идентификации линейных динамических объектов, отличающегося высокой помехоустойчивостью и простотой реализации. Этот метод также не лишен недостатков. В частности, при наличии помех он дает смещенные оценки. Однако смещение, как правило, незначительно и к тому же может быть уменьшено с помощью специальных процедур.

## 1. МЕТОД ЭКСПОНЕНЦИАЛЬНОЙ МОДУЛЯЦИИ

Пусть дан устойчивый объект с передаточной функцией

$$W(s) = \frac{b_m s^m + \dots + b_1 s + b_0}{a_n s^n + \dots + a_1 s + 1} = \frac{Y(s)}{X(s)}, \quad (1)$$

параметры  $a_1, \dots, a_n, b_0, \dots, b_m$  которого необходимо определить,  $X(s)$  и  $Y(s)$  — изображения по Лапласу входного и выходного сигналов.

В момент времени  $t = 0$  на вход объекта подается произвольный сигнал  $x(t)$  и в течение времени наблюдения от 0 до  $T_H$  на выходе регистрируется сигнал  $y(t)$ . Сформируем функцию вида  $z(t) = \exp(-t/\theta)$  такую, чтобы к моменту времени  $T_H$  она затухала бы до значений, много меньших 1. Эту функцию будем называть модулирующей, а коэффициент  $\theta$  — постоянной времени модулирующей функции (ПВМФ). Перемножим входной и выходной сигналы на эту функцию и вычислим площади  $S_x$  и  $S_y$  под образованными таким образом кривыми:

$$S_x = \int_0^{T_H} x(t)z(t)dt, \quad S_y = \int_0^{T_H} y(t)z(t)dt.$$

Поскольку  $x(t)$  — реальный сигнал, подаваемый на объект, то естественно считать, что он ограничен на всем временном интервале. Тогда с учетом затухания модулирующей функции по экспоненциальному закону можно записать:

$$\begin{aligned} S_x &= \int_0^{\infty} x(t)z(t)dt - \int_{T_H}^{\infty} x(t)z(t)dt \approx \int_0^{\infty} x(t)z(t)dt = \\ &= \int_0^{\infty} x(t)\exp(-t/\theta)dt = X(\theta^{-1}), \\ S_y &= \int_0^{\infty} y(t)z(t)dt - \int_{T_H}^{\infty} y(t)z(t)dt \approx \int_0^{\infty} y(t)z(t)dt = \\ &= \int_0^{\infty} y(t)\exp(-t/\theta)dt = Y(\theta^{-1}) \end{aligned} \quad (2)$$



В этих выражениях  $X(\theta^{-1})$  и  $Y(\theta^{-1})$  — значения изображений по Лапласу входного и выходного сигналов, которые были бы получены при подстановке вместо оператора  $s$  вещественного числа  $\theta^{-1}$ . Как следует из выражений (1) и (2):

$$W(\theta^{-1}) = \frac{b_m \theta^{-m} + \dots + b_1 \theta^{-1} + b_0}{a_n \theta^{-n} + \dots + a_1 \theta^{-1} + 1} = \frac{S_y}{S_x}. \quad (3)$$

Выражение (3) устанавливает связь между задаваемой ПВМФ  $\theta$ , вычисляемыми площадями  $S_x, S_y$  и неизвестными параметрами объекта  $a_1, \dots, a_n, b_0, \dots, b_m$ . Очевидно, что для вычисления этих параметров необходимо задать  $r$  модулирующих функций с постоянными времени  $\theta_1, \dots, \theta_r$  и вычислить  $r$  пар площадей  $S_{x1}, \dots, S_{xr}, S_{y1}, \dots, S_{yr}$ , где  $r = n + m + 1$ . В этом случае параметры объекта могут быть определены из решения системы линейных уравнений

$$\mathbf{A} = \mathbf{\Gamma}^{-1} \mathbf{\Theta}, \quad (4)$$

где  $\mathbf{A} = [a_n \dots a_1 \ b_m \dots b_0]^T$ ,  $\mathbf{\Theta} = [-\theta_1^n \dots -\theta_r^n]^T$ ,

$$\mathbf{\Gamma} = \begin{bmatrix} 1 & \theta_1 & \dots & \theta_1^{n-1} & -d_1 \theta_1^{n-m} & \dots & -d_1 \theta_1^n \\ \dots & \dots & \dots & \dots & \dots & \dots & \dots \\ 1 & \theta_r & \dots & \theta_r^{n-1} & -d_r \theta_r^{n-m} & \dots & -d_r \theta_r^n \end{bmatrix}, \quad d_i = S_{xi}/S_{yi}.$$

Если идентификация проводится в идеальных условиях (время наблюдения неограниченно, шаг дискретизации бесконечно мал, помехи отсутствуют), погрешность определения параметров равна нулю независимо от вида входного сигнала и выбора ПВМФ. На практике ни одно из перечисленных условий не выполняется, причем основным источником погрешности является помеха  $\eta(t)$ , действующая на объект (будем считать, что она приложена аддитивно к выходу объекта). При ее наличии погрешность зависит как от входного сигнала, так и от постоянных времени модулирующих функций. Значения ПВМФ должны задаваться, исходя из следующих соображений. С одной стороны, их нельзя выбирать слишком малыми, поскольку значения вычисляемых площадей в этом случае будут малоинформативными из-за большого удельного веса случайной составляющей. С другой стороны, они не должны быть слишком большими, иначе модулирующие функции не успеют затухнуть к моменту окончания наблюдения процесса, приближенные равенства в выражении (2) будут несправедливы и при вычислении оценок появится регулярная погрешность. Очевидно, должен существовать некоторый диапазон значений ПВМФ, обеспечивающий лучшие в определенном смысле оценки.

Как показано в работе [2], при идентификации методом экспоненциальной модуляции оптимальным входным сигналом является ступенчатое воз-

действие. Такой сигнал обладает рядом преимуществ, а именно:

- отношение «сигнал/шум» максимально среди всех сигналов заданной амплитуды в предположении, что шум имеет нулевое математическое ожидание (при произвольном законе распределения), а объект является фильтром низких частот;
- появляется возможность компенсировать регулярные составляющие погрешностей оценок, и выбор ПВМФ при этом ограничен только снизу;
- значения площадей  $S_x$  в выражении (2) априорно известны, что снижает объем вычислений;
- при проведении активного эксперимента ступенчатый сигнал реализуется наиболее просто.

В настоящей статье при исследовании вероятностных характеристик оценок параметров объекта будем считать, что на его вход подается единичное ступенчатое воздействие.

## 2. РАСПРЕДЕЛЕНИЕ ОЦЕНОК ПАРАМЕТРОВ ОБЪЕКТА ПРИ НАЛИЧИИ БЕЛОГО ШУМА

Пусть дан объект второго порядка с передаточной функцией

$$W(s) = \frac{1}{a_2 s^2 + a_1 s + 1}$$

и по результатам наблюдения переходного процесса в интервале времени от 0 до  $T_n$  необходимо определить параметры  $a_1$  и  $a_2$ . Согласно методу экспоненциальной модуляции, необходимо сформировать модулирующие функции  $z_1 = \exp(-t/\theta_1)$  и  $z_2 = \exp(-t/\theta_2)$ , которые затухали бы к моменту  $T_n$  до значений, много меньших 1, умножить на них выходной сигнал и вычислить площади  $S_1$  и  $S_2$ .

Тогда оценки  $\hat{a}_1$  и  $\hat{a}_2$  согласно системе уравнений (4) могут быть определены следующим образом:

$$\hat{a}_1 = \frac{1}{\theta_1 - \theta_2} [\theta_1^2 (\theta_1/S_1 - 1) - \theta_2^2 (\theta_2/S_2 - 1)],$$

$$\hat{a}_2 = \frac{\theta_1 \theta_2}{\theta_1 - \theta_2} [\theta_1 (\theta_1/S_1 - 1) - \theta_2 (\theta_2/S_2 - 1)].$$

Пусть на выходе объекта действует равномерно распределенный в интервале  $[-A, A]$  белый шум  $\eta(t)$  с функцией плотности распределения

$$f(\eta(t)) = \begin{cases} \frac{1}{2A} & \text{при } \eta(t) \in [-A; A], \\ 0 & \text{при } \eta(t) \notin [-A; A]. \end{cases}$$

Математическое ожидание этого сигнала  $M(\eta) = 0$ , дисперсия  $D(\eta) = A^2/3$ .

Обозначим через  $R_i$  значения площадей, которые были бы вычислены в отсутствие шума,

$R_i = \frac{\theta_i^3}{a_2 + a_1\theta_i + \theta_i^2}$ , а через  $\Delta S_i = S_i - R_i$  — погрешности, обусловленные действием шума. Тогда

$$\Delta S_i \approx \int_0^{\infty} \eta(t) \exp(-t/\theta_i) dt.$$

Пусть съём и обработка информации происходит с шагом  $\Delta t$ , настолько малым, что погрешностями численного интегрирования можно пренебречь для любого метода интегрирования. Вычислим значения  $\Delta S_i$ , применяя метод трапеций:

$$\begin{aligned} \Delta S_i &= \Delta t \left( \eta_0/2 + \sum_{k=1}^{\infty} \eta_k \exp(-k\Delta t/\theta_i) \right) = \\ &= \Delta t \left( \eta_0/2 + \sum_{k=1}^{\infty} \eta_k q_i^k \right), \end{aligned}$$

где  $\eta_k = \eta(k\Delta t)$ ,  $q_i = \exp(-\Delta t/\theta_i)$ .

Из этой формулы следует, что математические ожидания величин  $\Delta S_i$  равны 0, т. е.  $M(S_i) = R_i$ . Определим ковариации погрешностей вычисления площадей:

$$\begin{aligned} \text{cov}(\Delta S_1, \Delta S_2) &= \\ &= \Delta t^2 M \left[ \left( \frac{\eta_0}{2} + \sum_{k=1}^{\infty} \eta_k q_1^k \right) \left( \frac{\eta_0}{2} + \sum_{k=1}^{\infty} \eta_k q_2^k \right) \right] = \\ &= D(\eta) \Delta t \frac{\theta_1 \theta_2}{\theta_1 + \theta_2}. \end{aligned}$$

Отсюда можно найти дисперсии величин  $\Delta S_i$ :

$$D(\Delta S_1) = D(\eta) \Delta t \theta_1/2, \quad D(\Delta S_2) = D(\eta) \Delta t \theta_2/2.$$

Эти значения представляют собой дисперсии бесконечной суммы независимых случайных величин, однако вклад каждой величины в суммарную дисперсию неодинаков. Поэтому, строго говоря, закон распределения величин  $\Delta S_i$  отличен от нормального. В качестве меры близости закона к нормальному можно рассмотреть коэффициенты разложения в ряд Маклорена функции  $\ln[g_i(\lambda)]$ , где  $g_i(\lambda)$  — характеристическая функция случайной величины  $\Delta S_i$  (для нормально распределенной случайной величины все эти коэффициенты, начиная с третьего, равны нулю [3]):

$$\begin{aligned} g(\lambda) &= \prod_{k=0}^{\infty} \int_{-Aq^k}^{Aq^k} \frac{1}{2Aq^k} \exp(j\lambda \Delta t \eta) d\eta = \\ &= \prod_{k=0}^{\infty} \frac{\sin(\lambda \Delta t A q^k)}{\lambda \Delta t A q^k}. \end{aligned}$$

$$\begin{aligned} \ln[g(\lambda)] &= \sum_{k=0}^{\infty} \ln \left[ \frac{\sin(\lambda \Delta t A q^k)}{\lambda \Delta t A q^k} \right] \approx \\ &\approx \frac{\theta}{\Delta t} \left[ -\frac{(\lambda \Delta t A)^2}{2 \cdot 6} - \frac{(\lambda \Delta t A)^4}{4 \cdot 180} - \frac{(\lambda \Delta t A)^6}{6 \cdot 2835} - \dots \right]. \end{aligned}$$

Видно, что коэффициенты разложения убывают достаточно быстро, поэтому ограничимся лишь первым членом разложения:

$$\ln[g_i(\lambda)] \approx -\frac{1}{2} \Delta t \theta_i \frac{(\lambda A)^2}{6}.$$

Данная характеристическая функция соответствует нормальному закону распределения  $\Delta S_i$  с нулевым математическим ожиданием и дисперсией

$$D(\Delta S_i) = \Delta t \theta_i A^2/6 = D(\eta) \Delta t \theta_i/2.$$

Таким образом, можно считать, что при идентификации методом экспоненциальной модуляции влияние равномерно распределенного белого шума эквивалентно влиянию нормально распределенного шума. В этом случае закон распределения вектора  $\mathbf{S} = (S_1, S_2)$  имеет вид:

$$\begin{aligned} f(S_1, S_2) &= \frac{\sqrt{c_{11}c_{22} - c_{12}^2}}{2\pi} \exp \left\{ -\frac{1}{2} [c_{11}(S_1 - R_1)^2 + \right. \\ &\quad \left. + 2c_{12}(S_1 - R_1)(S_2 - R_2) + c_{22}(S_2 - R_2)^2] \right\}, \end{aligned}$$

$$\text{где } c_{11} = \frac{D(\Delta S_2)}{D(\Delta S_1)D(\Delta S_2) - \text{cov}^2(\Delta S_1, \Delta S_2)},$$

$$c_{22} = \frac{D(\Delta S_1)}{D(\Delta S_1)D(\Delta S_2) - \text{cov}^2(\Delta S_1, \Delta S_2)},$$

$$c_{12} = -\frac{\text{cov}(\Delta S_1, \Delta S_2)}{D(\Delta S_1)D(\Delta S_2) - \text{cov}^2(\Delta S_1, \Delta S_2)}.$$

После подстановки в эти выражения значений дисперсий и ковариации найдем закон распределения вектора  $\hat{\mathbf{a}} = (\hat{a}_1, \hat{a}_2)$ , воспользовавшись формулой [3]

$$f(\hat{\mathbf{a}}) = f(S(\hat{\mathbf{a}})) |J(\hat{\mathbf{a}})|,$$

где — якобиан,

$$J(\hat{\mathbf{a}}) = \left| \frac{\partial(S_1, S_2)}{\partial(\hat{a}_1, \hat{a}_2)} \right|.$$

Обозначим  $\Delta a_i = a_i - \hat{a}_i$  — отклонения оценок параметров от истинных значений. После подстановок и преобразований получим

$$\begin{aligned} f(\hat{a}_1, \hat{a}_2) &= \frac{(\theta_1 + \theta_2)(S_1 S_2)^2}{2\pi D(\eta) \Delta t (\theta_1 \theta_2)^{7/2}} \times \\ &\times \exp \left[ -\frac{B}{2} (d_{11} \Delta a_1^2 + d_{12} \Delta a_1 \Delta a_2 + d_{22} \Delta a_2^2) \right], \quad (5) \end{aligned}$$





где  $B = \frac{(\theta_1 + \theta_2)^2}{D(\eta)\Delta t(\theta_1 - \theta_2)^2}$ ,

$$d_{11} = \frac{R_1^4}{\theta_1^5} + \frac{8(R_1 R_2)^2}{(\theta_1 \theta_2)^2(\theta_1 + \theta_2)} + \frac{R_2^4}{\theta_2^5},$$

$$d_{12} = \frac{2R_1^4}{\theta_1^6} + \frac{8(R_1 R_2)^2}{(\theta_1 \theta_2)^3} + \frac{2R_2^4}{\theta_2^6},$$

$$d_{22} = \frac{R_1^4}{\theta_1^7} + \frac{8(R_1 R_2)^2}{(\theta_1 \theta_2)^3(\theta_1 + \theta_2)} + \frac{R_2^4}{\theta_2^7}.$$

Плотности проекций вектора  $\hat{\mathbf{a}}$  на оси  $\hat{a}_1$  и  $\hat{a}_2$  определяются, соответственно, следующими выражениями:

$$f_1(\hat{a}_1) = \int_{-\infty}^{\infty} f(\hat{a}_1, \hat{a}_2) d\hat{a}_2, \quad (6)$$

$$f_2(\hat{a}_2) = \int_{-\infty}^{\infty} f(\hat{a}_1, \hat{a}_2) d\hat{a}_1. \quad (7)$$

На рис. 1 изображен график плотности проекции вектора  $\hat{\mathbf{a}}$  на ось  $\hat{a}_1$  при истинных значениях параметров объекта  $a_1 = 3$  с,  $a_2 = 2$  с<sup>2</sup>, ПВМФ  $\theta_1 = 1$  с,  $\theta_2 = 1,1$  с, шаге дискретизации  $\Delta t = 0,01$  с, соотношении «шум/сигнал» 0,25; на рис. 2 — график плотности проекции вектора  $\hat{\mathbf{a}}$  на ось  $\hat{a}_2$  при тех же условиях. На этих же рисунках изображены гистограммы соответствующих случайных величин, полученные по результатам 50 тыс. экспери-

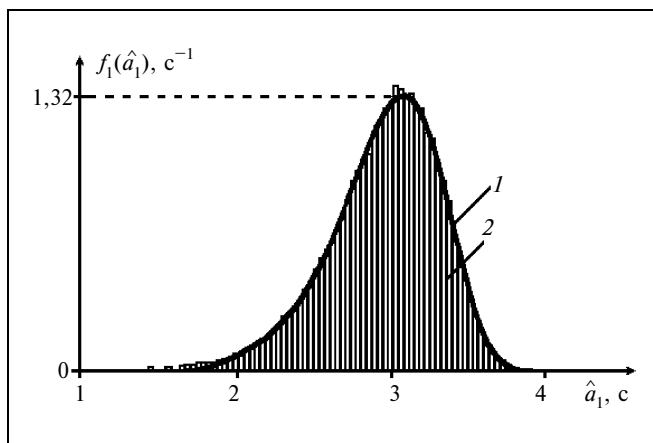


Рис. 1. График плотности проекции вектора  $\hat{\mathbf{a}}$  на ось  $\hat{a}_1$ , построенный в соответствии с выражениями (5) и (6) (кривая 1) и гистограмма, полученная экспериментально (ступенчатая кривая 2)

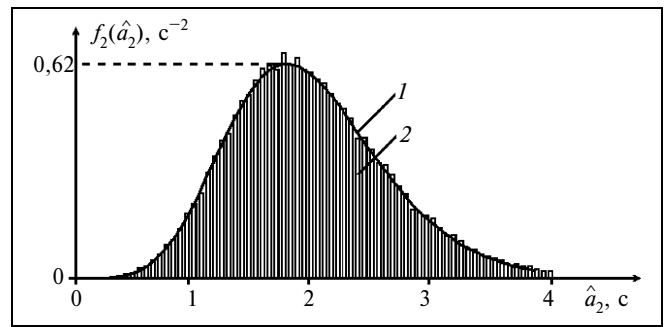


Рис. 2. График плотности проекции вектора  $\hat{\mathbf{a}}$  на ось  $\hat{a}_2$ , построенный в соответствии с выражениями (5) и (7) (кривая 1) и гистограмма, полученная экспериментально (ступенчатая кривая 2)

ментов. Видна близость результатов, полученных теоретически и экспериментально.

### 3. РАСПРЕДЕЛЕНИЕ ОЦЕНОК ПАРАМЕТРОВ ОБЪЕКТА ПРИ СИНУСОИДАЛЬНОЙ ПОМЕХЕ СО СЛУЧАЙНОЙ ФАЗОЙ

Для многих промышленных объектов характерно наличие помехи синусоидального вида, обусловленное обычно влиянием электрической сети переменного тока. Амплитуда и частота сигнала помехи, как правило, постоянные, а фаза является случайной величиной с равномерным распределением. Определим плотность распределения оценок для этого случая.

Пусть на вход объекта подается единичное ступенчатое воздействие, а к его выходу аддитивно приложено возмущающее воздействие  $\eta(t) = A \sin(\omega t + \varphi)$ , где  $A$  — амплитуда сигнала,  $\omega$  — его частота,  $\varphi$  — случайная фаза с плотностью распределения

$$f(\varphi) = \begin{cases} 1/2\pi & \text{при } -\pi \leq \varphi \leq \pi, \\ 0 & \text{в остальных случаях.} \end{cases}$$

Для простоты, без потери общности будем рассматривать объект первого порядка с передаточной функцией

$$W(s) = \frac{1}{as + 1}, \quad (8)$$

где  $a$  — неизвестный параметр объекта, который требуется оценить. В соответствии с методом экспоненциальной модуляции необходимо сформировать одну модулирующую функцию  $z(t) = \exp(-t/\theta)$ , умножить ее на выходной процесс  $y(t)$  и вычислить площадь  $S$  под образованной таким образом кривой.

Точное значение площади  $R$ , которое было бы вычислено в отсутствие помехи при ступенчатом входном воздействии, согласно выражениям (2) и (8), определяется как

$$R = \frac{\theta^2}{a + \theta}.$$

Рассмотрим распределение погрешности вычисления площадей  $\Delta S = S - R$ , обусловленной действием помехи. Легко показать, что

$$\begin{aligned} \Delta S_y &= \int_0^{\infty} \eta(t) \exp(-t/\theta) dt = \\ &= \int_0^{\infty} A \sin(\omega t + \varphi) \exp(-t/\theta) dt = \\ &= \frac{A\theta}{\sqrt{1 + (\omega\theta)^2}} \sin(\varphi + \arctg \omega\theta). \end{aligned}$$

Зная плотность распределения фазы  $\varphi$  и зависимость погрешности  $\Delta S$  от фазы, можно определить плотность распределения величины  $\Delta S$ :

$$f(\Delta S) = \frac{\mathbf{1}(\xi)}{\pi \sqrt{\frac{\theta^2 A^2}{1 + (\omega\theta)^2} - \Delta S^2}},$$

где  $\mathbf{1}(\xi) = \begin{cases} 0 & \text{при } \xi < 0, \\ 1 & \text{при } \xi \geq 0, \end{cases}$

$$\xi = \frac{A\theta}{\sqrt{1 + (\omega\theta)^2}} - |\Delta S|.$$

Обозначим  $\Delta a = a - \hat{a}$ ,  $n = a + \theta$ ,  $m = \frac{A(a + \theta)}{\theta \sqrt{1 + (\omega\theta)^2}}$ . Тогда

$$f(\Delta a) = \frac{\mathbf{1}(\xi)}{\pi |1 + \Delta a/n| \sqrt{m^2 (n + \Delta a)^2 - \Delta a^2}}. \quad (9)$$

Найдем область определения этой функции. Как следует из выражения (9):

$$m^2 (n + \Delta a)^2 - \Delta a^2 > 0 \text{ или } -m \leq \frac{\Delta a}{\Delta a + n} \leq m.$$

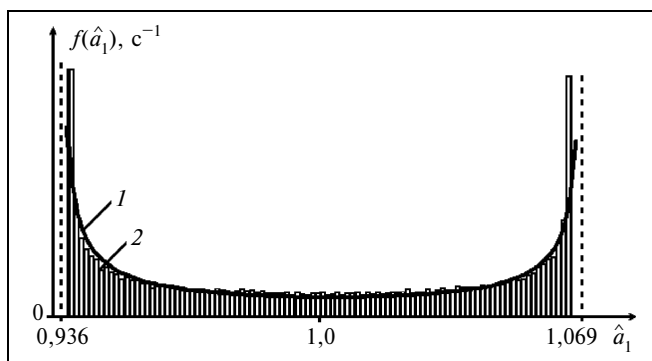


Рис. 3. График плотности распределения (кривая 1) и гистограмма (кривая 2) оценки параметра объекта при воздействии синусоидальной помехи со случайной фазой

В зависимости от значения  $m$  возможны три решения этого двойного неравенства:

- 1)  $\Delta a \in \left[ -\frac{mn}{1+m}, \frac{mn}{1-m} \right]$  при  $m < 1$ ;
- 2)  $\Delta a \in \left[ -\frac{n}{2}, \infty \right)$  при  $m = 1$ ;
- 3)  $\Delta a \in \left( -\infty, \frac{mn}{1-m} \right] \cup \left[ -\frac{mn}{1+m}, \infty \right)$  при  $m > 1$ .

Очевидно, второй случай возможен лишь при счетном множестве сочетаний значений  $A$ ,  $\omega$  и  $\theta$ , поэтому рассматривать его не имеет смысла. Третий случай является неблагоприятным, поскольку только правая часть области определения содержит желаемое значение  $\Delta a = 0$ . Что касается первого случая, то его всегда можно достичь правильным подбором значения  $\theta$ . Чем больше это значение, тем компактнее область определения функции плотности распределения и тем меньше разброс оценок параметра относительно его истинного значения. На рис. 3 изображены график плотности распределения и гистограмма оценки параметра  $\hat{a}$  при истинном значении  $a = 1$  с, ПВМФ  $\theta = 0,9$  с, соотношении «шум/сигнал» 0,15.

## ЗАКЛЮЧЕНИЕ

Рассмотрен новый метод идентификации динамических объектов, получивший название метода экспоненциальной модуляции. Показано, что он прост в реализации и обладает высокой помехоустойчивостью. Рассмотрены вопросы, связанные с выбором оптимального входного сигнала. Определены законы распределения оценок при наличии помех типа «белый шум» и «синусоида со случайной фазой». Совпадение полученных зависимостей и экспериментально построенных гистограмм свидетельствует о корректности проведенных расчетов.

## ЛИТЕРАТУРА

1. Анисимов Д.Н. Идентификация линейных динамических объектов методом экспоненциальной модуляции // Вестник МЭИ. — 1994. — № 2. — С. 68–72.
2. Анисимов Д.Н. Некоторые аспекты применения метода экспоненциальной модуляции для идентификации динамических объектов // Тр. междунар. конф. «Идентификация систем и задачи управления» SICPRO'03 / Ин-т пробл. упр. — М., 2003. — С. 1602–1618.
3. Пугачев В.С. Теория вероятностей и математическая статистика. — М.: Наука, 1979. — 496 с.

☎ (495) 362-74-07,

e-mail: AnisimovDN@mpei.ru, KhripkovAV@mpei.ru

Статья представлена к публикации членом редколлегии А.С. Манделем. □

# ИНТЕГРАЦИЯ И ГЛОБАЛИЗАЦИЯ ПРОГРАММНОГО ОБЕСПЕЧЕНИЯ ОБРАБОТКИ ДАННЫХ

М.В. Алексеевский

*РОСКОСМОС, филиал ФГУП ВНИИЭМ, г. Истра*

Предложен эффективный способ решения задачи глобализации и интеграции в единую информационную систему разобщенного корпоративного прикладного программного обеспечения, заключающийся в переходе к трехзвенной сетевой архитектуре доступа «Клиент → Провайдер → Сервер» с привлечением инструментальных программных средств технологии КОБРА/КОРИНФ.

## ВВЕДЕНИЕ

Корпоративные структуры (предприятия) являются активными потребителями современного программного обеспечения автоматизации сбора анализа и обработки компьютерной информации. Успешная деятельность крупных корпоративных структур и вовсе невозможна без программного обеспечения, относящегося к категориям автоматизированных информационных систем планирования и управления ресурсами предприятия и систем управления взаимоотношениями предприятий со своими партнерами и заказчиками.

Важное свойство корпоративного программного обеспечения состоит в его нестационарности, т. е. в процессе эксплуатации оно часто подвергается переработке, необходимой для его адаптации к внутренним корпоративным переменам и изменениям в корпоративном окружении. Например, в случае укрупнения предприятия и добавления к нему новых территориально нелокализованных подразделений возникает задача установки соединений доступа к корпоративным данным в режиме реального времени для пользователей, выпадающих за пределы ближайшего окружения, определяющего принадлежность к локальной компьютерной пользовательской сети, в которой соединение с источником данных гарантируется прямым кабельным подключением. Решение подобной за-

дачи путем преобразования существующего корпоративного программного обеспечения к состоянию, допускающему пользовательские обращения в реальном времени к удаленному источнику данных посредством подключений к глобальным сетевым ресурсам Интернета, будем условно называть глобализацией.

Под интеграцией будем понимать преобразование технологически неоднородного программного обеспечения, производящего пользовательские данные в строго разделенных интерактивных сеансах, к состоянию, допускающему синтезированную информационную отдачу в едином пользовательском сеансе соединения со множеством возможных источников корпоративных данных.

Вопросы глобализации и интеграции корпоративного программного материала исследуются крупнейшими мировыми производителями программного обеспечения, такими, как Microsoft, IBM, Sun Microsystems, Oracle и другими в концепции программной архитектуры SOA (Service oriented architecture) [1–5], рекомендуемой консорциумом стандартизации web-технологий W3C (см. <http://www.w3c.org>).

## 1. ГЛОБАЛИЗАЦИЯ И ИНТЕГРАЦИЯ ПО ТЕХНОЛОГИИ SOA

В концепции архитектуры SOA глобализация и интеграция корпоративного программного обеспечения может быть осуществлена путем преобра-

зования его в web-службы (сервисы). Тело web-сервиса представляет собой готовый программный модуль, загружаемый на выполнение и возвращающий полезные данные в ответ на пользовательский запрос через установленное HTTP соединение [6]. Полученные при этом данные моделируются в соответствии с запросом и XML инструкциями и укладываются в web-контейнер для передачи на сторону клиента, в общем случае через Интернет, для анализа и обработки в окне стандартного web-браузера. Тело web-сервиса обеспечивается специальным файлом WSD (Web Service Definition), составленным на XML подобном языке WSDL (Web Service Description Language) и содержащим необходимую информацию по конечному развертыванию пользовательского интерфейса для доступа к удаленным сервисным процедурам [1–5].

Итак, если исходить из того, что корпоративное программное обеспечение является множеством готовых программных модулей, способных производить полезные данные в режиме выполнения, то каждый такой модуль может быть преобразован в web-сервис с гарантированным для него файлом WSD, независимо от того, какими инструментальными программными средствами он был получен. Следовательно, в соответствии с концепцией SOA, можно избежать необходимости полной переработки неоднородного семейства корпоративных программ в целях обеспечить к нему широкий пользовательский доступ в масштабе Интернета и обобщенную, интегрированную функциональность.

В рамках концепции SOA глобализация и интеграция достигаются встраиванием в зону звена провайдера трехзвенной глобальной сетевой архитектуры доступа (Клиент → Провайдер → Сервер) готовых программных модулей, с заданной функциональностью. Проблема решается путем размещения готового программного материала в контейнерах сервера приложений, установленных на провайдере, с подключением контейнеров присоединенных внешних модулей режима гипертекстовой обработки данных перед их перемещением на сторону клиента. Целесообразность такого подхода обосновывается открывающейся возможностью повторного использования готового программного кода в зарезервированных позициях адекватного проблеме технологического шаблона.

Общая схема реализации условий глобализации и интеграции прикладного корпоративного программного обеспечения, согласно концепции SOA, показана на рис. 1. Программные модули, определяющие проектную функциональность и установленные на провайдере, представляют готовую часть информационной системы и обеспечивают производство полезных данных путем соединения со звеном сервера трехзвенной архитектуры дос-

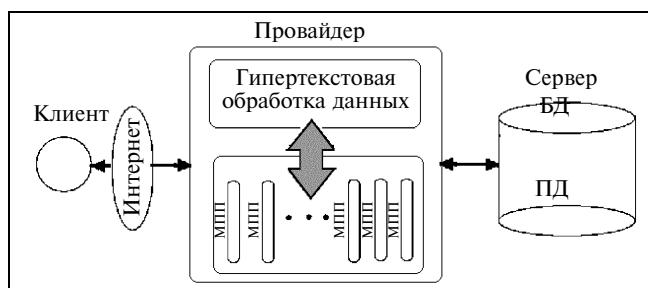


Рис. 1. Глобализация и интеграция в концепции SOA: МПП – модуль пользовательского приложения, БД – база данных, ПД – полезные данные

тупа. Гипертекстовая обработка данных (XML, HTML) является предметом дополнительного программирования и преследует две цели:

- обеспечение сообщения между клиентом и провайдером по эффективному протоколу HTTP передачи данных в масштабе Интернета;
- обеспечение при помощи стандартных инструментальных программных средств — web-браузеров — представления, анализа и обработки полезных данных.

Однако вполне очевидна неэкономичность такого решения проблем глобализации и интеграции, определяющего высокие темпы расходования оперативных ресурсов провайдера. Действительно, процедурные коды, интегрируемые в зону провайдера, выполняют подключения к базам данных, пользовательские запросы по выделению и моделированию данных в режиме параллельной обработки клиентских обращений, непрогнозируемая частота которых может стать причиной чрезмерности запрашиваемых провайдером ресурсов. В концепции SOA никакие меры по сдерживанию темпов расходования провайдером системных ресурсов не предусматриваются.

Встраиваемые модули воспринимаются в общем смысле готовности к загрузке и выполнению и обеспечиваются экспресс-механизмом реализации такой готовности в ответ на пользовательские обращения. Повышение экономичности режима выполнения параллельных пользовательских транзакций путем переработки процедурного кода может быть ощутимым разве что благодаря сужению суммарной функциональности интегрированных программных модулей, а привлекательность SOA-технологии при этом заметно снизится, ибо процедурное (алгоритмическое) программирование весьма трудоемко и требует подробного тестирования измененного кода, включая компиляцию и компоновку.

Следовательно, тактика встраивания набора готового процедурного кода в рамки одной компью-



терной операционной системы для выполнения в режиме параллельных пользовательских транзакций не вполне оправдана. Очевидно, что при таком подходе, из-за нехватки ресурсов, допустимая интенсивность клиентских обращений к провайдеру может оказаться недостаточной для удовлетворения корпоративных потребностей в информационном обеспечении.

## 2. ГЛОБАЛИЗАЦИЯ И ИНТЕГРАЦИЯ ПО ТЕХНОЛОГИИ КОБРА/КОРИНФ

При самых общих предположениях о составе корпоративного программного обеспечения рассматриваемого как произвольный набор подготовленных к загрузке и выполнению программных модулей, с заданной функциональностью по производству полезных данных, технология SOA решает вопрос о его глобализации и интеграции однозначно и безальтернативно. Однако, как показано в § 1, такое решение может оказаться неэффективным в силу того, что отведенные провайдеру ресурсы могут быть исчерпаны задолго до достижения наперед планируемой интенсивности клиентских обращений. В концепции, основу которой составляет технология КОБРА/КОРИНФ [7], глобализация и интеграция корпоративного программного обеспечения решается при дополнительном предположении о структуре модулей из обрабатываемого состава.

Предположение состоит в том, что модули, подлежащие глобализации и интеграции, являются приложениями, использующими соединения с базами данных при помощи стандартных SQL-запросов и инструкций. Исходя из мирового опыта автоматизированного компьютерного информационного строительства, можно утверждать, что работа подавляющего большинства программ управления корпоративными данными основана на взаимодействии с системами управления и обработки сетевых сообщений, удовлетворяющих стандарту SQL [8]. Это подтверждается тем, что первый стандарт языка SQL был принят ANSI (Американским национальным институтом стандартизации) в 1986 г. и ISO (Международной организацией по стандартизации) в 1987 г., а ведущими продуктами самых крупных мировых производителей программного обеспечения — IBM, Microsoft, Oracle — уже много лет являются системы управления базами данных (DB2, MS SQL, ORACLE), работа которых построена на обработке и выполнении SQL-инструкций. Итак, суть сделанного предположения состоит в том, что если не принимать во внимание программную реализацию графического пользовательского интерфейса, то полезная функциональность корпоративных информационных проектов

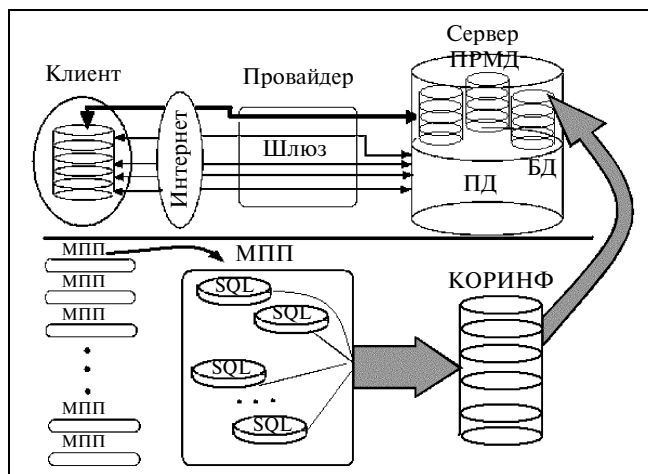


Рис. 2. Глобализация и интеграция в концепции КОБРА/КОРИНФ:

МПП — модуль пользовательского приложения, ПРМД — проектные метаданные, БД — база данных, ПД — полезные данные

обеспечивается в основном путем выполнения встроенных в процедурный код SQL-запросов и инструкций по обработке данных. Следовательно, не умаляя общности постановки проблемы предположением о том, что можно пренебречь программным обеспечением, не поддерживающим стандарт SQL, представим решение задачи глобализации и интеграции, основанное на использовании технологии непроектного программирования КОБРА/КОРИНФ [7].

Общая схема решения поставленной задачи приведена на рис. 2. Введем в рассмотрение набор прикладных программных модулей, каждый из которых в режиме открытого сеанса соединения с источником корпоративных данных в ответ на пользовательский ввод генерирует SQL-сообщения с инструкциями по манипулированию информационными объектами. Каждое встроенное SQL-обращение, в общем случае параметризованное и готовое к выполнению при условии замещения параметров фактическими значениями, может быть приведено к одному из трех видов (фильтры, консоли и отчеты) программируемых объектов языка КОРИНФ и включено в состав проектного дерева [7]. Путем несложного перевода в объекты языка КОРИНФ всех встроенных в преобразуемый программный модуль SQL-обращений и объединением их в общее проектное дерево получим в итоге конвертированный проект КОРИНФ, который с большой точностью воспроизводит функциональность отправного программного модуля. Поступая подобным образом со всеми модулями преобразуемого семейства, получим в результате семейство проектов КОРИНФ, которое в полной мере

способно обеспечить востребованную проектную функциональность. Для конвертированного семейства проектов КОРИНФ адекватных по функциональности исходному семейству программных модулей вопрос о глобализации и интеграции решается автоматически.

Обладая свойством мобильности [7], проекты КОРИНФ могут запоминаться и сохраняться в базе данных сервера, перемещаться в ответ на пользовательский запрос через Интернет и загружаться в браузер КОБРА клиента на выполнение. Следовательно, свойство мобильности, обеспечивая автоматическую загрузку проекта через Интернет на компьютер клиента, создает там же условия для компактного и монопольного режима выполнения проекта, освобождая провайдер от необходимости обрабатывать проектную функциональность в режиме параллельных пользовательских транзакций, как это происходит при использовании архитектуры SOA (см. рис. 1). Провайдер (см. рис. 2) эксплуатируется в режиме простого шлюзового приложения — брокера [7] — единого для проектов КОРИНФ вообще, выполняющего постоянные диспетчерские функции по управлению соединениями между клиентом и сервером и поэтому расходующего предельно малое количество компьютерных ресурсов.

Таким образом, вопрос глобализации и интеграции автоматически решается установкой метаданных конвертированного семейства проектов КОРИНФ в зоне сервера (см. рис. 2) (вместе с полезными данными), которые с помощью браузера КОБРА сразу становятся доступными каждому удаленному в масштабе Интернета привилегированному пользователю. Метаданные проекта, образуя набор программируемых текстовых модулей и управляющих текстовых файлов, как и полезные данные, загружаются на сторону клиента с помощью SQL-запроса, формируемого браузером КОБРА в ответ на выделение пользователем проекта из списка ему доступных в соответствии с заявленными привилегиями. После завершения загрузки метаданные проекта представляются в браузере КОБРА в виде проектного дерева. Выделение узла дерева (мышью или с клавиатуры) приводит соответствующий ему проектный модуль в состояние активности, обеспечивая подготовку SQL-запросов и инструкций и передачу их на сервер для согласованного чтения и обработки данных. В случае возврата данных полученный набор, без использования громоздкого XML моделирования, преобразуется в обычную строку с разделителями по выделению табличных ячеек, передается клиенту по протоколу HTTP [6] и восстанавливается в режиме электронной таблицы для анализа и обработки.

Если установкой метаданных конвертированного семейства проектов КОРИНФ в зоне сервера вопрос глобализации решается полностью, то вопрос интеграции можно считать решенным в первом приближении. Действительно, благодаря структурной однородности проектов КОРИНФ удаленный пользователь, при наличии у него достаточных привилегий, по собственному выбору загружает доступные ему проекты в браузер КОБРА с той же легкостью, с какой web-страница загружается пользователем в стандартный web-браузер. Следовательно, в браузере КОБРА стягивается в единый список разобщенное прежде семейство корпоративных программ, конвертированное в однородное семейство проектов КОРИНФ, представляемых в браузере многоуровневыми древовидными структурами модулей программируемых объектов — консолей, фильтров и отчетов.

Возможно, что первичное распределение модулей по проектам будет неоптимальным по информационной отдаче. Например, некоторый блок корпоративных данных, обеспечиваемый модулями одного из проектов семейства, может быть востребован в контексте другого проекта того же семейства, в общем случае с соблюдением отношений ссылочной целостности. Непроцедурное программирование на языке КОРИНФ, которое существенно проще процедурного (алгоритмического), позволяя полностью автоматизировать процесс составления разработчиком программных кодов [7], обеспечивает также возможность быстрого перераспределения модулей между проектами. Перераспределение может быть обеспечено как путем исключения модулей из одного проекта и перемещения их в другой, так и простым дублированием в разных проектах одной и той же группы востребованных модулей.

Следуя путем целенаправленного вторичного распределения модулей, конвертированное семейство проектов КОРИНФ, благодаря своим прогрессивным эксплуатационным свойствам, может полностью удовлетворить требованию интеграции корпоративного программного обеспечения. Следует также принять во внимание, что браузер КОБРА автоматически управляет ресурсами загруженного на выполнение проекта, а сам проект в объеме (по числу и составу модулей) практически неограничен, поскольку оперативные компьютерные ресурсы расходуются только на обслуживание активных модулей проекта.

Полное описание элементов языка программирования КОРИНФ и инструментальных программных средств разработки и эксплуатации проектов на этом языке содержится в монографии [7].



## ЗАКЛЮЧЕНИЕ

Обобщая сделанные выводы, выделим два основных показателя эффективности способов глобализации и интеграции корпоративного программного обеспечения. Этими показателями являются:

- производительность способа глобализации и интеграции корпоративного программного обеспечения;
- производительность конечного продукта в режиме его эксплуатации в зависимости от способа его получения путем глобализации и интеграции.

Получение конечного результата глобализации и интеграции семейства программ, производящих полезные данные для корпоративных клиентов, при проведении мероприятий согласно архитектуре SOA, требует процедурного (алгоритмического) программирования пользовательского интерфейса. Притом, что такое программирование наиболее эффективно и обеспечено средствами автоматизации либо для платформы J2EE (servlets, JSP) [1–3], либо для .NET (ASP.NET) [4, 5] в сочетании с XML, следует признать вхождение в него значительной доли трудоемкого ручного введения программных кодов. Напротив, конечный результат глобализации и интеграции, при использовании технологии КОБРА/КОРИНФ, достигается полностью автоматизированным составлением непроцедурных кодов при помощи рассмотренного в работе [5] дизайнера, с учетом того, что в распоряжении разработчика имеются готовые SQL-запросы и инструкции, извлеченные из отправных программных модулей. Вполне очевидно, что производительность непроцедурного программирования существенно выше процедурного, отягченного необходимостью слежения за правильным выделением и освобождением системных ресурсов, компиляции, компоновки и инсталляции в систему перед вводом в эксплуатацию результирующего кода. Впрочем, управление системными ресурсами для платформ J2EE и .NET выполняется автомати-

чески с помощью менеджеров сопровождения режима выполнения JVM [1–3] или CLR [4, 5], что, однако, снижает производительность информационной системы в целом.

Производительность конечного продукта глобализации и интеграции в архитектуре SOA терпит большие потери в силу того, что эксплуатация интегрированного семейства проектов реализуется в зоне провайдера в режиме параллельных пользовательских транзакций. Полный объем проектных вычислений в рамках одной транзакции может существенно ограничить допустимую интенсивность пользовательских обращений к провайдеру и удерживать её на уровне ниже запланированного. Напротив, в условиях применения технологии КОБРА/КОРИНФ, зона провайдера полностью разгружена от проектной функциональности и минимизирована по уровню потребления компьютерных ресурсов, чем обеспечивается оптимальная производительность конечного продукта в режиме эксплуатации.

## ЛИТЕРАТУРА

1. *Фанг Дж.* Введение в IBM Rational Application Developer (+ CD-ROM). — М.: Кудиц-образ, 2006. — 592 с.
2. *Ньюкомер Э.* Веб-сервисы. Для профессионалов. — СПб.: Питер, 2003. — 256 с.
3. *Хабибуллин И.Ш.* Разработка web-служб средствами Java. — СПб.: БХВ-Петербург, 2003. — 400 с.
4. *Шорт С.* Разработка XML web-сервисов средствами Microsoft .NET (+ CD-ROM). — СПб.: БХВ-Петербург, 2003. — 480 с.
5. *Шапошников И.В.* Web-сервисы Microsoft .NET. — СПб.: БХВ-Петербург, 2002. — 334 с.
6. *Семенов Ю.А.* Протоколы Интернет. — М.: Горячая линия — Телеком, 2005. — 1100 с.
7. *Алексеевский М.В.* Технология программирования КОБРА/КОРИНФ и корпоративная интернет-информатика. — М.: Изд-во МЭИ, 2006. — 169 с.
8. *Грабер М.* SQL. Описание SQL92, SQL99 и SQLJ. — М.: Лори, 2003. — 752 с.

☎ (495) 994-54-88, e-mail: [alemivabol@istranet.ru](mailto:alemivabol@istranet.ru)

Статья представлена к публикации членом редколлегии В.Л.Энштейном. □



Подписку на журнал «Проблемы управления» можно оформить с любого месяца в любом почтовом отделении (подписной индексы 80508 и 81708 в каталоге Роспечати или 38006 в объединенном каталоге «Пресса России»), а также через редакцию. Отдельные номера редакция высылает по первому же требованию.

**Подписка через редакцию — выгодное предложение для Вас!** Подписаться через редакцию можно с любого месяца, здесь можно приобрести любой номер журнала за прошедшие годы. Позвоните в редакцию по тел. (495) 330-92-00 или пришлите Ваш адрес по электронной почте [ru@iri.ru](mailto:ru@iri.ru) — и подписка будет оформлена за один рабочий день. Почтовые расходы по пересылке журнала редакция берет на себя.

**Оформить подписку** можно и через группу компаний «Урал-Пресс» (см. <http://www.ural-press.ru/>) и агентство «Коммерсант-Курьер» (см. [www.komsvr.ru](http://www.komsvr.ru)), которые проводят подписку на наш журнал во многих городах России.

*Не забудьте подписаться!*

# РАНЖИРОВАНИЕ WEB-СТРАНИЦ С ИСПОЛЬЗОВАНИЕМ ВЗАИМНОЙ ИНФОРМАЦИИ МЕЖДУ ГИПЕРССЫЛКАМИ

Р.М. Алгулиев, Р.М. Алыгулиев

*Институт информационных технологий Национальной академии наук Азербайджана, г. Баку*

Для повышения эффективности ранжирования web-страниц предложены три модификации алгоритма PageRank. Особенность первой из них состоит в измерении степени независимости гиперссылок, на основе которой определяется вес гиперссылки. Вторая и третья модификации, учитывающие тематическую близость web-страниц, представляют собой усовершенствованные варианты алгоритмов WPR и Topic-Centric, соответственно.

## ВВЕДЕНИЕ

Появление в конце прошлого столетия World Wide Web (WWW) сделало Интернет одним из основных источников информации. Сегодня WWW — это динамично изменяющаяся среда, а представленные в ней информационные ресурсы крайне разнородны. Распределенный характер WWW сильно затрудняет поиск нужной информации среды разнообразных материалов, охватывающих самые разные сферы человеческой деятельности. При поиске нужной информации в Web поисковым машинам приходится охватывать огромное число связанных гиперссылками страниц. По мере увеличения web-ресурсов и рассредоточения их источников поиск нужной информации в Интернете становится еще более трудоемким. При этом следует разрабатывать такие технологии и подходы, которые отвечали бы увеличивающимся потребностям пользователей. Один из таких подходов — это усовершенствование существующих технологий информационного поиска.

Поисковые машины по их методам индексирования делятся на два поколения. В поисковых машинах первого поколения, разработанных на ранних стадиях создания Web, для ранжирования web-страниц была использована частота слов или мера подобия, т. е. были использованы традиционные методы поиска документов. В отличие от обычных текстовых документов web-страницы имеют ряд специфических особенностей. В гипертекстовой среде носителями информации, кроме

web-страниц, являются и гиперссылки. Поэтому для улучшения поисковой точности поисковые машины второго поколения как дополнительный источник информации используют гиперссылки. Поисковые машины, следуя по гиперссылкам, посещают огромное web-пространство, в результате которого собирают дополнительную информацию о web-страницах. Потом эта информация ими используется при ранжировании web-страниц. Поэтому в последние годы многие исследования были посвящены анализам гиперссылок.

На сегодня разработаны некоторые алгоритмы для решения упомянутой проблемы. Среди них наиболее популярны алгоритмы PageRank [1, 2], HITS (Hypertext Induced Topic Search) [3] и SALSА (Stochastic Approach for Link Structure Analysis) [4]. Известные поисковые машины — Google, Yahoo и др. — в той или иной степени используют эти алгоритмы. Алгоритмы PageRank и HITS ранг web-страниц вычисляют итеративно. Алгоритм HITS с помощью обычных методов информационного поиска сначала идентифицирует web-страницы, релевантные запросу пользователя, а затем упорядочивает их таким образом, чтобы самые релевантные web-страницы были представлены в верхней части списка. Другой алгоритм PageRank ранжирует целые связанные гиперссылками web-страницы, затем среди них выбирает страницы, релевантные запросу пользователя, сохраняя при этом их ранги. Алгоритм SALSА является комбинацией алгоритмов PageRank и HITS.





Настоящая статья посвящена усовершенствованию алгоритма PageRank, где предлагаются три модификации.

## 1. КРАТКИЙ ОБЗОР АЛГОРИТМОВ PAGERANK И HITS

### Основные обозначения:

$G(W, E)$  — ориентированный web-граф;  
 $W$  — множество web-страниц;  
 $E$  — множество гиперссылок;  
 $u \rightarrow v$  — гиперссылка из страницы  $u$  в страницу  $v$ ;  
 $B(u) = \{v : v \rightarrow u\}$  — множество страниц, которые ссылаются на страницу  $u$  (Backwards links);  
 $F(u) = \{v : v \rightarrow u\}$  — множество страниц, на которые ссылается страница  $u$  (Forwards links);  
 $n$  — общее число страниц в web-графе  $G(W, E)$ ;  
 $PR(u)$  — PageRank страницы  $u$ ;  
 $A(u)$  — ранг страницы  $u$  как «авторитет»;  
 $H(u)$  — ранг страницы  $u$  как «концентратор»;  
 $|U|$  — мощность множества  $U$ ;  
 $\text{sim}(u, v)$  — мера близости страниц  $u$  и  $v$ , которая определяется метрикой косинуса;  
 $\omega(v \rightarrow u)$  — вес гиперссылки  $u \rightarrow v$ ;  
 $r(q, t)$  — степень релевантности запроса  $q$  тематике  $T$ .

Для извлечения информации из структуры гиперссылок и перекрестных ссылок, отслеживания дефектов их структуры, анализа связей между ссылками и объектами ссылок (Web Structure Mining) широко применяется алгоритм PageRank. Это статический алгоритм, предназначенный для оценки качества страниц, не зависящий от каких-либо запросов, т. е. с его помощью вычисляется «глобальная значимость» страниц. Суть алгоритма заключается в следующем. Представьте себе случайного пользователя, перемещающегося по Web. Пусть пользователь посещает страницу  $v$ . На каждом шаге пользователь либо «перепрыгивает» на другую страницу в Web, выбранную случайным образом, либо он следует по гиперссылке на текущей странице, при этом не возвращаясь и не посещая одну и ту же страницу дважды. Если через  $(1 - d)$  обозначить вероятность случайного прыжка, то вероятность перехода по ссылке будет  $d$ . Таким образом, показатель PageRank страницы  $u$  можно вычислить по следующей рекурсивной формуле:

$$RP(u) = \frac{(1-d)}{n} + d \sum_{v \in B(u)} \frac{1}{|F(v)|} RP(v). \quad (1)$$

В правой части  $1/n$  соответствует тому, что среди  $n$  страниц каждая страница выбирается с одинаковой вероятностью. Здесь также предпо-

лагается, что исходящие ссылки на странице  $v$  выбираются с одинаковой вероятностью, равной  $1/|F(v)|$ . Для сходимости процесса (1) вероятность  $d$  (ее называют коэффициентом демпфирования [1, 2]) выбирается из интервала  $d \in [0,8; 1]$ . Из алгоритма (1) видно, что чем больше ссылок на страницу, тем она становится «важнее».

Алгоритм HITS [3], как и алгоритм PageRank, основан на анализе web-структуры, но в отличие от него у каждой страницы выделяются две роли: роль «авторитета» (authority) и роль «концентратора» (hub). Алгоритм HITS, анализируя входящие и исходящие гиперссылки, ранжирует web-страницы. По этому алгоритму страница, на которую ссылаются другие страницы, называется «авторитетом», а страница, которая ссылается на другие страницы, называется «концентратором»:

$$A(u) = \sum_{v \in B(u)} H(v), \quad (2)$$

$$H(u) = \sum_{v \in F(u)} A(v). \quad (3)$$

Цель алгоритма HITS заключается в поиске наиболее качественных «авторитетов» и наиболее качественных «концентраторов». Из формул (2) и (3) видно, что для каждой страницы алгоритм HITS вычисляет два ранга: ранг  $A(u)$ , показывающий качество страницы как «авторитета», и ранг  $H(u)$ , показывающий качество страницы как «концентратора». Как и в алгоритме PageRank, в первом приближении рангам страниц присваивается произвольное ненулевое значение и затем производится итерационный процесс, состоящий из последовательного применения операций (2) и (3).

Несмотря на то, что поисковые машины второго поколения достигли более высокой точности, чем машины первого поколения, в дальнейшем их эффективность снизилась, были выявлены уязвимости перед недобросовестными методами манипулирования рейтингом (спамдексингом). Причина состояла в том, что в первоначальных вариантах алгоритмов степень значимости страниц определялась числом входящих гиперссылок. Другими словами, при вычислении показателя значимости страниц эти алгоритмы не учитывали тематической близости страниц. В настоящее время перспективно направление, связанное с комбинированным учетом информации о гиперссылочной связности web-страниц и результатов контентного анализа этих страниц.

Для улучшения качества ранжирования предложены некоторые модифицированные варианты алгоритмов PageRank и HITS. Далее приводятся некоторые модифицированные варианты алгоритма PageRank.

Например, при ранжировании web-страниц алгоритм Topic-Centric [5] учитывает меру близости страниц:

$$PR(u) = \frac{(1-d)}{n} + d \sum_{v \in B(u)} \frac{\text{sim}(u, v)}{\sum_{x \in F(v)} \text{sim}(x, v)} PR(v). \quad (4)$$

Согласно определению,  $\text{sim}(u, v) = 1$  соответствует максимальной близости страниц  $u$  и  $v$ , а  $\text{sim}(u, v) = 0$  — их полному различию.

В алгоритме TSPR (Topic-Sensitive PageRank) [6] web-страницы сначала группируются по тематике, а потом на каждом тематическом разделе вычисляется ранг страницы. Тематические разделы отбираются из верхнего уровня ODP (Open Directory Project). Пусть  $U_k$  означает множество URL-адресатов страниц в тематическом разделе  $T_k$ . При вычислении PageRank вектора для тематического раздела  $T_k$  алгоритм TSPR предполагает, что случайный пользователь двигается (следует по гиперссылкам или «прыгает») только по страницам из множества  $U_k$ , т. е. не выходит за рамки тематики. Тогда показатель PageRank страницы  $u$  на тематическом разделе  $T_k$  будет определен так:

$$PR_k(u) = d \sum_{v \in B(u)} \frac{PR_k(v)}{|F(v)|} + \begin{cases} (1-d)/|U_k|, & \text{если } u \in U_k, \\ 0, & \text{если } u \notin U_k. \end{cases}$$

Тогда зависящий от запроса показатель значимости (query sensitive importance score) страницы  $u$

$$S_q(u) = \sum_k PR_k(u) r(q, T_k).$$

Подчеркнем, что результат поиска ранжируется с учетом этих счетов.

В модели Intelligent Surfer [7] для вычисления ранга страницы  $u$ , зависящего от запроса  $q$ , алгоритм PageRank преобразуется к следующему виду:

$$PR_q(u) = (1-d) \frac{r(q, u)}{\sum_{x \in W} r(q, x)} + d \sum_{v \in B(u)} \frac{r(q, u)}{\sum_{y \in F(v)} r(q, y)} PR_q(v).$$

В моделях Topical PageRank и Topical HITS [8] каждой странице  $u$  сопоставляется два вектора: контент-вектор  $C(u)$  и авторитет-вектор  $A(u)$ .

Контент-вектор  $C_u$ :  $[C(u^1), \dots, C(u^i), \dots, C(u^m)]$  является вероятностью распределения контента страницы  $u$ , где компонента  $C(u^i)$  представляет собой относительный вклад  $i$ -го тематического раздела в контент страницы  $u$ .

Авторитет-вектор  $A_u$ :  $[A(u^1), \dots, A(u^i), \dots, A(u^m)]$  определяет степень значимости страницы  $u$  по тематическим разделам, где компонента  $A(u^i)$  определяет степень значимости страницы  $u$  относительно  $i$ -го тематического раздела (как известно, в алгоритме HITS, кроме авторитет-вектора определяется и концентратор-вектор).

Степень значимости  $A(u^i)$  страницы  $u$  на  $i$ -м тематическом разделе вычисляется таким образом:

$$A(u^i) = d(1-\alpha) \sum_{v \in B(u)} \frac{1}{|F(v)|} A(v^i) + d\alpha \sum_{v \in B(u)} \frac{1}{|F(v)|} C(v^i) \sum_{k \in T} A(v^k) + \frac{(1-d)}{n} C(u^i) \sum_{v \in G} \sum_{k \in T} A(v^k), \quad (5)$$

где  $i \in T = \{1, 2, \dots, m\}$ ,  $\alpha$  — вероятность перехода пользователя на  $i$ -й тематический раздел на странице.

Пусть  $A(v) = \sum_{k \in T} A(v^k)$  и  $\sum_{v \in G} A(v) = 1$ , тогда формула (5) примет более компактный вид:

$$A(u^i) = d \sum_{v \in B(u)} \frac{(1-\alpha)A(v^i) + \alpha C(v^i)A(v)}{|F(v)|} + \frac{(1-d)}{n} C(u^i).$$

Когда на странице тематические переходы скрыты, т. е. при  $\alpha = 0$ , данная модель сводится к оригинальному PageRank-алгоритму.

В Topical HITS-алгоритме зависимость между векторами  $A(u)$  и  $H(u)$  задается следующими соотношениями:

$$A(u^i) = \sum_{v \in B(u)} \frac{(1-\alpha)H(v^i) + \alpha C(v^i)H(v)}{|F(v)|},$$

$$H(v^j) = \sum_{v \in F(u)} \frac{(1-\alpha)A(u^i) + \alpha C(u^i)A(u)}{|B(u)|},$$

где  $H(v) = \sum_{k \in T} H(v^k)$  и  $A(u) = \sum_{k \in T} A(u^k)$ .

В случае, когда страница не разбита на тематические разделы, т. е. если на странице тематические детали скрыты, эта модель сводится к нормализованному алгоритму HITS или упрощенному варианту алгоритма SALSA [4]:

$$A(u) = \sum_{v \in B(u)} \frac{1}{|F(v)|} H(v),$$

$$H(v) = \sum_{v \in F(u)} \frac{1}{|B(v)|} A(u).$$

В работе [9] была предложена модель WPSS (Web Page Scoring Systems), обобщающая все вышеупомянутые модели. Прежде чем произвести



ранжирование, алгоритм WICER (Weighted Inter-Cluster Edge Rank) [10] предлагает кластеризацию страниц по тематике. Для уточнения весов гиперссылок в алгоритме WLRank (Weighted Links Rank) [11] предлагается учитывать различные атрибуты. А именно, полагается, что атрибуты — tag, anchor text и др. — дают гиперссылкам дополнительные веса, в результате чего улучшается точность поисковых машин.

Заметим, что во всех перечисленных работах главная цель — найти web-страницы с максимальным рангом, отвечающим потребностям пользователей. В этом контексте и работа [12] не исключение. Для нахождения web-страниц с максимальным рангом в ней, на основе теории потока в сетях, предложена оптимизационная модель ранжирования web-страниц, которая была сведена к задаче линейного программирования.

Изложенный краткий обзор позволяет сделать вывод, что исследование проблемы ранжирования web-страниц является перспективным направлением разработки технологий информационного поиска.

## 2. ТРИ МОДИФИКАЦИИ АЛГОРИТМА PAGERANK

**Первая модификация.** Как было отмечено, вероятность нахождения страницы с максимальным рангом непосредственно связана с выбором гиперссылок. В данной модификации для определения вероятности выбора гиперссылки используется точечная взаимная информация (Pointwise Mutual Information — PMI) [13]. При определении вероятности выбора гиперссылки точечной взаимной информацией учитывается, что если web-страницы указываются (цитируются) одной и той же страницей, то каждая гиперссылка (ссылочная страница) содержит долю информации о других гиперссылках (ссылочных страницах).

Пусть  $F(v)$  — множество страниц, на которые ссылается страница  $v$ . Тогда, следуя работе [13], точечная взаимная информация между гиперссылками  $v \rightarrow u_i$  и  $v \rightarrow u_j$  определяется формулой:

$$PMI(v \rightarrow u_i, v \rightarrow u_j) = \log_2 \left( \frac{p(u_i, u_j)}{p(u_i)p(u_j)} \right), \quad (6)$$

где  $p(u_i)$  — вероятность цитирования страницы  $u_i$  (вероятность следования пользователя по гиперссылке  $v \rightarrow u_i$ ), а  $p(u_i, u_j)$  — вероятность коцитирования страниц  $u_i$  и  $u_j$  (совместная вероятность следования пользователя по гиперссылкам  $v \rightarrow u_i$  и  $v \rightarrow u_j$ ).

В дальнейшем для простоты записи вместо обозначения  $PMI(v \rightarrow u_i, v \rightarrow u_j)$  будем применять обозначение  $PMI(u_i, u_j)$ .

Интуитивно, точечной взаимной информацией между гиперссылками  $v \rightarrow u_i$  и  $v \rightarrow u_j$  измеряется количество информации по отношению друг к другу.

Из определения (6) следует, что если вероятности выбора гиперссылок  $v \rightarrow u_i$  и  $v \rightarrow u_j$  независимы, то нет приращения информации. Это означает, что если гиперссылки независимы, то гиперссылка  $v \rightarrow u_i$  не содержит никакой информации о гиперссылке  $v \rightarrow u_j$ , и наоборот. Следовательно, их точечная взаимная информация равна нулю, т. е. информация о гиперссылке  $v \rightarrow u_i$  не дает никакой информации о гиперссылке  $v \rightarrow u_j$  (и наоборот). Действительно, если гиперссылки  $v \rightarrow u_i$  и  $v \rightarrow u_j$  независимы, то  $p(u_i, u_j) = p(u_i)p(u_j)$  и, следовательно,

$$PMI(u_i, u_j) = \log_2 \left( \frac{p(u_i, u_j)}{p(u_i)p(u_j)} \right) = \log_2 1 = 0.$$

Для определения вероятности выбора гиперссылки  $v \rightarrow u_i$  вычисляется суммарная точечная взаимная информация  $PMI(u_i)$ , которая определяется между ней и остальными исходящими гиперссылками страницы  $v$ , и общая точечная взаимная информация  $PMI(F(v))$ , которая вычисляется между всевозможными парами гиперссылок, исходящими из страницы  $v$ .

Суммарную точечную взаимную информацию гиперссылки  $v \rightarrow u_i$  будем определять формулой:

$$PMI(u_i) = \sum_{\substack{u_j \in F(v) \\ u_j \neq u_i}} PMI(u_i, u_j). \quad (7)$$

Общая точечная взаимная информация  $PMI(F(v))$  получается суммированием формулы (7):

$$\begin{aligned} PMI(F(v)) &= \sum_{u_i \in F(v)} PMI(u_i) = \\ &= \sum_{u_i \in F(v)} \sum_{\substack{u_j \in F(v) \\ u_j \neq u_i}} PMI(u_i, u_j). \end{aligned} \quad (8)$$

Тогда доля ранга  $PR(v)$ , которая распределяется по страницам  $u_i \in F(v)$ , т. е. вес гиперссылки  $v \rightarrow u_i$  будет вычисляться отношением формул (7) и (8):

$$\begin{aligned} \omega(v \rightarrow u_i) &= \frac{PMI(u_i)}{PMI(F(v))} = \\ &= \frac{\sum_{\substack{u_j \in F(v) \\ u_j \neq u_i}} PMI(u_i, u_j)}{\sum_{u_i \in F(v)} \sum_{\substack{u_j \in F(v) \\ u_j \neq u_i}} PMI(u_i, u_j)}, \quad u_i \in F(v). \end{aligned} \quad (9)$$

Легко видеть, что если вероятности выбора гиперссылок  $v \rightarrow u_i$  и  $v \rightarrow u_j$  независимы, то приходим к неопределенности. Во избежание неопределенности формула (9) преобразуется к виду:

$$\omega(v \rightarrow u_i) = \frac{1 + \sum_{\substack{u_j \in F(v) \\ u_j \neq u_i}} PMI(u_i, u_j)}{|F(v)| \sum_{\substack{u_i \in F(v) \\ u_j \in F(v) \\ u_j \neq u_i}} PMI(u_i, u_j)}, \quad (10)$$

$u_i \in F(v).$

Таким образом, оригинальный PageRank-алгоритм имеет вид:

$$PR(u) = \frac{(1-d)}{n} + d \sum_{v \in B(u)} \omega(v \rightarrow u) PR(v). \quad (11)$$

Из формулы (11), как следствие, легко можно получить оригинальный PageRank-алгоритм. Действительно, если предположить, что выбор гиперссылок независим, тогда из формулы (10) вытекает, что вероятность выбора каждой гиперссылки равна величине  $\omega(v \rightarrow u_i) = 1/|F(v)|$ , которая совпадает с вероятностью выбора гиперссылки в оригинальном PageRank-алгоритме, выраженным формулой (1).

Теперь переходим к вычислению вероятности  $p(u_i, u_j)$ , которую можно определить так:

$$p(u_i, u_j) = p(u_j|u_i)p(u_i). \quad (12)$$

Поскольку каждая страница представляется как «мешок слов», то условная вероятность

$$p(u_j|u_i) = \sum_{k=1}^K p(u_j|u_i, w_k)p(w_k|u_i),$$

где  $K$  означает общее число слов в наборе страниц  $\{v\} \cup F(v)$ .

При допущении независимости появления слова  $w_k$  в страницах  $u_i$  и  $u_j$  последнюю формулу можно выразить так:

$$p(u_j|u_i) = \sum_{k=1}^K p(u_j|u_i, w_k)p(w_k|u_i) \approx \sum_{k=1}^K p(u_j|w_k)p(w_k|u_i),$$

и формула (12) принимает вид:

$$p(u_i, u_j) = p(u_i) \sum_{k=1}^K p(u_j|w_k)p(w_k|u_i). \quad (13)$$

Пусть  $f_{ik}$  — число появления слова  $w_k$  в странице  $u_i$ . Тогда вероятность появления слова  $w_k$  в странице  $u_i$

$$p(w_k|u_i) = f_{ik} / \sum_{s=1}^K f_{is}. \quad (14)$$

Условная вероятность  $p(u_j|w_k)$  определяется согласно формуле Байеса:

$$p(u_j|w_k) = \frac{p(w_k|u_j)p(u_j)}{p(w_k)}, \quad (15)$$

где  $p(w_k) = \sum_{i=1}^{|F(v)|} p(w_k|u_i)p(u_i)$ .

С учетом формул (13)–(15)

$$p(u_i, u_j) = \frac{p(u_i)p(u_j)}{K} \sum_{k=1}^K \frac{f_{ik}f_{jk}}{p(w_k)}.$$

Подставляя последнее выражение в формулу (6), для точечной взаимной информации получим следующую формулу:

$$PMI(u_i, u_j) = \log_2 \left( \frac{1}{\sum_{s=1}^K f_{is} \sum_{s=1}^K f_{js}} \sum_{k=1}^K \frac{f_{ik}f_{jk}}{p(w_k)} \right).$$

**Вторая модификация.** В работе [14] показано, что при вычислении PageRank-вектора следует учитывать не только входящие гиперссылки, но и исходящие. Этот алгоритм, как и все перечисленные модификации алгоритма PageRank, позволяет избежать равномерного распределения ранга страницы между ссылочными страницами. Здесь для каждой гиперссылки вычисляются два веса:  $\omega^+(v \rightarrow u)[0, 1]$  и  $\omega^-(v \rightarrow u)$ .

Вес  $\omega^+(v \rightarrow u)$  определяется числом входящих гиперссылок страницы  $u$  и ссылочных страниц страницы  $v$ :

$$\omega^+(v \rightarrow u) = \frac{|B(u)|}{\sum_{x \in F(v)} |B(x)|}, \quad (16)$$

а вес  $\omega^-(v \rightarrow u)$  — числом исходящих гиперссылок страницы  $u$  и всех страниц множества  $F(v)$ :

$$\omega^-(v \rightarrow u) = \frac{|F(u)|}{\sum_{x \in F(v)} |F(x)|}. \quad (17)$$

С учетом формул (16) и (17) оригинальный PageRank-алгоритм модифицируется так [14]:

$$PR(u) = \frac{(1-d)}{n} + d \sum_{v \in B(u)} \omega^+(v \rightarrow u) \omega^-(v \rightarrow u) PR(v).$$

Как можно заметить из формул (16) и (17), в этой модификации не учитывается тематическая близость страниц. Для учета тематической близос-



ти страниц нами предлагаются следующие определения весов (16) и (17):

$$\omega^+(v \rightarrow u) = \frac{\sum_{x \in B(u)} \text{sim}(u, x)}{\sum_{y \in F(v)} \sum_{z \in B(y)} \text{sim}(y, z)}, \quad (18)$$

$$\omega^-(v \rightarrow u) = \frac{\sum_{x \in F(u)} \text{sim}(u, x)}{\sum_{y \in F(v)} \sum_{z \in F(y)} \text{sim}(y, z)}. \quad (19)$$

Если предположить, что для любой гиперссылки ( $x \rightarrow y$ ) мера подобия между страницами  $x$  и  $y$  принимает одно и то же значение, то из нашего определения, как следствие, получается результат работы [14]. Действительно, пусть для любой гиперссылки ( $x \rightarrow y$ )  $\text{sim}(x, y) = a = \text{const}$ , тогда из формул (18) и (19) получаются соответствующие формулы (16) и (17).

**Третья модификация.** Предлагаемая модификация является усовершенствованным вариантом алгоритма Topic-Centric [5]. В отличие от алгоритма Topic-Centric в нашем варианте вероятность выбора гиперссылки не только зависит от степени близости страниц  $v$  и  $u \in F(v)$ , она также зависит от степени близости страниц множества  $F(v)$ :

$$PR(u) = \frac{(1-d)}{n} + d \sum_{v \in B(u)} \left( \frac{\lambda \sum_{x \in F(v)} \text{sim}(u, x) + (1-\lambda) \sum_{x \neq u} \text{sim}(u, x)}{\lambda \sum_{x \in F(v)} \text{sim}(x, v) + (1-\lambda) \sum_{x \in F(v)} \sum_{y \in F(v), y \neq x} \text{sim}(x, y)} \right) \times PR(v), \quad (20)$$

где  $0 \leq \lambda \leq 1$ .

Если влияние мер близости страниц множества  $F(v)$  свести к нулю, т. е. если в формуле (20) положить  $\lambda = 1$ , то получается формула (4).

Можно предложить и другую модификацию:

$$PR(u) = \frac{(1-d)}{n} + d \sum_{v \in B(u)} \left( \frac{\text{sim}(u, v) \sum_{x \in F(v)} \text{sim}(u, x)}{\sum_{x \in F(v)} \text{sim}(x, v) \cdot \sum_{x \in F(v)} \sum_{y \in F(v), y \neq x} \text{sim}(x, y)} \right) PR(v).$$

### ЗАКЛЮЧЕНИЕ

Исследования последних лет показали, что точность алгоритмов анализа гиперссылок — HITS, PageRank и др. — непосредственно зависит от выбора гиперссылки. Другими словами, степень точности ранжирования результатов поиска непосредственно зависит от вероятности выбора гиперссылки. В оригинальном PageRank-алгоритме и в некоторых других его модификациях вероятности выбора гиперссылок считались равными, т. е. гиперссылки выбирались с одинаковой вероят-

ностью. Проведенные эксперименты показали, что такой подход не гарантирует нахождение релевантных страниц, отвечающих потребностям пользователей. Одна из главных причин заключается в том, что в традиционных алгоритмах ранг страницы определяется числом гиперссылок. Дальнейшие исследования подтверждают, что без контентного анализа документов невозможно решить проблему эффективности поисковых машин. Эффективность поиска оценивается степенью релевантности отобранных документов к запросу пользователя. Для повышения эффективности поиска в работе предложены три модификации алгоритма PageRank. Каждая модификация при определении веса гиперссылки учитывает тематическую близость страниц и их соседей, связанных гиперссылками. Предложенные в данной статье модификации алгоритма PageRank представляют собой усовершенствованные варианты результатов работ [1, 2, 5, 14].

### ЛИТЕРАТУРА

1. Brin S., Page L. The anatomy of a large-scale hyper-textual Web search engine // Computer Networks and ISDN systems. — 1998. — Vol. 30, — N 1–7. — P. 107–117.
2. Berkhin P. A survey on PageRank computing // Internet Mathematics. — 2005 — 2006. — Vol. 2, N 1. — P. 73–120.
3. Kleinberg J.M. Authoritative sources in a hyperlinked environment // Journal of the ACM. — 1999. — Vol. 46, N 5. — P. 604–632.
4. Lempel R., Moran S. SALSA: the stochastic approach for link-structure analysis // ACM Trans. on Information Systems. — 2001. — Vol. 19, N 2. — P. 131–160.
5. Ingongngam P., Rungsawang A. Topic-centric algorithm: a novel approach to Web link analysis // Proc. of the 18th Intern. Conf. on Advanced Information Networking and Applications (AINA'04). — Fukuoka, Japan, 2004. — Vol. 2. — P. 299–301.
6. Haveliwala T.H. Topic-sensitive PageRank: a context-sensitive ranking algorithm for Web search // IEEE Trans. on Knowledge and Data Eng. — 2003. — Vol. 15, N 4. — P. 784–796.
7. Richardson M., Domingos P. The intelligent surfer: probabilistic combination of link and content information in PageRank // Advances in Neural Information Processing Systems. MIT Press. — 2002. — Vol. 14. — P. 1441–1448.
8. Nie L., Davison B. D., Qi X. Topical link analysis for Web search // Proc. of the 29<sup>th</sup> Annual International ACM SIGIR Conference on Research and Development in Information Retrieval. — Seattle, USA, 2006. — P. 91–98.
9. Diligenti M., Gori M., Maggini M. A unified probabilistic framework for Web page scoring systems // IEEE Trans. on Knowledge and Data Engineering. — 2004. — Vol. 16, N 1. — P. 4–16.
10. Padmanabhan D., Desikan P., Srivastava J. WICER: a weighted inter-cluster edge ranking for clustered graphs // Proc. of the 2005 IEEE/WIC/ACM Intern. Conf. on Web Intelligence (WI'2005). — Compiègne, France, 2005. — P. 522–528.
11. Baeza-Yates R., Davis E. Web page ranking using link attributes // Proc. of 13<sup>th</sup> World Wide Web Conference (WWW13). — New York, USA, 2004. — P. 328–329.
12. Алыгулиев Р.М. Оптимизационная модель ранжирования Web-страниц // Системы управления и информационные технологии. — 2006. — № 3(25). — С. 4–7.
13. Efron M. Using cocitation information to estimate political orientation in Web documents // Knowledge and Information Systems. — 2006. — Vol. 9, N 4. — P. 492–511.
14. Xing W., Ghorbani A. Weighted PageRank algorithm // Proc. of the Second Annual Conf. Communication Networks and Services Research (CNSR'04). — Fredericton, Canada, 2004. — P. 305–314.

e-mail: rasim@science.az; a.ramiz@science.az

Статья представлена к публикации членом редколлегии В.Л. Эпштейном. □

# МНОГОПРОЦЕССОРНАЯ АРХИТЕКТУРА ПС-2000 НА КРИСТАЛЛЕ СБИС

Ю.С. Затуливетер, Е.А. Фищенко

*Институт проблем управления им. В.А. Трапезникова РАН, г. Москва*

Рассмотрены современные проблемы построения многопроцессорных архитектур на одном кристалле. Предложено развитие архитектуры хорошо зарекомендовавшей себя отечественной многопроцессорной вычислительной системы ПС-2000 для реализации на одном кристалле.

## ВВЕДЕНИЕ

Длительное время микроэлектронная промышленность обеспечивала быстрый рост производительности микропроцессоров. До середины 1990-х гг. производительность увеличивалась пропорционально росту как числа транзисторов на кристалле, так и рабочей частоте. Но в последнее десятилетие, несмотря на экспоненциальный рост числа транзисторов, производительность микропроцессоров увеличивалась, главным образом, лишь благодаря увеличению тактовой частоты, сопровождавшемуся быстрым увеличением тепловыделения.

В линейке микропроцессоров i286, i386 и i486 повышение производительности посредством увеличения числа транзисторов от нескольких десятков до нескольких сотен тысяч производилось, в основном, за счет увеличения разрядности машинных слов и аппаратного распараллеливания алгоритмов выполнения арифметических операций. Другой резерв повышения производительности с увеличением числа транзисторов от единиц до десятков миллионов — динамически раскрываемый параллелизм на основе анализа связей по данным на уровне соседних команд — был задействован во всех микропроцессорах линии Pentium.

Однако резервы параллелизма, скрывавшегося в рамках модели последовательных вычислений фон Неймана, изначально ограничены. Быстро растущий избыток транзисторов на кристалле, не укладывавшихся в ограниченные резервы параллелизма, отдавались буферной памяти (кэши разных

уровней), вклад которых в рост производительности был относительно небольшим.

Производительность микропроцессоров линии Pentium в расчёте на отдельный транзистор с ростом числа транзисторов только падала. Максимальное значение этого показателя достигнуто на самой первой модели Pentium (3,1 млн. транзисторов), что привело к структурному насыщению микропроцессорных архитектур и свидетельствовало о начавшемся (в скрытой форме) уже к середине 1990-х гг. кризисе классической модели последовательного счёта [1].

Естественный выход из этого кризиса — однокристальные многопроцессорные архитектуры [1, 2]. Однако на этом пути имеются два принципиальных препятствия:

- известные решения в части многопроцессорных архитектур не предназначались для однокристальной реализации, поскольку принципы балансировки однокристалльных многопроцессорных архитектур существенно иные;
- трудоёмкость программирования высокопараллельных архитектур на порядки выше, чем последовательных, поскольку требуется максимальный учет их архитектурных особенностей.

Ведущие изготовители микроэлектронных компонентов, пытаясь сохранить тенденцию роста производительности, предлагают радикальные изменения в базовой архитектуре персональных компьютеров — многоядерные однокристалльные процессоры. Но столь существенные архитектурные изменения предполагают кардинальное изменение подходов к индустриальному программированию, а именно — активное использование



чрезвычайно трудоемкого программирования разнообразных архитектур с массовым параллелизмом [3]. Многоядерность ясно и масштабно, но, к сожалению, с большим опозданием (почти на десятилетие [1]), обозначила на компьютерном рынке угрозу системного кризиса индустрии программирования. Сейчас, на пороге массового производства многоядерных кристаллов проявление «архитектурного голода» [1] на тщательно сбалансированные высокопараллельные архитектуры под стремительно выходящее количество от сотен миллионов до миллиарда и более транзисторов проявилось уже в острых формах.

Проблемы построения многопроцессорных систем общего и специального назначения, распараллеливания вычислений, программирования и эффективного применения таких систем относятся к наиболее сложным в компьютерной тематике. Отыскание лучших по одновременно многим системным и пользовательским характеристикам или хотя бы приемлемых вариантов составляет чрезвычайно сложную задачу.

Переход на микропроцессорную элементную базу, произошедший в 1980-х гг., сыграл злую шутку с архитекторами, занимавшимися разработкой многопроцессорных систем. Ввиду высокой сложности многопроцессорных систем темпы их разработки существенно отставали от темпов смены поколений и роста производительности новых микропроцессоров.

Сначала из сферы полномочий разработчиков многопроцессорных систем полностью ушла арифметика с регистровым уровнем хранения промежуточных данных, на которых обеспечивается максимально возможная скорость выполнения массовых операций. Им оставались только сферы межпроцессорной коммутации с принудительными запретами осуществления обмена данными между процессорами на регистровом уровне. Возможности распараллеливания для многопроцессорных систем сузились, поскольку они вынужденно сместились, главным образом, на уровень слабо взаимодействующих процессов с утратой колоссального резерва массового параллелизма на уровне арифметических операций.

Затем по мере повышения степени интеграции элементной базы блоки коммутации, собираемые из «россыпи», в целях снижения стоимости и сроков разработок стали вытесняться аппаратными средствами поддержки сетевых протоколов. Так многопроцессорные архитектуры эволюционировали в кластерные системы. В очень успешной борьбе за снижение стоимости и сроков крупнопанельная «сборка» таких систем свелась к приобретению и монтажу готовых многолезвийных ком-

пьютерных блоков и сетевых модулей. Необходимость в глубоких архитектурных разработках на уровне мелких «кирпичиков» отпала сама собой.

В результате более чем за двадцатилетний период быстрого и непрерывного укрупнения элементной базы в ходе смены поколений разработчиков была во многом утрачена культура и навыки комплексной балансировки всех возможных аппаратных уровней параллелизма. Именно такая культура становится совершенно необходимой при переходе к многопроцессорным архитектурам на одном кристалле, поскольку в полное ведение архитекторов поступает совокупность всех предоставляемых на кристалле транзисторов, каждый из которых — мельчайший и самый универсальный из возможных «кирпичиков».

Период длительного забвения архитекторов, работающих одновременно на всех уровнях аппаратного параллелизма, завершен. На транзисторном уровне для них открываются беспрецедентные возможности для повышения на порядки эффективности параллельных архитектур, как общего, так и специального назначения.

В связи с завершением архитектурного «ледникового периода» сохранившиеся островки культуры и навыков построения параллельных архитектур не из «крупнопанельной» элементной базы, а из мельчайших элементов, вне всяких сомнений составляют золотой фонд мирового компьютеростроения.

Одним из них является отечественная компьютерная школа, сформировавшаяся в 1972—1990 гг. в ходе создания и широкого промышленного применения многопроцессорной вычислительной системы (МВС) ПС-2000.

## 1. ПРЕИМУЩЕСТВА АРХИТЕКТУРЫ МВС ПС-2000

Оригинальная высокопараллельная архитектура ПС-2000 [4—7] обладает уникальными возможностями гибкой перестраиваемости под структуры решаемых задач. Она не только не утратила своей актуальности, но обладает реализованным лишь в малой части потенциалом, как в расширении классов решаемых задач, так и в существенном повышении эффективности параллельной обработки, благодаря балансировке всех доступных видов параллелизма, от уровня массовых операций до уровня системных процессов.

Система ПС-2000 разрабатывалась в 1975—1980 гг. Институтом проблем управления (г. Москва) совместно с НПО «Импульс» (г. Северодонецк), где с 1981 по 1988 г. она серийно выпускалась. Это была первая в мире выпущенная большим тиражом (242 шт.) высокопроизводительная

(200 Mips, где ips — команда (инструкция) в секунду) многопроцессорная вычислительная система, которая характеризовалась как «самый быстродействующий и самый продвинутый компьютер Советского Союза» [8, 9].

Система ПС-2000 широко и с большим экономическим эффектом применялась в различных областях народного хозяйства. Одно из главных промышленных применений — обработка данных геофизической сейсморазведки нефти и газа. Системы ПС-2000 входили в состав экспедиционных геофизических вычислительных комплексов ЭГВК ПС-2000, оснащенных системой обработки геофизической информации (сейсмоданных) СОС ПС (ВНИИ «Геофизика», г. Москва). Эти комплексы успешно конкурировали с геофизическими вычислительными центрами, оснащенными на порядок более дорогими суперкомпьютерами зарубежного производства [6].

Положенные в основу МВС ПС-2000 архитектурные принципы особенно перспективны сегодня. Тогда они позволили при очень малой по нынешним временам тактовой частоте 3 МГц снимать со всех параллельно работающих 64-х процессорных элементов (ПЭ) ПС-2000 пиковую производительность 200 Mips (или 50 Mflops, где flops — операция с плавающей запятой в секунду), получаемую умножением тактовой частоты на число ПЭ. При этом совокупность из 64 медленных (3 МГц) ПЭ работала как один «сверхбыстрый» процессор с частотой 200 МГц. Близкая к 100 % средняя загрузка всех ПЭ в реальных применениях обеспечивала среднюю производительность, близкую к пиковой, что достигалось на решении важнейших задач промышленной обработки данных [10].

В настоящее время реализация архитектурных принципов ПС-2000 на одном кристалле с тактовой частотой 2 ГГц может обеспечить производительность более 120 Gips, т. е. работать с приведённой к одному процессору частотой более 120 ГГц.

Наличие большого и далее быстро растущего резерва транзисторов, предоставляемого современными СБИС-технологиями, даёт возможность существенно увеличить производительность выполнения арифметических операций с числами в форматах с плавающей запятой, что позволяет оценить производительность отдельного кристалла с 64 ПЭ в 120 Gflops.

Масштабируемая по числу ПЭ архитектура ПС-2000 позволяет наращивать число ПЭ пропорционально росту числа транзисторов на кристалле. В указанных предположениях производительность, близкая к 1 Tflops, может достигаться на 512 ПЭ в одном кристалле.

Элементная база из многоядерных микросхем с архитектурой ПС-2000 откроет возможности для построения программно реконфигурируемых параллельно-конвейерных систем сверхвысокой производительности как специального, так и общего назначения. С помощью таких кристаллов производительность в 1 Pflops достижима в нескольких стойках.

Пиковая производительность ПС-2000, равная произведению быстродействия арифметико-логического устройства на число ПЭ, достигалась путем одновременного исполнения вычислительных операций и управляющих действий в других функциональных устройствах, таких как считывание команд, считывание операндов и запись результатов, модификация адресов операндов, активация ПЭ, межпроцессорные обмены, управление конфигурацией и т. п.

Эффективное управление обеспечивается реализацией принципов [11]:

- применения отдельных блоков памяти для хранения команд и данных, как векторных, так и скалярных с возможностью одновременного доступа;
- одновременного исполнения независимых действий чтение/запись в памяти для векторных и скалярных данных;
- применения значительных объемов регистровой памяти в разнотипных функциональных устройствах;
- применения форматов команд одинакового размера, разбитых на поля, в которых задаются независимые трехадресные регистровые операции;
- совмещения программным путем нескольких параллельно исполняемых действий разной длительности.

Эти принципы получили распространение в более поздних зарубежных разработках компьютеров различных классов [12, 13]. Развитие предложенных принципов в применении к высокопараллельным компьютерным архитектурам актуально и в настоящее время.

Как известно, МВС с SIMD-архитектурой весьма эффективны для обработки мультимедийных данных [14, 15] и цифровой обработки сигналов [16], для решения научных задач, которые хорошо описываются параллельными алгоритмами, для обработки акустической, радиолокационной и телеметрической информации. В этой связи весьма актуальной становится проблема создания универсальной программируемой МВС с SIMD-архитектурой на одном кристалле, которая могла бы во многих областях заменить специализированные ускорители для различных применений.





## 2. МИКРОАРХИТЕКТУРА ПС-2000М

Микроархитектура ПС-2000М, т. е. архитектура МВС ПС-2000, предназначенная для реализации на одном кристалле, весьма эффективна для решения задач, в которых на одно вводимое данное приходится много промежуточных действий.

Требования к балансировке микроархитектуры:

- параллелизм на уровне массовых операций требует гибкой системы межпроцессорного обмена промежуточными данными на регистровом уровне;
- балансировка общего транзисторного ресурса на кристалле при его разделении между ресурсами обработки, хранения и коммутации;
- балансировка времени и совмещение ввода-вывода данных со временем счета;
- балансировка параллелизма на уровне массовых операций с уровнем системных процессов;
- масштабируемость микроархитектуры внутри кристалла;
- масштабируемость архитектуры на межкристальном уровне (наращивание производительности путем прямого соединения между собой многопроцессорных кристаллов на уровне плат, блоков плат, стоек).

На рис. 1 (см. цветную вклейку) представлена микроархитектура МВС ПС-2000М. Она содержит общее устройство управления (ОУУ),  $m$  устройств обработки (УО) и устройство коммутации (УК).

Каждое устройство обработки  $УО.j$ ,  $j \in \{1, 2, \dots, m\}$ , содержит  $n$  ПЭ, банк динамической памяти  $M.j$  с доступом к строкам размером  $n$  32-разрядных слов, канал прямого доступа  $DMA.j$  к банку памяти  $M.j$ , входной буфер  $MW.j$  строки банка памяти  $M.j$ , содержащий  $n$  32-разрядных регистров  $MW.j.1, MW.j.2, \dots, MW.j.n$ , выходной буфер  $MR.j$  строки банка памяти  $M.j$ , содержащий  $n$  32-разрядных регистров  $MR.j.1, MR.j.2, \dots, MR.j.n$ , устройство локальной адресации  $L.j$  к строке банка памяти  $M.j$ . Каждый ПЭ  $ПЭ.j.i$ ,  $j \in \{1, 2, \dots, m\}$ ,  $i \in \{1, 2, \dots, n\}$ , содержит арифметико-логическое устройство  $S.j.i$  с регистрами общего назначения, устройство активации  $T.j.i$ .

Общее устройство управления содержит банк динамической памяти  $H$  с произвольным доступом, входной буфер  $HW$  банка памяти  $H$ , выходной буфер  $HR$  банка памяти  $H$ , устройство адресации  $HL$  к банку памяти; арифметико-логическое устройство  $W$  с регистрами общего назначения, память команд  $G$ .

Устройство коммутации содержит широковещательный и кольцевой каналы. Широковещательный канал содержит  $m$  регистров  $K.j$  и регистр  $KU$ . Регистр  $K.j$  соединен с устройством обработки  $УО.j$ , регистр  $KU$  соединен с ОУУ. Широковеща-

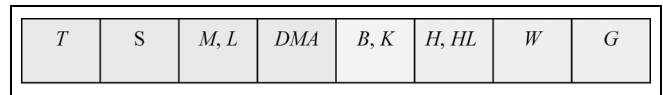


Рис. 2. Структура команды МВС ПС-2000М

тельный канал предназначен для передачи всем ПЭ общего данного либо из ОУУ, либо из любого активированного ПЭ. Кольцевой канал состоит из  $m$  колец  $B.j$ , каждое из которых состоит из  $n$  регистров  $B.j.1, B.j.2, \dots, B.j.n$ .

Вход каждого устройства  $S.j.i$  соединен с регистрами  $MR.j.i, T.j.i, B.j.i, L.j$  и  $K.j$ , его выход соединен с регистрами  $B.j.i, L.j$  и  $K.j$ . Выбор процессорного элемента  $ПЭ.j.i$ , который соединяется с общими для всех ПЭ, входящими в УО, регистрами  $L.j$  и  $K.j$ , осуществляется с помощью устройства активации  $T.j.i$ .

На рис. 2 приведена структура команды, которая состоит из полей, задающих одновременную работу векторных функциональных устройств  $T, S, M$  и  $L, DMA$ , кольцевого  $B$  и магистрального  $K$  каналов, скалярных функциональных устройств  $H, HL, W$  и  $G$ .

В настоящее время пропускная способность памяти существенно ограничивает эффективность многоядерных вычислительных систем. В микроархитектуре ПС-2000М пропускная способность памяти  $V$  значительно увеличивается, поскольку она пропорциональна всем основным параметрам, характеризующим параллелизм и быстродействие архитектуры:  $V = n * m * p * f$ , где  $m$  — число локальных банков динамической памяти  $M.j$  в  $m$  устройствах обработки,  $n$  — число слов в строке банка памяти  $M.j$ ,  $p$  — число разрядов слова,  $f$  — внутренняя тактовая частота памяти в кристалле.

Пропускная способность памяти значительно возрастает, благодаря как значительному увеличению числа одновременно считываемых бит информации, так и увеличению тактовой частоты памяти. Так, при  $m = 64$ ,  $n = 8$ ,  $p = 32$  и  $f = 1000$  МГц пропускная способность памяти составит 160 Тбит/с. Время доступа к динамической памяти становится постоянным благодаря адресации только к строкам банков памяти, что позволяет ей работать в синхронном режиме.

Нетрудно видеть, что и функциональные возможности программно управляемой памяти архитектуры ПС-2000М, размещаемой на одном кристалле, много превосходят возможности буферной кэш-памяти, размещаемой на кристалле микропроцессоров, в том числе и многоядерных.

Масштабирование микроархитектуры МВС ПС-2000М обеспечивается благодаря возможности программной настройки конфигурации кольце-

вого и широкоэмитательного каналов, что позволяет отключать неисправные устройства обработки и тем самым увеличивать процент выхода годных микросхем.

Применение динамической памяти позволяет снизить уровень энергопотребления на кристалле МВС ПС-2000М. Применение высокопараллельной масштабируемой SIMD-архитектуры для организации вычислений и распределенной памяти позволяет регулировать соотношение производительность/энергопотребление путем изменения внутренней тактовой частоты, а также путем балансировки количества оборудования для памяти и вычислений. Это делает микроархитектуру ПС-2000М легко масштабируемой относительно энергопотребления и дает возможность ее применения в мобильных устройствах.

### 3. ОБ АВТОМАТИЗАЦИИ ПРОГРАММИРОВАНИЯ

Предложенная для реализации на одном кристалле микроархитектура ПС-2000М является синхронной с заранее определенным количеством тактов при выполнении команд в различных функциональных устройствах. В ней аппаратно разделены и совмещены во времени команды памяти, массовых вычислений, межпроцессорных обменов, модификации адресов операндов, вычисления условий, подготовки настроек, ввода/вывода. Она предоставляет открытый для программистов параллелизм на уровне команд и данных, который доступен для формализации посредством структурных параметров. Система команд обладает гибкой системой настроек, позволяющей составлять программы, инвариантные к числу ПЭ.

Указанные свойства открывают возможности для автоматического синтеза программ машинного уровня, эффективно использующих все возможности высокопараллельной архитектуры, доступные на низовом программном уровне, прежде всего, параллелизм массовых операций. В рамках исследовательских работ по автоматизации программирования высокопараллельной архитектуры ПС-2000 сформировался общий подход [17], в котором параллельные компьютерные архитектуры в полнофункциональном виде задаются формальным структурным параметром.

Параметризованный синтез программ можно представить формулой  $p = S(f, r)$ , где  $p$  — синтезируемая параллельная программа машинного уровня;  $f$  — математическое описание программируемой функции, сохраняющее ее естественный параллелизм и не зависящее от компьютерной архитектуры;  $r$  — структурный параметр, представляющий параллельную архитектуру с уровнем детализации до всех машинных действий открытых

на программном уровне;  $S$  — универсальная относительно допустимых значений  $f$  и  $r$  процедура, осуществляющая автоматический синтез высокоэффективных программ машинного уровня, предназначенных для выполнения в компьютере с архитектурой, представленной структурным параметром  $r$ .

Это качественно новый путь построения средств индустриального программирования параллельных компьютеров, который открывает возможности не только автоматического синтеза программ машинного уровня, но и решения острых проблем межархитектурной переносимости многопроцессорных приложений.

### ЗАКЛЮЧЕНИЕ

Появление на массовом рынке многоядерных кристаллов легализовали долго скрывавшиеся за массовыми тиражами фоннеймановских компьютеров и технологий индустриального программирования до сих пор нерешенные фундаментальные проблемы «архитектурного голода» на высокопараллельные архитектуры и технологии их программирования. Это признак серьезных перемен на мировом компьютерном рынке.

Прежде доступные только в США опережающие микропроцессорные СБИС-технологии обеспечивали им подавляющие односторонние преимущества в конкурентной борьбе. Пример тому — яркое, но краткое по времени (конец 1980-х — начало 1990-х гг.) противостояние замечательной во многих отношениях параллельной транспьютерной архитектуры, предложенной и совместно доведенной до коммерческого воплощения ведущими европейскими (Англия, Франция, Германия) фирмами. Но запаса опережающих преимуществ этой архитектуры хватило лишь на несколько лет. Европа не смогла противостоять заокеанской лавине транзисторов. Выдающаяся во многих отношениях транспьютерная архитектура уступила грубой силе закона Мура.

Теперь, спустя десятилетие, продолжающий своё действие закон Мура привел к несбалансированным многоядерным архитектурам, демонстрирующим не столько прогресс компьютерных архитектур, сколько вынужденный (без должной замены) откат компьютерной индустрии от своего главного «стандарта» — модели фон Неймана, исчерпавшей свои системообразующие возможности [1]. На высшие приоритеты возвращаются высокопараллельные архитектуры, но уже в более изощренных требованиях многоуровневой балансировки и легко масштабируемого СБИС-погружения многопроцессорных сегментов с однокристалльной реализацией.



## ЛИТЕРАТУРА

1. *Затуливетер Ю.С.* Компьютерные архитектуры: неожиданные повороты // *Hard 'n' Soft*. — 1996. — № 2. — С. 86—94 <[http://zvt.hotbox.ru/p2\\_z1.htm](http://zvt.hotbox.ru/p2_z1.htm)>.
2. *Параллельные высокопроизводительные ЭВМ на основе заказных СБИС с многопроцессорной архитектурой / Ю.С. Затуливетер, Е.А. Фищенко, В.А. Кротов, В.А. Лементуев // Приборы и системы управления*. — 1996. — № 12. — С. 24—26.
3. *Зюбин В.* Многоядерные процессоры и программирование // *Открытые системы*. — 2005. — № 7-8. — С. 12—19.
4. *Прангишвили И.В., Виленкин С.Я., Медведев И.Л.* Многопроцессорные вычислительные системы с общим управлением. — М.: Энергоатомиздат, 1983. — 312 с.
5. *А. с. 751238 (СССР)*. Многопроцессорная вычислительная система / А.Я. Бирюков, С.Я. Виленкин, В.А. Жуков и др. // *Бюл.* — 1983. — № 22.
6. *Москва компьютерная / И.В. Прангишвили, Ю.С. Затуливетер, А.Д. Смирнов, А.Н. Томилин // Развитие инженерного дела в Москве: Исторические очерки*. — М., 1998. — С. 338—354. <<http://www.computer-museum.ru/histussr/moscomp.htm>>.
7. *Затуливетер Ю.С., Фищенко Е.А.* Многопроцессорная вычислительная система ПС-2000 <<http://www.computer-museum.ru/histussr/11-1.htm>>.
8. *Wolcott P., Goodman S. E.* High-Speed computers of the Soviet Union // *IEEE Computer*. — 1988. — Vol. 21, N 9. — P. 32—41.
9. *Wolcott P., Goodman S.E.* International perspectives: under the stress of reform high-performance computing in the former Soviet Union // *Communications of the ACM*. — 1993. — Vol. 36, N 10. — P. 21—24.
10. *Миграция временного разреза земли на параллельной ЭВМ ПС-2000 (Быстрое сейсмологографическое преобразова-*
- ние Кирхгофа) / Е.Т. Амелина, Ю.С. Затуливетер, Е.П. Лазебник и др. — М.: Ин-т пробл. управл. 1992. — 36 с.
11. *Фищенко Е. А.* Выбор системы команд для многопроцессорной вычислительной системы с общим потоком команд // *Многопроцессорные вычислительные системы с общим потоком команд: Сб. науч. тр. / Ин-т пробл. упр.* — М., 1978. — № 19. — С. 33—39.
12. *Fisher J.* Very long instruction word architectures and the ELI512 // *Proc. of the 10-th Annual International Symposium on Computer Architecture*. — Stockholm, — 1983. — P. 140—150.
13. *Philips Semiconductors.* Introduction to VLIW Computer Architecture <[http://www.semiconductors.philips.com/acrobat\\_download/other/vliw-wp.pdf](http://www.semiconductors.philips.com/acrobat_download/other/vliw-wp.pdf)>.
14. *Kozyrakis C., Patterson D.* Vector vs superscalar and VLIW architectures for embedded multimedia benchmarks // *Proc. of the 35th Intl. Symp. on Microarchitecture / Istanbul, Turkey, Nov., 2002*. — P. 283—293 <<http://citeseer.ist.psu.edu/kozyrakis02vector.html>>.
15. *Nguyen T., Zakhor A., Yelick K.* Performance Analysis Of An H.263 Video Encoder For Viram <<http://citeseer.ist.psu.edu/correct/416764>>.
16. *Elliott D. G., Snelgrove W. M., Stumm M.* Computational RAM: A Memory-SIMD Hybrid and its Application to DSP // *In Custom Integrated Circuits Conference*. — Boston, MA, 1992. — P. 30.6.1—30.6.4.7 <<http://www.dissonance.com/archive/rtcapers/Elliott92.pdf>>.
17. *Затуливетер Ю.С.* Введение в проблему параметризованного синтеза программ для параллельных компьютеров. — М.: Ин-т. пробл. упр., 1993. — 88 с.

☎ (495) 334-92-09, e-mail: [zvt@ipu.ru](mailto:zvt@ipu.ru), [fish@ipu.ru](mailto:fish@ipu.ru)

Статья представлена к публикации членом редколлегии В.Д. Малюгиным. □



## УВАЖАЕМЫЕ ДАМЫ И ГОСПОДА!

Приглашаем Вас принять участие в седьмой международной конференции и выставке «СИСТЕМЫ ПРОЕКТИРОВАНИЯ, ТЕХНОЛОГИЧЕСКОЙ ПОДГОТОВКИ ПРОИЗВОДСТВА И УПРАВЛЕНИЯ ЭТАПАМИ ЖИЗНЕННОГО ЦИКЛА ПРОМЫШЛЕННОГО ПРОДУКТА (CAD/CAM/PDM-2007)»,

которая состоится с 23 по 25 октября 2007 г. в Москве, в Институте проблем управления им. В.А. Трапезникова РАН.

## Тематика конференции:

- организация структур технических и программных средств проектирования и управления; средства взаимодействия, структуры данных, международные стандарты;
- компьютерная графика и CAD/CAM/PDM-системы в учебных процессах (программы обучения по дисциплинам, методические материалы, тестирование); средства виртуальной реальности в промышленных системах;
- интегрированные производственные системы и управление технологическими процессами PDM-системы;
- проектирование в машиностроении и строительстве;
- проектирование в радиоэлектронике;
- аналитика и творчество.

**Ключевые даты:** до 1 сентября 2007 г. — заявка на участие в конференции и (или) выставке, перевод оргвзноса, представление тезисов докладов (более одной страницы формата А5);

до 20 сентября 2007 г. — представление полных текстов докладов.

Заявки, тезисы и доклады высылаются по e-mail: [conf18@ipu.ru](mailto:conf18@ipu.ru). Форма заявки, правила оформления тезисов и докладов представлены на сайте <http://lab18.ipu.rssi.ru>.

**Адрес оргкомитета:** 117997 г. Москва, Профсоюзная ул., 65, Институт проблем управления им. В.А. Трапезникова РАН.

Председатель Оргкомитета д-р техн. наук Е.И. Артамонов, учёный секретарь канд. техн. наук С.В. Смирнов.

Тел. для справок: (495) 334-93-50; факс (495) 334-91-29.

ОРГКОМИТЕТ

# ПРИЛОЖЕНИЕ КЛАСТЕРНОГО АНАЛИЗА К АВТОМАТИЧЕСКОМУ РАСПОЗНАВАНИЮ ДЕФЕКТОВ<sup>1</sup>

И.Л. Васильев<sup>(1)</sup>, Д.Н. Сидоров<sup>(2)</sup>

<sup>(1)</sup>Институт динамики систем и теории управления СО РАН, г. Иркутск;

<sup>(2)</sup>Институт систем энергетики СО РАН, г. Иркутск

Рассмотрена задача автоматического распознавания и классификации дефектов, возникающая при разработке систем машинного зрения. Предложен подход к ее решению, в соответствии с которым на первом этапе осуществляется автономный процесс обучения системы, во время которого обучающие выборки анализируемых признаков дефектов разбиваются на кластеры, в которых находятся медианы. На втором этапе в режиме реального времени реализуется процесс инспектирования, при котором входящий образец классифицируется в соответствии с типом ближайшей медианы. Отмечено, что сравнительный анализ с другими методами показал эффективность предложенного подхода.

## ВВЕДЕНИЕ

Важную роль в производственных системах машинного зрения, применяемых для контроля качества выпускаемой продукции, играют автоматические распознавание и классификация объектов. Современные аппаратные средства позволяют с высокой скоростью получать и хранить изображения с высоким разрешением. В связи с этим актуальна задача разработки быстрых и эффективных методов обработки изображений, основанных на визуальной схожести объектов. На их основе возможно создание автоматических систем распознавания дефектов на производственных линиях, работающих в режиме реального времени.

Как правило, процесс классификации изображений состоит из двух этапов. На первом этапе из изображения объекта выделяются его основные признаки и характеристики, т. е. осуществляется переход от реального образа в пространство признаков (дескрипторов). Существуют различные подходы, позволяющие описывать такие признаки

изображений, как текстура, форма, контур и цвет [1]. В отличие от реальных изображений, данные признаки представляют собой числовые векторы и поэтому достаточно просто их сравнивать. На втором этапе происходит непосредственно классификация объектов в соответствии с их признаками. Человеческий мозг отлично справляется с задачами классификации изображений даже в условиях больших помех и отсутствия существенных признаков. Однако искусственные классификаторы все еще далеки от совершенства.

В данной статье рассматривается практическая производственная задача распознавания и классификации образов, которая возникает в системе автоматического распознавания дефектов (АРД), работающей в режиме реального времени при изготовлении полупроводниковых микросхем.

Система АРД должна находить дефект на изображении объекта и определять: является ли данный дефект браком в производстве (например, царапина, трещина или наплыв), или же это всего лишь так называемый «оверкилл» — посторонний предмет, не влияющий на качество продукции (например, волос, волокно или пыль). Таким образом, система АРД должна классифицировать входящий образец как «брак» или «оверкилл». Од-

<sup>1</sup> Исследование выполнено при финансовой поддержке РФФИ (проекты 05-01-00110 и 05-01-00336) и НАТО (проекты RIG981258 и RIG98127).



но из главных требований систем реального времени — достаточное быстроедействие. Поэтому только недорогие в вычислительном плане алгоритмы могут быть использованы при извлечении признаков изображений и их классификации. Примеры систем АД можно найти в энергетике (инспектирование трубопроводных систем [2]), целлюлозно-бумажном производстве [3], деревоперерабатывающей промышленности [4] и др.

Проведенные эксперименты показали, что методы, хорошо зарекомендовавшие себя в задачах классификации, такие как метод нейронных сетей [5] и метод SVM (Support Vector Machine) [6], не подходят для рассматриваемого случая. При анализе ряда пространств признаков изображений было замечено, что в общем случае образцы брака и оверкиллов не могут быть разделены на два подпространства. Тем не менее, они концентрируются в нескольких слабо пересекающихся областях. В связи с этим возникла идея применения кластерного анализа при классификации. Задача кластерного анализа заключается в разбиении исходного множества объектов на заданное число подмножеств схожих объектов (кластеров), и затем выбирается наиболее «типичный» представитель (медиана) из каждого кластера. Предлагается следующий подход, состоящий из двух основных этапов: 1) автономный этап обучения, 2) этап классификации, выполняемый в режиме реального времени. На первом этапе обучающие выборки брака и оверкиллов разбиваются на кластеры. На втором этапе входящий образец классифицируется в соответствии с типом ближайшей медианы.

Задача кластерного анализа может быть сформулирована в виде хорошо известной задачи комбинаторной оптимизации — задачи о  $p$ -медиане [7]. Рассмотрим граф, в котором каждой вершине сопоставлено изображение объекта. Расстояние между вершинами графа (вес дуг) определяется схожестью соответствующих изображений. Под схожестью изображений понимается расстояние между векторами признаков изображений, вычисленных по определенной метрике в пространстве признаков. Таким образом, чтобы разбить исходное множество на определенное число кластеров (скажем,  $p$ ), необходимо найти  $p$  медианных вершин графа, минимизирующих сумму весов дуг до остальных вершин.

Известно, что задача о  $p$ -медиане является  $NP$ -трудной [7]. Обширный обзор литературы по этой задаче можно найти в работе [8], а также в работах [9, 10]. Отдельно отметим последние работы, касающиеся решения задачи о  $p$ -медиане большой размерности: VNS (Variable Neighborhood Search) [11], GRASP эвристики [10], метод ветвей, отсечений и оценок [9].

Для решения задачи кластеризации в настоящей работе используется подход, предложенный в работе [9] с ограничением по времени счета. В результате численного эксперимента на данных производства полупроводниковых микросхем были получены решения, уменьшающие ошибки классификации SVM и нейронных сетей на 10—20 %.

## 1. СИСТЕМА АВТОМАТИЧЕСКОГО РАСПОЗНАВАНИЯ ДЕФЕКТОВ

Структура системы АД представлена на рис. 1. Объект снимается цифровой камерой высокого разрешения, и изображения передаются программному блоку «сегментатор». В нем в изображении выделяются сегменты, отличающиеся от «идеального» изображения, т. е. от изображения объекта без дефектов. Далее, из каждого сегмента извлекаются их числовые признаки, которые передаются в блок «классификатор». Описание процесса сегментации и извлечения признаков дано в § 2.

Основным компонентом классификатора является так называемая функция классификации. Останемся на ее описании более подробно.

Пусть дано  $m$  признаков изображений  $x^i$  размерности  $n_i$ , т. е.  $x^i \in R^{n_i}$ ,  $i = 1, \dots, m$ . Функция классификации  $f(\cdot): R^{\sum_{i=1}^m n_i} \mapsto \{-1, 1\}$  определяется следующим образом:

$$f(x) = f(x^1, \dots, x^m) = \begin{cases} -1, & \text{если } x \text{ соответствует оверкиллу,} \\ 1, & \text{если } x \text{ соответствует браку.} \end{cases}$$

На практике для построения функции классификации широко используются метод SVM [6] и нейронные сети [5]. Процессу классификации предшествует процесс обучения, в ходе которого на обучающей выборке подбираются параметры функции.

При классификации методами SVM и нейронными сетями пространство признаков в некото-

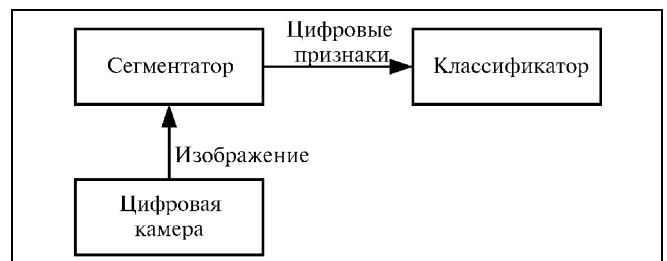


Рис. 1. Структура системы АД

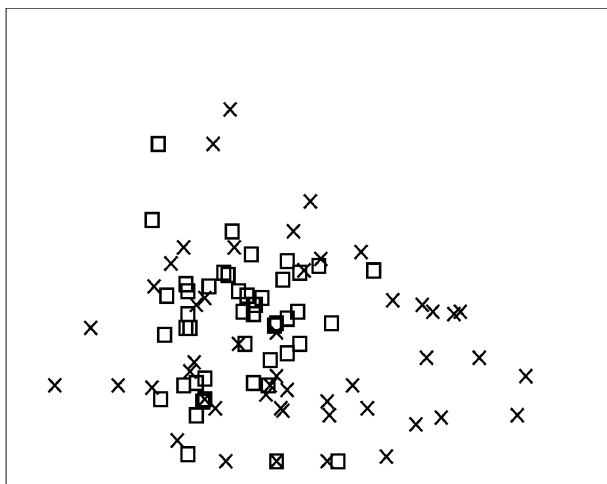


Рис. 2. Характерный пример пространства признаков:  
 × — оверкиллы; □ — брак

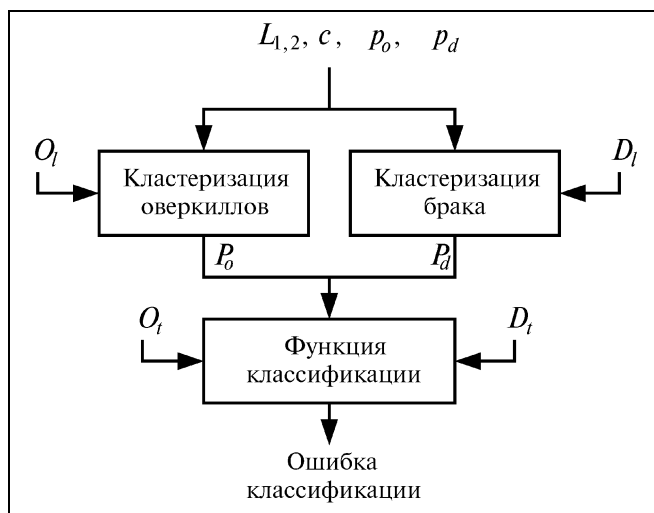


Рис. 3. Схема обучения

ром смысле делится на два подпространства, что в рассматриваемом случае сделать невозможно. Проиллюстрируем это на примере двух проекций данных на плоскость одного из признаков (рис. 2). Видно, что точки брака и оверкиллов не разделяются. Тем не менее, можно заметить, что точки разного типа группируются в несколько кластеров.

Поэтому предлагается построение функции классификации на основе кластерного анализа. Степень схожести между изображениями определяется как взвешенная сумма расстояний (в определенной метрике) между соответствующими признаками, т. е.

$$d(x, y) = \sum_{i=1}^m c_i \|x^i - y^i\|,$$

где  $c_i$  — заданные веса признаков. Рассматривались метрики  $L_1$  (Манхеттен) и  $L_2$  (Евклида). Схема процесса обучения представлена на рис. 3. Обучение производится на обучающей выборке оверкиллов  $O_t$  и брака  $D_t$ . В данных выборках соответственно выбираются  $p_o$  и  $p_d$  медиан (подробности алгоритма кластеризации см. в § 3).

Обозначим их множества через  $P_o$  и  $P_d$ , соответственно. Входящий образец классифицируется в соответствии с типом ближайшей к нему медианы:

$$f(x) = \begin{cases} -1, & \text{если } \min_{y \in P_o} d(x, y) \leq \min_{y \in P_d} d(x, y), \\ 1, & \text{в противном случае.} \end{cases}$$

Таким образом, процесс обучения заключается в выборе метрики при подсчете расстояния между признаками изображений, подборе весов  $c = (c_1, \dots, c_m)^T$  и выборе числа медиан  $p_o$  и  $p_d$ . Качество подобранных параметров определяется по числу ошибок классификации на тестирующих выборках оверкиллов  $O_t$  и брака  $D_t$ . Процесс обучения может производиться автономно и не имеет жестких временных ограничений. Поэтому в процессе обучения выбирались лучшие по критерию качества параметры путем полного перебора из дискретных ограниченных множеств параметров. Детали практических экспериментов описываются в § 4.

## 2. ПРОСТРАНСТВА ПРИЗНАКОВ

Пространства признаков (дескрипторов) формируются на этапе, следующем за выделением (сегментацией) дефекта из исходного изображения продукта. При выборе признаков возникают два основных подхода к описанию выявленной области изображения: по характеру границы и по структуре расположения элементов изображения внутри области. Отметим, что при построении дескрипторов дефектов необходимо учитывать их инвариантность к перемещению по полю изображения (сдвиг, поворот), а также в некоторых случаях масштаб.

### 2.1. Дескрипторы границ

Существует ряд методов, позволяющих компактно представлять контуры областей, подробный обзор которых представлен в работах [1, 12]. Простейшие геометрические признаки контуров, такие как периметр, кривизна, площадь и др., широко используются в практических приложениях. В данной работе использовались следующие признаки для описания контуров.



- *Нормализованный цепной код.* Компактный дескриптор контуров, основанный на цепном коде Фримена [1, 13] был предложен в середине прошлого века и получил широкое распространение в силу вычислительной простоты и инвариантности относительно поворота и сдвига. Суть метода заключается в линейном приближении контура последовательностью соединенных отрезков, для которых указана длина и направление. В качестве дескриптора использовалась нормализованная функция распределения направлений этих отрезков. В настоящей работе также использовался более компактный дифференциальный цепной код [13], в котором кодируются не длины, а только изменения направлений.
- *Приближение кривой контура рядами Фурье* [1].
- *Инвариантные моменты.* Данный признак основан на теореме, гарантирующей возможность описания кусочно-непрерывной функции последовательностью ее моментов [14]. В качестве признака, обладающего свойством инвариантности относительно сдвига, поворота и изменения масштаба, был выбран набор из семи моментов, представляющих собой специальную комбинацию моментов второго и третьего порядков [15].

## 2.2. Текстурные дескрипторы

Для описания внутренних элементов изображений использовались следующие дескрипторы.

- *Нормализованная гистограмма изображения,* которая подсчитывается по всем элементам изображения, включая границу [13].
- *Дескрипторы Харлика.* Отметим, что гистограмма не несет никакой информации о взаимном расположении элементов изображения. Один из способов учета подобной информации заключается в построении матриц совместного распределения яркостей и вычислении по ним текстурных характеристик, таких как контраст, энергия, энтропия, однородность, максимум вероятности и обратный момент разности. Такие дескрипторы известны как статистические дескрипторы Харлика [1, 2]. Практическое использование дескрипторов Харлика состоит в обучении на выборке из различных текстур, состоящем в выборе способа построения матрицы совместного распределения яркостей. Для анализа текстуры изображений используются также признаки Габора [1, 16], основанные на неортогональном вейвлет-преобразовании, но их вычисление требует значительных временных затрат и, следовательно, не подходит для работы в режиме реального времени.
- *Признак совокупной направленности текстуры.* Такой признак, основанный на внешнем восприятии текстуры изображения, был впервые предложен в работе [17]. На первом этапе распознаются

перепады (границы) в анализируемой области и для каждого элемента изображения подсчитывается угол и значение перепада. Далее составляется функция распределения вероятностей перепадов (гистограмма), характеризующая все значения перепадов, превышающие определенное пороговое значение, и квантованная по значениям углов вычисленных перепадов. В итоге, острота всплесков такой функции, подсчитанная по их вторым моментам, была использована в качестве признака совокупной направленности.

Отметим, что задача построения эффективных дескрипторов методически близка к задаче сжатия информации, так как ставит целью компактное представление тех или иных признаков. Различные ортогональные преобразования (Фурье, косинус, Хаара, вейвлет) использовались для поиска признаков в задачах обработки сигналов и изображений (см. например, работы [13, 18] и др). Однако в случае построения эффективных систем АРД необходимо также учитывать специфику тех или иных дефектов. Поэтому ключевым моментом является автоматизация выбора пространства признаков по обучающей и тестирующей выборкам.

## 3. ЗАДАЧА КЛАСТЕРИЗАЦИИ

Один из подходов к решению задачи кластерного анализа состоит в использовании моделей дискретной оптимизации [19—24]. На практике данный подход мало распространен, так как его применение относительно сложно с вычислительной точки зрения. Действительно, задача о  $p$ -медиане, как уже отмечалось, относится к классу  $NP$ -трудных задач, т. е. сложность задачи растет экспоненциально относительно ее размерности. Тем не менее, современный уровень развития теории методов дискретной оптимизации и вычислительной техники позволяет успешно решать эту задачу достаточно большой размерности [9—11].

Рассмотрим задачу кластерного анализа в терминах комбинаторной оптимизации. Пусть множества объектов  $V = \{1, \dots, n\}$  — суть признаки изображений и расстояния между ними, как определено в § 1. Рассмотрим полный, взвешенный ориентированный граф  $G(V, A)$ , где  $V$  — множество вершин графа,  $A = \{uv : u \in V, v \in V, u \neq v\}$  — множество дуг графа и  $w_{uv} = d(u, v)$  — веса дуг (расстояние между объектами). Введем бинарные переменные  $y_u$  ( $x_{uv}$ ), которые будут соответствовать вершинам  $u \in V$  (дугам  $uv \in A$ ). Переменная  $y_u$  равна 1, если вершина  $u$  является медианой. Переменная  $x_{uv}$  равна 1, если вершина  $v$  относится к кластеру, определяемому медианой  $u$ . В этих обозначениях задача о  $p$ -медиане может быть сформу-

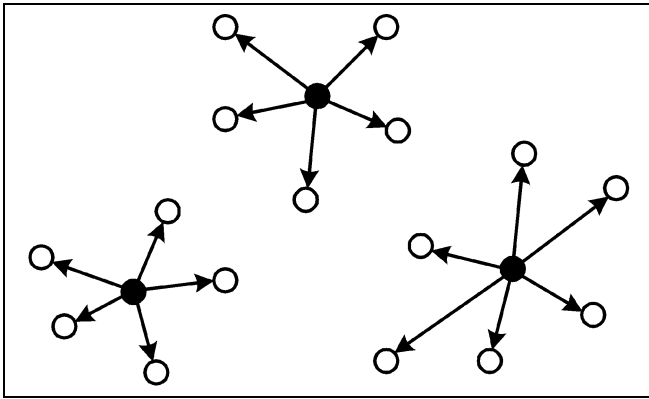


Рис. 4. Допустимая точка

лирована как задача целочисленного программирования:

$$\sum_{uv \in A} w_{uv} x_{uv} \rightarrow \min_{(x, y)}, \quad (1)$$

$$\sum_{u \in V, u \neq v} x_{uv} + y_v = 1, \quad \forall v \in V, \quad (2)$$

$$x_{uv} \leq y_u, \quad \forall uv \in A, \quad u \in V \quad (3)$$

$$\sum_{u \in V} y_u = p \quad (4)$$

$$y_u \in \{0, 1\}, \quad \forall v \in V, \quad (5)$$

$$x_{uv} \in \{0, 1\}, \quad \forall uv \in A \quad (6)$$

Ограничения (2) гарантируют, что каждая вершина  $v$  является либо медианой, либо имеет одну входящую дугу из медианной вершины. Неравенства (3) исключают существование выходящих дуг из немедианных вершин. Число медиан определяется уравнением (4).

Допустимая точка задачи на графе представляет из себя  $p$  «звезд», вершины которых образуют кластеры с медианами посередине, как проиллюстрировано на рис. 4. Для решения задачи применялся подход, основанный на алгоритме ветвей, отсечений и оценок, детально описанный в работе [9]. Данный подход хорошо зарекомендовал себя при решении задач о  $p$ -медиане большой размерности. В этой статье отметим только основные компоненты этого подхода (общая теория и методы дискретной оптимизации может быть найдена в книге [25]).

• *Лагранжевая эвристика.* Данный метод позволяет находить как нижние, так и верхние оценки целевой функции. Кроме того, информация из этого алгоритма использовалась при оценивании переменных на следующем этапе.

• *Методы генерации столбцов и строк* для решения задачи линейного программирования (ЛП), полученной из (1)–(6) путем ослабления условий

целочисленности. Данный метод позволяет рассматривать большинство переменных  $x_{uv}$  и ограничений (3) в неявном виде, тем самым значительно уменьшая объем используемой памяти и ускоряя работу алгоритмов решения ЛП задач.

• *Метод отсечений.* В ходе генерации столбцов и строк также ведется поиск специальных ограничений, не входящих в формулировку задачи (1)–(6), которые позволяют улучшить нижние оценки метода ветвей и границ.

• *Метод ветвей и границ* — хорошо известный метод неявного перебора, который широко применяется для решения задач дискретной оптимизации.

Результаты, представленные в работе [9], показывают, что данный подход позволяет на ранних этапах вычисления находить хорошие решения. Поэтому в настоящей работе было установлено ограничение в 30 с при решении одной задачи кластеризации. Численный эксперимент показал, что в рассматриваемых задачах при таком временном ограничении относительная погрешность решения не превосходит 1 %, что вполне допустимо.

#### 4. ЧИСЛЕННЫЙ ЭКСПЕРИМЕНТ

Предложенный подход к классификации для системы АРД был протестирован на примере производства полупроводниковых микросхем. Была рассмотрена база данных, состоящая из 1194 изображений дефектов. Примеры брака и оверкилла представлены на рис. 5. База данных была случайным образом разбита на обучающую и тестирующую выборки, как представлено в табл. 1. Из каждого изображения было извлечено 6 признаков (нормализованные цепочки кодов, коэффициенты Фурье, нормализованная гистограмма, признаки по матрице смежности, моменты и признак направленности).

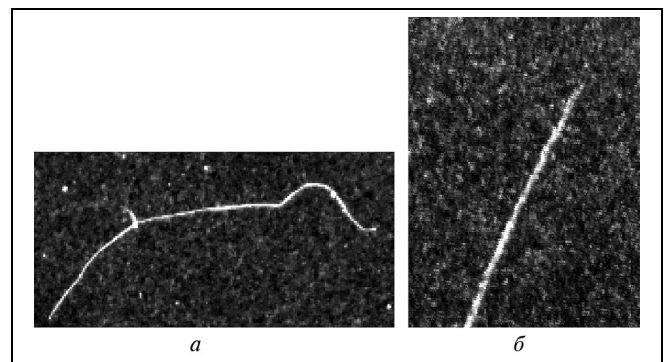


Рис. 5. Примеры оверкиллов и брака на печатной плате:  $a$  — посторонние частицы (нить и пыль);  $b$  — критический брак (царапина)





Таблица 1

**База данных изображений**

Число изображений в выборке	Брака	Оверкиллов
обучающей	257	322
тестирующей	269	346
Всего	526	668

Таблица 2

**Ошибка классификации**

$p_0$	$p_d$	Ошибка, %	$p_0$	$p_d$	Ошибка, %
16	11	<b>26,5</b>	11	10	28,3
	9	26,8	12	9	
13	11	27,0		10	
16	10	27,2	13	7	
13	9	27,3	14	10	
20	11	27,5	17	9	
11		27,6	20	10	
13	10		15		28,5
12	11	27,8	17	8	
14			16		12
15				13	
17			9		28,0
20	12	28,8			
14			13	28,9	
17	12	28,1	15		12
13	12		17	7	
15	9		19	13	
17	14	28,3	13	6	29,1
11	9		28,3	13	6

Таблица 3

**Ошибки классификации методом SVM, %**

Кернфункция SVM	SVM-Lite	LIBSVM
Линейная	38,2	38,1
Полином степени 2	38,1	43,7
" " 3	37,6	
Радиальный базис	<b>37,4</b>	
Сигмоид	44,3	

На первом этапе экспериментов при обучении рассматривались все возможные варианты параметров из множеств, определенных следующим образом:

- метрики  $L_1$  и  $L_2$ ;
- веса  $c_i = 0, 1, 2$ , исключая из рассмотрения эквивалентные векторы весов (т. е. пропорциональные векторы);
- число медиан  $p_o$  и  $p_d$  от 5 до 20 с шагом 5.

Результаты вычислений показали, что дефекты наилучшим образом характеризуются контуром и наиболее адекватным признаком является нормализованный цепной код. Данное заключение также подтверждается мнением экспертов в данной области. Использование различных метрик не выявило существенной разницы в результатах. Только в некоторых случаях евклидова метрика показала незначительно лучшие результаты. С учетом этих замечаний, на втором этапе рассматривались только цепной код, евклидова норма, значения  $p_o$  и  $p_d$  изменялись от 5 до 20 с шагом 1. Лучшие результаты вычислений представлены в табл. 2.

Предложенный подход сравнивался с результатами классификации, полученными методом SVM (табл. 3). Были использованы программные пакеты SVM-Lite [26] и LIBSVM [27]. Основным компонентом в методе SVM является кернфункция, с помощью которой определяется вид разделяющей области. Самые распространенные кернфункции:

- линейная – с линейным ядром  $K(x, y) = x * y$ .
- полином  $K(x, y) = (x * y + c)^d$ .
- радиальный базис  $K(x, y) = \exp(-\gamma \|x - y\|^2)$ .
- сигмоид – с нейронным ядром  $K(x, y) = \text{th}(ax * y + b)$ ,

где «звездочка» означает скалярное произведение.

Другие пакеты, реализующие метод SVM [6] и нейронные сети [5], не улучшают полученные ошибки классификации. Таким образом, применяя кластерный анализ для решения задачи классификации промышленных дефектов, нам удалось снизить ошибку классификации с 37,4 до 26,5 % по сравнению с методом SVM.

**ЗАКЛЮЧЕНИЕ**

В данной работе предложена новая методика классификации образов при идентификации промышленного брака. Численные результаты на практических данных показали ее эффективность, обусловленную учетом неоднородности распределения пространства признаков благодаря построению функции классификации на основе кластерного анализа и автоматическим выбором опти-

мальных параметров функции классификации из широкой области их допустимых значений.

Однако выбор параметров функции классификации осуществляется простым перебором из большой области значений, что требует значительных временных затрат. Для уменьшения времени счета целесообразно исследовать возможность применения методов искусственного интеллекта для поиска оптимальных параметров и распараллеливания алгоритма. Отметим, что на практике важна не только классификация на два множества (брак и оверкилл), но и более детальная классификация каждого из этих множеств для анализа производственного процесса. Например, знание преобладающих видов брака и оверкиллов может помочь в поиске и устранении причин, способствующих их возникновению.

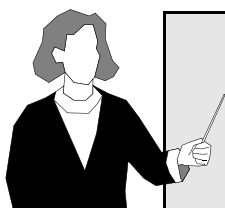
Авторы благодарят компанию «Microview Technologies Ltd.» (Сингапур) за любезно предоставленные экспериментальные материалы, профессоров Pasquale Avella и Jean-Francois Lerallut за полезные дискуссии и критические замечания.

## ЛИТЕРАТУРА

1. Гонзалес Р., Вудс Р. Цифровая обработка сигналов. — М: Техносфера, 2005. — 1072 с.
2. Mery D., Berti M.A. Automatic detection of welding defects using texture features // CTIP2003 Proceedings. — Berlin, 2003. — P. 676—681.
3. Kunttu I., Lepisto L., Visa A. Efficient Fourier shape descriptor for industrial defect images using wavelets // SPIE Optical Engineering Letters. — 2005. — Vol. 44. — P. 345—351.
4. Niskanen M., Silven O., Kauppinen H. Wood inspection with non-supervised clustering // Machine Vision and Applications Journal. — 2003. — Vol. 13. — P. 275—285.
5. Хайкин С. Нейронные сети: полный курс. — М: Изд. дом «Вильямс», 2006.
6. Scholkopf B., Burges C.J.C., Smola A.J. Advances in Kernel Methods. Support Vector Learning. — The MIT Press, 1999.
7. Kariv O., Hakimi L. An algorithmic approach to network location problems. ii: the  $p$ -medians // Operations Research. — 1979. — Vol. 37. — P. 539—560.
8. Mladenovic N., Brimberg J., Hansen P., Perez J.  $p$ -median problem: A survey of metaheuristic approaches // European Journal of Operations Research. — 2007. — Vol. 179. — P. 927—939.
9. Avella P., Sassano A., Vasil'ev I. Computational study of large-scale  $p$ -median problems // Mathematical Programming. — 2006. In press, available online.
10. Resende M.G.C., Werneck R.F.F. A hybrid heuristic for the  $p$ -median problem // Journal of Heuristics. — 2004. — Vol. 10. — P. 59—88.
11. Hansen P., Mladenovic N., Perez-Brito D. Variable neighbourhood decomposition search // Journal of Heuristics. — 2001. — Vol. 7. — P. 335—350.
12. Marshall S. Review of shape coding techniques // Image and Vision Computing. — 1989. — Vol. 7. — P. 281—294.
13. Павлидис Т. Алгоритмы машинной графики и обработки изображений. — М.: Радио и связь, 1986.
14. Papoulis A. Characteristic Functions. — N.-Y.: McGraw-Hill, 1991. — P. 153—162.
15. Hu M. Visual pattern recognition by moment invariants // IRE Trans. Information Theory. — 1962. — Vol. 8. — P. 179—187.
16. Ohanian P., Dubes R. Performance evaluation for four classes of textural features // Pattern Recognition. — 1992. — Vol. 25. — P. 819—833.
17. Tamura H., Mori S., Yamawaki T. Texture features corresponding to visual perception // IEEE Transactions on Systems, Man and Cybernetics. — 1978. — Vol. 8. — P. 460—472.
18. Sidorov D.N., Lerallut J.-F., Cocquerez J.-P., Azpiroz J. Robust retrieval from compressed medical image archives // Proc. of SPIE: Medical Imaging 2005. — San Jose, 2005. — Vol. 5748. — P. 419—426.
19. Bradley P.S., Mangasarian O.L. Feature selection via concave minimization and support vector machines // Proc. of the Fifteenth Intern. Conf. on Machine Learning (ICML'98). San Francisco, California, — 1998. — P. 82—90.
20. Fung G., Mangasarian O.L. Semi-supervised support vector machines for unlabeled data classification // Optimization Methods and Software. — 2000. — Vol. 1. — P. 29—44.
21. Hansen P., Jaumard B. Cluster analysis and mathematical programming // Mathematical Programming. — 1997. — Vol. 79. — P. 191—215.
22. Mulvey J.M., Crowder H.P. Cluster analysis: an application of Lagrangian relaxation // Management Science. — 1979. — P. 329—340.
23. Vinod H.D. Integer programming and the theory of groups // Journal of the American Statistical Association. — 1969. — Vol. 6. — P. 506—519.
24. Rao M.R. Cluster analysis and mathematical programming // Ibid. — 1971. — Vol. 6. — P. 622—626.
25. Nemhauser G.L., Wolsey L.A. Integer and Combinatorial Optimization. — N.-Y.: Wiley, — 1988.
26. Joachims T. Making large-scale SVM learning practical. — The MIT Press, 1999.
27. Chang C.-C., Lin C.-J. LIBSVM: a library for support vector machines. Department of Computer Science and Information Engineering. — National Taiwan University Publ., — 2001.

☎ +7-395-251-13-98, +7-395-242-84-40;  
e-mail: vil@icc.ru, dsidorov@isem.sei.irk.ru

Статья представлена к публикации членом редколлегии  
С.Н. Васильевым. □



Читайте в следующем номере статью В.В. Баранова и В.В. Матросова «Модели полезности и риска в задачах управления деградирующими системами», в которой строятся модель полезности, определяющая априорные предпочтения на альтернативах управления, и модель риска при выборе альтернатив диагностики ситуаций, ориентированные на применения в задачах управления деградирующими системами.

## **ОБ ОЦЕНИВАНИИ ВКЛАДА ВЕДОМСТВЕННЫХ ЦЕЛЕВЫХ ПРОГРАММ В ДОСТИЖЕНИЕ ЦЕЛЕЙ ПРАВИТЕЛЬСТВА**

П.Ю. Чеботарев<sup>(1)</sup>, А.Н. Митькин<sup>(2)</sup>, Д.С. Шмерлинг<sup>(3)</sup>

<sup>(1)</sup>Институт проблем управления им. В.А. Трапезникова РАН, г. Москва;

<sup>(2)</sup>Министерство экономического развития и торговли РФ, г. Москва;

<sup>(3)</sup>Государственный университет – Высшая школа экономики, г. Москва

Для формирования перечня ведомственных целевых программ, финансируемых из бюджета, предложена процедура расчетно-экспертного оценивания их вкладов в достижение целей Правительства Российской Федерации.

### **ВВЕДЕНИЕ**

Проводимая в стране реформа государственно-го управления предусматривает внедрение современных методов управления, ориентированных на достижение общественно значимых результатов и стимулирующих органы власти разных уровней и бюджетные организации максимально эффективно использовать бюджетные средства. Одна из главных целей реформы состоит в смещении акцента с «управления затратами» на «управление результатами» [1–5].

В рамках имеющихся бюджетных ограничений ежегодно возникает задача формирования перечня ведомственных целевых программ, финансируемых из федерального бюджета. В соответствии с нормативной правовой базой критерием выбора программ для финансирования (частично или полностью) из бюджета принимаемых обязательств служит их вклад в реализацию приоритетных направлений политики Правительства Российской Федерации [5, 6]. Однако регламентированная законодательно процедура формирования и реализации ведомственных целевых программ не осуществляется уже два бюджетных цикла.

В ряде экономически развитых стран разработаны и реализованы экспертные подходы к выбору

проектов для финансирования. Так, в Великобритании, Канаде, Австралии, Новой Зеландии в основе взаимодействия между правительством и ведомствами лежит контрактная модель, в рамках которой основой деятельности ведомства является контракт о предоставлении услуг. При этом ведомства несут ответственность за недостижение поставленных целей и поощряются за превышение запланированных значений показателей результативности (см., например, книги [7, 8]). В ряде других стран, в частности, в США и Франции, реализована административная модель, при которой финансирование происходит на основе бюджетных заданий и целевых показателей. При разработке системы бюджетирования в Российской Федерации успешный зарубежный опыт [8–11] должен использоваться с учетом российской специфики и в контексте проводимой реформы государственного управления [1–6].

В данной работе предлагается процедура расчетно-экспертного оценивания вкладов ведомственных целевых программ в достижение целей Правительства Российской Федерации. Оценивание производится для поддержки принятия решений высшими органами государственной власти при формировании перечня ведомственных целевых программ, предлагаемых к финансированию

(полностью или частично) из бюджета принимаемых обязательств [5, 6].

## 1. ИСХОДНЫЕ ДАННЫЕ И ПОСТАНОВКА ЗАДАЧИ

К исходным относятся следующие данные.

- Набор целевых программ, претендующих на получение средств из федерального бюджета.
- Трехуровневая система целей Правительства Российской Федерации ([12, Приложение 1.0]).
- Показатели, характеризующие уровни достижения целей третьего уровня. Показатели могут быть сопоставлены также целям второго и первого уровней. Для каждого показателя, характеризующего уровень достижения цели Правительства, в Сводном докладе о результатах и основных направлениях деятельности Правительства Российской Федерации [12, Приложение 1.1] указано целевое значение, соответствующее концу планового периода. Рассматриваются плановые периоды в 3 года и 1 год.
- Для каждого показателя, характеризующего уровень достижения цели Правительства Российской Федерации и одновременно характеризующего степень достижения цели программы, в описании программы указано:

— значение, которое будет достигнуто к концу планового периода программы в случае ее реализации;

— прогнозное инерционное значение, которое, согласно прогнозу, будет достигнуто к концу планового периода программы в случае, если ее инновационная часть не будет реализована.

Задача состоит в следующем: необходимо разработать процедуру оценивания вкладов ведомственных целевых программ в достижение целей Правительства, на реализацию которых эти программы направлены, а также — в общую систему целей Правительства [5, 12].

Полученные оценки должны отражать сравнительную ценность программ и сравнительную важность целей Правительства. Оценивание производится для формирования перечня ведомственных целевых программ, получающих финансирование из бюджета принимаемых обязательств [6, 13, 14].

## 2. ПРЕДЛАГАЕМАЯ ПРОЦЕДУРА ОЦЕНИВАНИЯ ВКЛАДОВ ВЕДОМСТВЕННЫХ ЦЕЛЕВЫХ ПРОГРАММ В ДОСТИЖЕНИЕ ЦЕЛЕЙ ПРАВИТЕЛЬСТВА

### 2.1. Задание плановых, инерционных и программных значений показателей достижения целей Правительства

Реализация бюджетных целевых программ направлена, в частности, на достижение целевых значений по показателям, характеризующим уровень достижения целей Правительства Российской

Федерации. Вклад целевой программы может быть оценен по отклонению динамики показателей в случае реализации программы от *инерционной динамики*.

Среди бюджетных целевых программ есть полностью инновационные — такие программы представляют собой проекты, которые не будут реализованы, если программы не будут включены в перечень программ, предлагаемых к финансированию из бюджета принимаемых обязательств. Программы другого типа объединяют как элементы инноваций, так и мероприятия, проводимые на постоянной основе. Для таких программ, как правило, предусматриваются два варианта финансирования: первый из них предполагает финансирование только мероприятий, проводящихся на постоянной основе, второй предполагает, кроме того, финансирование инновационных мероприятий. Финансирование в первом варианте обычно осуществляется из бюджета действующих обязательств, во втором варианте инновационные мероприятия финансируются из бюджета принимаемых обязательств.

Понятие «инерционной динамики» для двух указанных типов бюджетных целевых программ различно. Для полностью инновационных программ инерционная динамика соответствует случаю, когда ни данная программа, ни какие-либо ее аналоги не реализуются. Для программ, основу которых составляют мероприятия, проводящиеся на постоянной основе, инерционная динамика соответствует случаю, когда реализуются и финансируются только такие мероприятия. Таким образом, оценка вклада программы этого типа по отклонению от инерционной динамики учитывает лишь вклад инновационных мероприятий. Этот подход оправдан, поскольку оценивание производится для поддержки принятия решений о финансировании именно инновационных мероприятий.

Даже в случае чисто инновационной программы инерционному сценарию соответствует динамика, а не постоянные значения показателей, поскольку учитывается действие естественных процессов, происходящих в стране и в мире. Значения показателей, соответствующие инерционному сценарию, будем называть также *базовыми прогнозными значениями показателей*.

Для каждой целевой программы задается ряд значений тех показателей уровня достижения целей Правительства, на которые влияет реализация данной программы. Все целевые и прогнозные значения задаются на конец планового периода целевой программы.

По каждому показателю указывается значение за текущий год, а также значения: целевое ( $A$ ), инерционное ( $a$ ) и программные — без учета риска и с учетом риска ( $a'$ ) за последний год планового



периода. Эти данные представлены в Сводном докладе о результатах и основных направлениях деятельности Правительства Российской Федерации [12] (см. также документ [15]) и в Докладе о результатах и основных направлениях деятельности соответствующего субъекта бюджетного планирования [16]; значения показателей за последний год планового периода в случае реализации программы (без учета риска) содержатся в тексте целевой программы. Необходимо выделить значение, достигаемое именно благодаря реализации рассматриваемой целевой программы. Если достижение этого значения есть результат совместной реализации комплекса целевых программ и вычленив исключительный вклад рассматриваемой программы в изменение показателя невозможно, то следует ограничиться оценением вклада этого комплекса программ в достижение целей Правительства. Действительно, вклад каждой из программ невозможно вычленив в том и только том случае, когда программы настолько тесно взаимодействуют, что не могут реализовываться по отдельности. Но тогда и рассматривать их следует в комплексе.

Необходимо учесть не только показатели, значения которых улучшаются при реализации целевой программы, но и показатели (например, экологические, социологические и др.), на значения которых реализация программы может оказать негативное влияние. Вклад программы в изменение значений этих показателей отрицателен.

## 2.2. Учет факторов риска

Оптимистический сценарий выполнения бюджетной целевой программы может быть нарушен реализацией факторов риска. Поэтому следует выделить основной сценарий реализации программы, при котором факторы риска не реализуются, и не менее одного сценария реализации факторов риска. Дается экспертная оценка вероятности каждого из сценариев:  $p_0$  (вероятность основного сценария),  $p_1, \dots, p_k$ . Сценарии образуют полный набор взаимоисключающих событий, поэтому  $p_0 + p_1 + \dots + p_k = 1$ .

Для каждого из показателей уровня достижения целей Правительства, на которые влияет реализация данной целевой программы, указываются:  $z_0$  — значение для основного сценария, взятое из текста программы, а также экспертные прогнозные значения для выделенных сценариев при реализации факторов риска:  $z_1, \dots, z_k$ .

Среднее значение показателя с учетом факторов риска находится по формуле

$$z' = p_0 z_0 + p_1 z_1 + \dots + p_k z_k. \quad (1)$$

## 2.3. Расчет вклада целевой программы в изменение значений показателей

Введем следующие обозначения:  $A_i$ ,  $a_i$  и  $a'_i$  — соответственно целевое, инерционное (базовое прогнозное) и программное с учетом риска (см. формулу (1)) значения  $i$ -го показателя за последний год планового периода ведомственной целевой программы. Здесь индекс  $i$  нумерует те показатели уровня достижения целей Правительства, на которые влияет реализация данной целевой программы.

Прогнозируемый вклад целевой программы в изменение значения  $i$ -го показателя может быть оценен величиной (см., например, [17, 18])

$$\delta_i = \frac{a'_i - a_i}{A_i - a_i}. \quad (2)$$

Знаменатель дроби (2) есть приращение, необходимое для достижения цели по  $i$ -му показателю. Числитель есть приращение значения показателя, которое ожидается в результате реализации данной программы. Таким образом, найденное отношение  $\delta_i$  показывает, в какой степени достижение цели будет обеспечено реализацией данной программы. Эта интерпретация, по существу, сохраняется и в случае, если числитель или знаменатель (или они оба) отрицательны.

## 2.4. Получение приблизительной оценки вклада целевой программы в достижение цели Правительства

Цели Правительства РФ представляют собой трехуровневую систему [12]. Обозначим через  $m$  общее число показателей уровня достижения некоторой цели третьего уровня. Пусть для рассматриваемой целевой программы найдены значения  $\delta_1, \delta_2, \dots, \delta_m$  прогнозируемых вкладов в изменение значений этих показателей. Для показателей, на значения которых реализация данной программы не влияет,  $\delta_i = 0$ .

Вклад рассматриваемой целевой программы в достижение данной цели может быть в первом приближении оценен величиной

$$\delta = \frac{\delta_1 + \delta_2 + \dots + \delta_m}{m}, \quad (3)$$

имеющей ясную интерпретацию. Если достижение цели в среднесрочной перспективе [19] сводится к достижению целевых значений по всем показателям уровня достижения цели к концу периода планирования, то величина  $\delta$  представляет собой оценку долевого вклада рассматриваемой программы в решение этой задачи.

Оценка (3) весьма приближительна в частности потому, что она основана на допущении равной значимости достижения целевых значений по всем показателям уровня достижения цели. Более точная оценка имеет вид

$$\delta' = w_1\delta_1 + w_2\delta_2 + \dots + w_m\delta_m, \quad (4)$$

где  $w_1, w_2, \dots, w_m$  — весовые коэффициенты важности показателей, удовлетворяющие ограничению  $w_1 + w_2 + \dots + w_m = 1$ . Это наиболее распространенная аддитивная модель ценности (полезности) [18, 20–23]. Метод нахождения весовых коэффициентов описан далее.

### 2.5. Определение весовых коэффициентов важности для показателей уровня достижения цели Правительства

Коэффициенты важности показателей не могут быть получены сравнением содержательных категорий, выражаемых показателями. Это связано с тем, что весовые коэффициенты обратно пропорциональны масштабам шкал показателей, а прямое сравнение содержательных категорий едва ли может обеспечить должную чувствительность к масштабам шкал. Корректный метод получения весовых коэффициентов основан на сравнении важности приращений по шкалам показателей [23–25].

Сравнения производятся экспертом, имеющим четкое представление о приоритетах Правительства, в два этапа.

На первом этапе среди показателей уровня достижения данной цели эксперт выбирает показатель, который будет служить базой для сравнения. Это должен быть достаточно наглядный показатель, желательно, такой, что продвижение от инерционного значения к целевому по этому показателю представляется не самым большим по значимости среди показателей (в противном случае указание эквивалентных продвижений по другим показателям может вызвать трудности). Обозначим его  $B$  (от слова «базовый»), инерционное значение этого показателя через три года обозначим  $b$ , целевое значение через три года обозначим  $B$ .

На втором этапе эксперт сравнивает показатели уровня достижения данной цели с выбранным базовым показателем. Пусть  $\Pi$  — показатель, сравниваемый с базовым. Обозначим через  $x$  и  $X$ , соответственно, его инерционное и целевое значения через три года. Выбирая оценку, эксперт руководствуется следующей инструкцией.

Представьте себе, что Вам нужно сравнить две гипотетических ситуации, которые могут возникнуть к концу планового периода.

В ситуации 1 базовый показатель выведен на целевой уровень  $B$ , но показатель  $\Pi$  остался на инерционном уровне  $x$ .

В ситуации 2 базовый показатель остался на инерционном уровне  $b$ , а показатель  $\Pi$  выведен на некоторый уровень  $x'$ .

Укажите численно как можно точнее значение  $x'$ , при котором Вы сочтете, что ситуации 1 и 2 при прочих равных условиях равноценны — нет оснований предпочесть ни одну из них.

Чем меньше экспертная оценка  $x'$  отличается от инерционного значения  $x$ , тем более важен показатель  $\Pi$  по сравнению с базовым показателем  $B$ . Действительно, чем важнее показатель  $\Pi$ , тем меньше его приращение  $x' - x$ , приравненное по важности к приращению базового показателя  $B - b$ . Пусть  $\delta_B^{(1)}$  и  $\delta_\Pi^{(1)}$  — соответственно вклады целевой программы в изменение значений базового показателя и показателя  $\Pi$  в ситуации 1,  $\delta_B^{(2)}$  и  $\delta_\Pi^{(2)}$  — вклады программы в изменение значений базового показателя и показателя  $\Pi$  в ситуации 2. Приняв за единицу коэффициент важности показателя  $B$  и обозначив через  $\alpha$  коэффициент важности показателя  $\Pi$ , условие равноценности ситуаций 1 и 2 можно записать в виде

$$1 \cdot \delta_B^{(1)} + \alpha \delta_\Pi^{(1)} = 1 \cdot \delta_B^{(2)} + \alpha \delta_\Pi^{(2)}. \quad (5)$$

Заметим теперь, что  $\delta_B^{(1)} = \frac{B-b}{B-b} = 1$ ,

$$\delta_\Pi^{(1)} = \frac{x-x}{X-x} = 0, \quad \delta_B^{(2)} = \frac{b-b}{B-b} = 0, \quad \delta_\Pi^{(2)} = \frac{x'-x}{X-x}.$$

Подставив эти значения в формулу (5), получим

$$\alpha = \frac{X-x}{x'-x}.$$

Этим методом могут быть найдены относительные весовые коэффициенты  $\alpha_1, \alpha_2, \dots, \alpha_m$  всех показателей уровня достижения данной цели (при этом относительный весовой коэффициент базового показателя равен единице). Далее они нормируются делением на их сумму:

$$w_i = \frac{\alpha_i}{\alpha_1 + \dots + \alpha_m}, \quad i = 1, \dots, m, \quad (7)$$

тем самым получаем нормированные весовые коэффициенты показателей уровня достижения цели.

*Замечание 1.* Задача, которая ставится перед экспертом в данной процедуре оценивания, далеко не проста и на первых порах потребует от него определенных интеллектуальных усилий. Однако, приобретая навык оценивания, он сможет двигаться вперед быстрее и увереннее. Важно, что эта процедура оценивания, рассматриваемая как про-



цедура принятия решений, «приближена к реальности». Действительно, в ее основе лежит операция выбора более предпочтительной из двух ситуаций — это базовая операция принятия решений любым управленцем, представителем законодательной или исполнительной власти. К серии таких операций сравнения сводится ключевая для данной процедуры операция нахождения равноценных по предпочтительности ситуаций. Отметим, что без подобной процедуры экспертного оценивания обойтись, по сути, невозможно. Действительно, альтернативными подходами являются присвоение всем показателям одинаковых весов, что приводит к существенным искажениям, и «умозрительное» сравнение важности показателей (без сравнения конкретных приращений), что практически невозможно осуществить корректно с точки зрения анализа данных. Действительно, вопросы типа «Какой из двух показателей Вы считаете более важным?» не имеют точного верифицируемого смысла. Как можно сказать, например, что важнее, сокращение бедности или повышение безопасности? Ответ зависит от того, насколько сокращается бедность и насколько повышается безопасность. А если эти приращения заданы, то возникает корректная и лежащая в основе процедур принятия решений операция сравнения ситуаций, которая и используется в рассматриваемой процедуре экспертного оценивания.

Без экспертного оценивания можно было бы обойтись, если бы был задан выходной показатель качества работы Правительства — тогда можно было бы построить и исследовать статистическую модель его зависимости от частных входных показателей. Но такие выходные показатели пока не заданы.

*Замечание 2.* Качество итоговых оценок, вообще говоря, повысится при работе нескольких экспертов. Это связано, прежде всего, с тем, что цели Правительства затрагивают разные стороны человеческой деятельности и невозможно быть экспертом во всём. Далее, использование оценок нескольких экспертов с последующим усреднением оценок или взаимодействием экспертов [26] позволяет снизить влияние субъективных факторов и повысить точность. Желательно, чтобы, кроме числовых экспертных оценок, эксперты представляли их краткое обоснование. Точность оценок также может быть повышена путем проведения ряда контрольных сравнений показателей: эти сравнения дополняют сравнения показателей с базовым показателем. При этом возникает задача агрегирования результатов парных сравнений; методы решения этой задачи представлены в работах [27—31].

*Замечание 3.* Если эксперту трудно дать точечную оценку величины  $x'$ , задание можно несколько упростить, включив в инструкцию следующую заключительную часть.

Если для Вас очень сложно указать значение  $x'$ , Вы можете указать интервал вида  $[x'_{\min}, x'_{\max}]$ , содержащий это значение. Следует учесть, однако, что чем шире интервал, тем менее информативен Ваш ответ. Поэтому, выбрав интервал, рассмотрите возможность указать в качестве оценки его центральную точку. Если всё же такой ответ кажется Вам неадекватным, постарайтесь сузить интервал настолько, насколько это возможно.

Данные интервального вида потребуют несколько более громоздких методов обработки, но принципиально эти методы не будут отличаться от обсуждаемых в настоящей работе.

*Замечание 4.* Предлагаемый подход легко может быть модифицирован таким образом, что целевые значения показателей не будут необходимы для расчетов. Однако при наличии целевых значений результаты всех вычислений, включая промежуточные, легко интерпретируются. В частности, после нормировки показателей, использующей их целевые (плановые) значения, вклады (2) в достижение целей и весовые коэффициенты показателей приобретают легко интерпретируемую безразмерную форму.

## 2.6. Уточнение оценки вклада целевой программы в достижение цели Правительства

Полученные весовые коэффициенты показателей уровня достижения цели могут быть подставлены в формулу (4), с помощью которой оценивается вклад целевой программы в достижение цели, характеризуемой  $m$  показателями. Это приводит к повышению точности оценивания. Дальнейшее уточнение может быть получено посредством учета степени независимости показателей, характеризующих достижение одной и той же цели.

Под независимостью здесь понимается следующее. Если изменение значения первого показателя само по себе не означает и в среднем не приводит к изменению значения второго показателя, и наоборот, то эти показатели независимы. При независимых показателях формула (4) с оцененными ранее весовыми коэффициентами не нуждается в уточнении (в предположении линейности связи  $\delta'$  с  $\delta_1, \dots, \delta_m$ ) [18].

Пусть теперь показатель  $P_1$  с инерционным значением через три года  $a_1$  и целевым значением через три года  $A_1$  и показатель  $P_2$  с инерционным значением через три года  $a_2$  и целевым значением через три года  $A_2$  таковы, что приращение  $A_1 - a_1$  значения показателя  $P_1$  само по себе в среднем приводит к приращению  $\beta(A_2 - a_2)$  значения показателя  $P_2$ , где  $\beta$  — некоторое число, которое устанавливается анализом набора временных ря-

дов или экспертной оценкой. Тогда вычисление вклада ведомственной целевой программы в изменение значения показателя  $P_2$  нуждается в уточнении. Согласно изложенному в п. 2.3 этот вклад исчисляется по формуле (2).

Теперь учтем, что из достигнутого в результате реализации программы приращения  $a'_2 - a_2$  часть, равная  $\beta(A_2 - a_2) \frac{a'_1 - a_1}{A_1 - a_1}$ , получена в результате достижения приращения  $a'_1 - a_1$  по показателю  $P_1$ . Поэтому независимая часть приращения по показателю  $P_2$  равна  $(a'_2 - a_2) - \beta(A_2 - a_2) \frac{a'_1 - a_1}{A_1 - a_1}$ . Подстановка ее в качестве числителя в формулу для  $\delta_2$  дает уточненное выражение

$$\delta'_2 = \frac{a'_2 - a_2}{A_2 - a_2} - \beta \frac{a'_1 - a_1}{A_1 - a_1} = \delta_2 - \beta \delta_1.$$

Это уточненное значение подставляется вместо  $\delta_2$  в формулу (4) вычисления вклада программы в достижение цели.

Учитывая таким образом влияние показателя  $P_1$  на показатель  $P_2$ , уже не следует одновременно учитывать влияние показателя  $P_2$  на показатель  $P_1$ . Учитывать влияние на показатель  $P_2$  более одного показателя можно только если влияющие на него показатели сами независимы – в противном случае требуются более сложные процедуры учета.

### 2.7. Оценка вклада целевой программы в достижение целей Правительства, отнесенных ко второму и первому уровням, а также в общую систему целей

Реализация процедуры оценивания вклада целевой программы в достижение целей второго и первого уровней зависит от того, имеются ли в системе целей и показателей деятельности Правительства показатели уровня достижения целей первых двух уровней, и если да, то каковы они. В системе целей, представленной в Сводном докладе о результатах и основных направлениях деятельности Правительства Российской Федерации на 2006–2008 гг. [12], таких показателей за редким исключением нет, и вклад программ в достижение целей второго и первого уровней может быть оценен только агрегированием вкладов в достижение целей третьего уровня.

Если показатели, характеризующие первые два уровня целей, будут добавлены, станет возможным исследовать вопрос о применимости статистических алгоритмов анализа данных (регрессионного анализа и др.) для прогнозирования изменений показателей второго и первого уровней по измене-

ниям входных показателей третьего уровня, характеризующих реализацию целевых программ.

Общая процедура оценивания вклада целевой программы в достижение цели второго или первого уровня или в общую систему целей Правительства состоит в следующем.

Обозначим через  $\delta_1, \delta_2, \dots, \delta_M$  прогнозируемые вклады целевой программы в изменение значений всех показателей уровня достижения целей третьего уровня, относящихся к данной цели второго уровня, цели первого уровня или входящих в общую систему целей. Здесь  $M$  – число рассматриваемых показателей. Если некоторые показатели встречаются несколько раз, характеризуя уровни достижения разных целей, каждый из них включается в список лишь один раз. Метод нахождения вкладов  $\delta_1, \delta_2, \dots, \delta_M$  изложен выше; для показателей уровня достижения цели третьего уровня их вклады уточнены посредством использования информации о зависимости показателей. Поскольку каждая целевая программа направлена на решение конкретных задач, большинство вкладов  $\delta_1, \delta_2, \dots, \delta_M$  равны нулю.

Вклад программы в достижение цели второго или первого уровня или в общую систему целей Правительства оценивается выражением  $\delta = w'_1 \delta_1 + w'_2 \delta_2 + \dots + w'_M \delta_M$ , где  $w'_1, w'_2, \dots, w'_M$  – весовые коэффициенты, удовлетворяющие ограничению  $w'_1 + w'_2 + \dots + w'_M = 1$ . Эти весовые коэффициенты находятся с помощью того же алгоритма, что и коэффициенты  $w_1, w_2, \dots, w_m$  в п. 2.5.

Для каждой цели третьего уровня выделяется базовый показатель. Для каждого показателя, характеризующего ту же цель третьего уровня, находится коэффициент  $\alpha$ , как в п. 2.5 (нужно воспользоваться найденными ранее значениями).

Далее сравниваются базовые показатели целей третьего уровня. Среди них выбирается один наглядный и легко интерпретируемый *главный базовый показатель* и находятся относительные весовые коэффициенты  $\alpha''$  сравнения других базовых показателей с ним.

Коэффициенты  $\alpha$  всех показателей пересчитываются по отношению к главному базовому показателю: если коэффициент  $\alpha_i$  получен в результате сравнения данного показателя с базовым показателем, характеризующим ту же цель третьего уровня, а  $\alpha''$  получен в результате сравнения последнего с главным базовым показателем, то сравнение данного показателя с главным базовым показателем характеризуется коэффициентом  $\alpha'_i = \alpha_i \cdot \alpha''$ .

Когда найдены все относительные весовые коэффициенты  $\alpha'_1, \alpha'_2, \dots, \alpha'_M$  сравнения показателей с главным базовым показателем (сам он при





этом получает относительный весовой коэффициент 1), весовые коэффициенты нормируются делением на их сумму:

$$w'_i = \frac{\alpha'_i}{\alpha'_1 + \dots + \alpha'_M}, \quad i = 1, \dots, M.$$

Тем самым найдены весовые коэффициенты показателей уровня достижения целей, и вклад целевой программы в достижение цели второго или первого уровня или в общую систему целей Правительства оценивается выражением  $\delta = w'_1 \delta_1 + w'_2 \delta_2 + \dots + w'_M \delta_M$ .

Достоинство этой оценки в том, что она имеет ясную интерпретацию. Это оценка долевого вклада целевой программы в решение полной совокупности задач, характеризующих имеющейся системой показателей и набором инерционных и целевых значений. Частной задачей является достижение целевого значения по отдельному показателю, а общая задача сводится к полной совокупности частных задач. Значение  $\delta$ , даже выраженное в процентах, может оказаться малым. Например, если получена итоговая оценка  $\delta = 0,033 \%$ , это интерпретируется так: данная целевая программа позволяет решить примерно одну трехтысячную часть всех задач, стоящих перед Правительством в трехлетней перспективе.

### 2.8. Об оценивании вклада целевой программы в достижение целей Правительства в долгосрочной перспективе

Выше рассмотрена процедура оценивания вклада ведомственной целевой программы в достижение целей Правительства в среднесрочной, а именно, трехлетней, перспективе [19]. Аналогично можно оценить вклад программы в достижение долгосрочных целей правительства. Для этого нужно в формулах, в которые входят целевые значения показателей через три года, заменить их на долгосрочные целевые значения тех же показателей. Последние для многих показателей указаны в Сводном докладе о результатах и основных направлениях деятельности Правительства [12] и в докладах о результатах и основных направлениях деятельности субъектов бюджетного планирования [16].

Разумеется, вклад каждой программы в достижение долгосрочных целей окажется меньше, чем ее вклад в реализацию среднесрочных целей, и ранжирования целевых программ по двум этим показателям могут не совпасть. Последнее обстоятельство связано с тем, что при долгосрочном планировании целевых значений по разным показателям зачастую применяются совершенно различные подходы; сроки планирования также могут существенно различаться.

### 2.9. Показатели экономической эффективности целевой программы

При формировании перечня целевых программ, финансируемых из бюджета, необходимо знать не только оценки вклада целевых программ в достижение целей Правительства, но и значения показателей экономической эффективности программ. Коснемся использования простейших показателей экономической эффективности: чистого приведенного дохода, срока окупаемости и рентабельности инвестиций (см., например, рекомендации [32]).

При расчете чистого приведенного дохода измерение разновременных показателей осуществляется путем приведения (дисконтирования) их к начальному периоду. Для приведения разновременных затрат, результатов и эффектов используется норма дисконта ( $g$ ). При оценивании экономической эффективности целевых программ значение  $g$  обычно выбирают равным прогнозируемому уровню инфляции.

Приведение к начальному моменту затрат, результатов и эффектов, относящихся к концу  $k$ -го года реализации программы, производится  $k$ -кратным умножением на дисконтный множитель (коэффициент дисконтирования), определяемый в случае постоянной процентной ставки (нормы дисконта)  $g$  как  $V = 1/(1 + g)$ . Напомним, что чистый приведенный доход — это разность дисконтированных на один и тот же момент времени показателей дохода и капиталовложений. Он может быть вычислен по формуле

$$W = \sum_{i=1}^n E_i V^i - \sum_{i=1}^n K_i V^{i-1}, \quad (6)$$

где  $E_i$  и  $K_i$  — соответственно, доход и инвестиционные расходы за год  $i$ ,  $V$  — коэффициент дисконтирования,  $n$  — срок реализации программы в годах. Формула (6) записана для случая, когда инвестиционные расходы относятся к началу каждого года, а доходы — к концу года. Если величина  $W$  отрицательна, т. е. приведенные расходы превосходят приведенные доходы, удобно говорить о *чистых приведенных расходах*, равных ( $-W$ ).

Срок окупаемости без учета дисконтирования рассчитывается как  $N = K/R$ , где  $K$  — общий объем инвестиций,  $R$  — средний ежегодный доход.

Рентабельность инвестиций есть отношение приведенных доходов к приведенным инвестиционным расходам. Для трехлетней программы рентабельность инвестиций может быть рассчитана для каждого из трех этапов инвестирования по формуле  $u_i = E_i V_i / K_i$ ,  $i = 1, 2, 3$ , где коэффициенты дисконтирования  $V_1, V_2, V_3$ , вообще говоря, могут различаться.

Найденные значения вкладов программ в достижение целей Правительства и значения показателей экономической эффективности программ являются исходными данными для формирования перечня целевых программ, финансируемых из бюджета. Принципам формирования указанного перечня мы планируем посвятить отдельную публикацию.

### ЗАКЛЮЧЕНИЕ

Предложена процедура расчетно-экспертного оценивания вкладов ведомственных целевых программ в достижение целей Правительства Российской Федерации. Процедура основана на оценивании вкладов каждой целевой программы в изменение значений показателей уровня достижения каждой цели Правительства. Вклад определяется соотношением инерционных, целевых и программных (с учетом риска) значений по этим показателям. Разработана экспертная процедура определения весовых коэффициентов важности показателей, а также процедура учета их взаимосвязи. Итоговая оценка вклада целевой программы в достижение целей Правительства интерпретируется как взвешенная доля задач, решаемых реализацией данной программы, в общем объеме задач, стоящих перед Правительством в среднесрочной перспективе.

### ЛИТЕРАТУРА

1. Концепция реформы бюджетного процесса в РФ в 2004 — 2006 гг., одобренная Постановлением Правительства Российской Федерации от 22 мая 2004 г. № 249.
2. Реформа исполнения бюджета: бюджетирование, ориентированное на результат (БОР), в условиях в условиях казначейского исполнения бюджета / Под ред. М.А. Клишиной. — М.: РОХОС, 2004.
3. Практика внедрения бюджетирования, ориентированного на результат, в субъектах Российской Федерации / Под ред. М.А. Клишиной. — М.: ЛЕНАНД, 2005.
4. Постановление Правительства Российской Федерации от 30 апреля 2004 г. № 225 «О Правительственной комиссии по повышению результативности бюджетных расходов».
5. Положение о разработке, утверждении и реализации ведомственных целевых программ, утвержденное Постановлением Правительства Российской Федерации от 19 апреля 2005 г. № 239.
6. Положение о разработке перспективного финансового плана Российской Федерации и проекта федерального закона о федеральном бюджете на очередной финансовый год, утвержденное Постановлением Правительства Российской Федерации от 6 марта 2005 г. № 118.
7. *The Green Book: Appraisal and evaluation in central government.* — London: TSO, 2004.
8. *Merna T., Njiru C. Financing infrastructure projects.* — London: Thomas Telford, 2002.
9. *Cost-Benefit Analysis: Concepts and practice.* 3rd Ed. / A.E. Boardman, et al. — Prentice Hall, 2006.
10. *Aravossis K., Koutsiana E. Program evaluation methodologies. A comparative assessment // Social Science Tribune, Special Issue «Economy and Society», Papazisis Edition.* — 2003. — Vol. 10, N 37. — P. 97—115.
11. *Rogers M. Engineering project appraisal: The evaluation of alternative development schemes,* Oxford, UK: Blackwell Science, 2001.
12. Сводный доклад о результатах и основных направлениях деятельности Правительства Российской Федерации на 2006 — 2008 гг., рассмотренный на заседании Правительства Российской Федерации 15 июня 2007 г. <www.economy.gov.ru>.
13. *Липсиц И.В., Коссов В.В. Инвестиционный проект: методы подготовки и анализа.* — М.: БЕК, 1996.
14. *Ковалев В.В. Методы оценки инвестиционных проектов.* — М.: Финансы и статистика, 1998.
15. *Положение о разработке Сводного доклада о результатах и основных направлениях деятельности Правительства Российской Федерации на 2006 — 2008 гг., утвержденное Постановлением Правительства Российской Федерации от 27 апреля 2005 г. № 259.*
16. *Методические рекомендации по подготовке Докладов о результатах и основных направлениях деятельности субъектов бюджетного планирования на 2006 — 2008 гг. (Утверждены Минэкономразвития России и Минфином России).*
17. *Волкович В.Л. Многокритериальные задачи и методы их решения // Тр. семинара «Кибернетика и вычислительная техника», Сложные системы управления.* — Киев: Наукова думка, 1969. — Вып. 1
18. *Дубов Ю.А., Травкин С.И., Якимец В.Н. Многокритериальные модели формирования и выбора вариантов систем.* — М.: Наука, 1986.
19. *Программа социально-экономического развития Российской Федерации на среднесрочную перспективу (2006—2008 гг.), утвержденная Распоряжением Правительства РФ от 19.01.2006 № 38-р.*
20. *Готов В.А., Павельев В.В. Экспертные методы определения весовых коэффициентов // Автоматика и телемеханика.* — 1976. — № 12. — С. 95—107.
21. *Емельянов С.В., Ларичев О.И. Многокритериальные методы принятия решений.* — М.: Знание, 1985.
22. *Кини Р., Райфа Х. Принятие решений при многих критериях: предпочтения и замещения.* — М.: Радио и связь, 1981.
23. *Поддиновский В.В., Ногин В.Д. Парето-оптимальное решение многокритериальных задач.* — М.: Наука, 1982.
24. *Поддиновский В.В. Коэффициенты важности критериев в задачах принятия решений. Порядковые или ординальные коэффициенты важности // Автоматика и телемеханика.* — 1978. — № 10. С. 130—141.
25. *Поддиновский В.В. Количественная важность критериев // Там же.* — 2000. — № 5. — С. 110—123.
26. *Панкова Л.А., Петровский А.М., Шнейдерман М.В. Организация экспертизы и анализ экспертной информации.* — М.: Наука, 1984.
27. *Экспертные оценки. Методы и применение (обзор) / Шмерлинг Д.С. и др. // Статистические методы анализа экспертных оценок.* — М.: Наука, 1977. — С. 290—382.
28. *Литвак Б.Г. Экспертная информация: методы получения и анализа.* — М.: Радио и связь, 1981.
29. *Пригарина Т.А., Чеботарев П.Ю. Методы экспертных оценок и определение предпочтительности объектов // Экспертные оценки в социологических исследованиях.* — Киев: Наукова думка, 1990. — С. 190—225.
30. *Пригарина Т.А., Чеботарев П.Ю., Шмерлинг Д.С. Парные сравнения объектов (Аналитический обзор) // Научно-Техническая Информация, Сер. 2. Информационные процессы и системы.* — 1996. — № 2. — С. 20—32.
31. *Ларичев О.И., Мошкович Е.М. Качественные методы принятия решений.* — М.: Физматлит, 1996.
32. *Методические рекомендации по оценке эффективности инвестиционных проектов, утвержденные Минэкономки РФ, Минфином РФ, Госстроем РФ 21.06.1999, № ВК 477.*

☎(495) 334-88-69, 650-84-76; e-mail: pchv@rambler.ru, Mitkin@economy.gov.ru, schmerling@hse.ru

Статья представлена к публикации членом редколлегии Р.М. Нижегородцевым. □



# КЛАССИФИКАЦИЯ ОБЪЕКТОВ ПРОФЕССИОНАЛЬНОЙ ДЕЯТЕЛЬНОСТИ СПЕЦИАЛИСТА ПРИ ПРОЕКТИРОВАНИИ ПРОФЕССИОНАЛЬНЫХ И ОБРАЗОВАТЕЛЬНЫХ СТАНДАРТОВ

В.В. Никитин<sup>(1)</sup>, С.В. Мальцева<sup>(1)</sup>, А.А. Дорофеев<sup>(2)</sup>, А.С. Мандель<sup>(2)</sup>, А.Л. Чернявский<sup>(2)</sup>

<sup>(1)</sup>Государственный университет — Высшая школа экономики, г. Москва;

<sup>(2)</sup>Институт проблем управления им. В.А. Трапезникова РАН, г. Москва

Показано, что задача формирования обобщающих понятий, характеризующих сферу профессиональной деятельности специалиста, сводится к задаче диагонализации некоторой матрицы связей. Предложена человеко-машинная процедура её решения.

## ВВЕДЕНИЕ

Одна из центральных задач управления процессами подготовки профессиональных кадров заключается в разработке государственных профессиональных и образовательных стандартов. Решение этой задачи представляет собой многоэтапный процесс анализа, оценивания и обработки больших массивов информации [1].

Важный начальный этап этого процесса состоит в определении объектов профессиональной деятельности (ОПД) специалиста, относительно которых «выстраиваются» его профессиональная деятельность и профессиональные компетенции. Образовательные программы, в свою очередь, должны быть построены таким образом, чтобы обеспечивать формирование у обучающихся заданного набора компетенций.

Формирование множества ОПД для определенной предметной области и их классификация представляют собой слабо формализуемую и трудоемкую задачу. Это, в первую очередь, сказывается на формировании содержания образования в высокотехнологичных, динамически развивающихся отраслях, для которых характерны частое появление новых объектов и связанных с ними понятий, существование большого числа синонимичных понятий, быстрое устаревание некоторых типов объектов, изменения в описании и интерпретации отдельных понятий. Особенно это про-

является на комплексных объектах, которые являются сложной агрегацией множества простых объектов и определяются обобщающими понятиями. Важно, что именно эти понятия, как правило, используются в качестве определений ОПД в стандартах высшего профессионального образования.

Сегодня общепризнанным способом идентификации таких сложных объектов служат онтологии, на основе которых можно определить не только набор более простых понятий, составляющих сложное понятие, но и связи между сложными понятиями, а также степень сходства и различия между понятиями, что позволяет выделить набор обобщающих понятий, на основе которых можно идентифицировать основные ОПД. Эти понятия являются образующими понятиями для формирования набора компетенций специалиста.

Создание онтологий в различных предметных областях представляет собой одну из самых актуальных задач. Её решением на разных уровнях занимаются исследовательские центры и международные организации. Примером одного из наиболее крупных проектов может служить проект UN SPSC (United Nations Standard Products and Services Code — Стандартный классификатор товаров и услуг), разработка которого ведется экспертной группой UNDO (United Nations Development Organization) и крупной компанией «Dun & Bradstreet». Проект развивается на основе опыта использования транснациональных систем и стандартов

(SWIFT, EDIFACT и др.) [1]. В настоящее время этот классификатор насчитывает более 13 тыс. категорий товаров и услуг.

Задача формирования обобщающих понятий, используемых при проектировании профессиональных и образовательных стандартов для той или иной предметной области, может быть понята как задача формирования классов простых (исходных) ОПД. Далее под ОПД всюду понимаются именно исходные ОПД. Сложность этой задачи связана, прежде всего, с объемом исходной информации. Число ОПД исчисляется сотнями, а число классов, в которые их нужно объединить, — десятками. При решении задачи вручную эксперт вынужден резко ограничивать число рассматриваемых вариантов классификации, основываясь на своих субъективных оценках. Для повышения объективности и обоснованности классификации желательно использовать формализованные критерии и многовариантные процедуры.

### 1. ФОРМАЛИЗАЦИЯ ПОСТАНОВКИ ЗАДАЧИ

Вначале введем необходимые обозначения:

$V = \{v_n \mid n = 1, \dots, N_V\}$  — множество ОПД, элемент множества представляет собой код ОПД;

$\chi = \{\chi_n \mid n = 1, \dots, N_V\}$  — вектор наименований ОПД, элементы вектора — наименования ОПД;

$\mu = \{\mu_n \mid n = 1, \dots, N_V\}$  — вектор описаний ОПД, элементы вектора — текстовые описания ОПД.

Пусть  $R = \{r_{ij}\}$ ,  $i, j = 1, \dots, N_V$ ,  $i \neq j$ , — матрица связей между ОПД, элемент  $r_{ij}$  которой служит численной характеристикой степени связи между  $i$ -м и  $j$ -м ОПД ( $0 \leq r_{ij} \leq 1$ ).

Цель обработки матрицы связей состоит в разбиении всего множества элементов  $v_1, \dots, v_{N_V}$  на такие непересекающиеся подмножества (агрегаты)  $q'_1, \dots, q'_{N'}$ , чтобы связи  $r_{ij}$  между элементами, попавшими в одно подмножество, были возможно больше, а между элементами, попавшими в разные подмножества, — возможно меньше. Такую задачу будем называть задачей классификации (агрегирования) исходного множества элементов, а полученные классы (агрегаты)  $q'_1, \dots, q'_{N'}$  интерпретировать как классы объектов профессиональной деятельности.

Эту же задачу можно описать и несколько иначе. Если различным образом нумеровать элементы  $v_1, \dots, v_{N_V}$ , это будет соответствовать перестановкам строк и столбцов матрицы связей  $R$ . Рассмотрим задачу нахождения такой нумерации элементов (т. е. такого расположения строк и соответствующих столбцов матрицы  $R$  и выделения таких непересекающихся квадратных подматриц вдоль

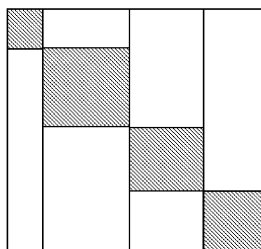


Рис. 1. Диагонализированная матрица связей

главной диагонали преобразованной подобным образом матрицы связей, чтобы элементы — числа  $r_{ij}$  — каждой выделенной подматрицы были возможно больше, а числа  $r_{ij}$ , расположенные вне этих подматриц, — возможно меньше. Будем также требовать, чтобы выделенные квадратные подматрицы полностью покрывали главную диагональ матрицы связи. Условно выделенные подматрицы с большими компонентами можно изобразить в виде квадратов, вообще говоря, разного размера (рис. 1), расположенных вдоль главной диагонали матрицы  $R$ .

Подмножество элементов, которым соответствуют строки (или столбцы) преобразованной матрицы связей  $R$ , образующие одну из выделенных подматриц, — это и есть один из классов (сильно связанных агрегатов), о которых шла речь. Имея в виду такую интерпретацию, рассматриваемую задачу называют также задачей диагонализации матрицы связей.

Приведенное описание задачи ещё не есть точная её постановка, так как последняя предполагает формальное определение того, что означает требование, чтобы в выделенных подматрицах «компоненты были возможно большими, а вне их — возможно меньшими». С этой целью обычно вводится в рассмотрение подходящий функционал (критерий качества классификации), зависящий от того, как именно разбиты элементы на  $N'$  классов  $q'_1, \dots, q'_{N'}$ , и такой, что «хороший» в интуитивном понимании способ классификации элементов соответствует экстремальному значению функционала. Как только подходящий функционал сформулирован, задача сводится к конструированию процедуры его экстремизации. Такой подход к решению задачи мы будем называть вариационным.

Этот подход, однако, наталкивается на ряд трудностей, особенно если число классифицируемых элементов достаточно велико (порядка нескольких сотен).

Прежде всего, не удастся предложить функционал, который отражал бы все аспекты нашего интуитивного представления о «хорошей» классификации элементов в самых разных задачах. Тем самым далеко не всегда результаты классификации, получаемые путем экстремизации того или иного



функционала, будут устраивать пользователя. Далее, в такого рода задачах, за весьма редкими и специфическими исключениями, не существует эффективных процедур (т. е. процедур, не сводящихся к полному перебору всех возможных вариантов), которые доставляли бы глобальный экстремум соответствующему функционалу. Поэтому, как правило, можно надеяться только на достижение его локального экстремума. Вместе с тем, с ростом числа элементов растет число локальных экстремумов функционала и тем самым уменьшается вероятность достижения такого локального экстремума функционала, который был бы достаточно близок к его глобальному экстремуму.

В связи с отмеченными недостатками вариационного подхода широкое распространение при решении задачи классификации получили алгоритмы, относительно которых нельзя сказать, какой формальный критерий качества (функционал) они оптимизируют, но которые, по самому своему характеру, отражают те или иные аспекты нашего представления о «хорошей» классификации. Такой подход к решению задачи диагонализации матрицы связи мы будем называть эвристическим.

Далее рассматриваются алгоритмы решения задачи диагонализации матрицы связи, реализующие как вариационный, так и эвристический подходы.

Построение классов ОПД — это лишь один из этапов проектирования профессиональных и образовательных стандартов. В связи с этим к алгоритмам классификации ОПД предъявляется дополнительное требование: на выходе они должны давать не только классификацию (т.е. совокупность классов ОПД), но и матрицу связей между классами  $\Phi = \{\varphi_{ij}\}$ , где  $N$  — число классов, а  $\varphi_{ij}$  — численные характеристики степени связи между классами ОПД, аналогичные характеристикам степени связи  $r_{ij}$  между отдельными ОПД.

## 2. АЛГОРИТМЫ РЕШЕНИЯ ЗАДАЧИ

Описываемые далее алгоритмы состоят из двух этапов:

- 1) построение численных характеристик степени связи (матрицы связей  $R$ );
- 2) классификация (диагонализация матрицы связей).

Первый этап общий для всех алгоритмов.

### 2.1. Построение численных характеристик элементов множества $R$ (матрицы связей)

Набор численных характеристик степеней связи между разными ОПД (элементы матрицы связей  $R = \{r_{ij}\}$ ,  $i, j = 1, \dots, N_v$ ,  $i \neq j$ ) определяется экспертным путем.

Связи между любыми двумя объектами могут быть:

— ассоциативные (близость объектов может быть охарактеризована качественно и (или) количественно);

— агрегатно-ассоциативные (оба объекта входят в состав некоторого агрегата, или объекта «более высокого уровня»);

— композитные (один объект является составной частью второго объекта);

— обобщающие (один объект можно рассматривать как частный случай второго объекта).

Для получения численной оценки связи между ОПД зададим следующие правила экспертного оценивания:

— если эксперт считает, что связь между объектами  $v_i$  и  $v_j$  композитная, полагаем  $r_{ij} = 1$ ;

— если эксперт считает, что связь между объектами  $v_i$  и  $v_j$  агрегатно-ассоциативная или обобщающая, полагаем  $r_{ij} = 0,75$ ;

— если эксперт считает, что связь между объектами  $v_i$  и  $v_j$  ассоциативная, он должен оценить ее числом, лежащим в диапазоне  $0 < r_{ij} < 0,75$ ;

— если связь между объектами  $v_i$  и  $v_j$  отсутствует, полагаем  $r_{ij} = 0$ .

После экспертного оценивания по указанным правилам множество связей между объектами профессиональной деятельности  $v_1, \dots, v_{N_v}$  становится числовым множеством, которое задается матрицей связей  $R = \{r_{ij}\}$ ,  $i, j = 1, \dots, N_v$ ,  $i \neq j$ . Именно матрица связей и есть тот эмпирический материал, который обрабатывается с помощью описываемых далее алгоритмов.

### 2.2. Вариационный подход к решению задачи классификации

Пусть имеются  $N_v$  элементов  $v_1, \dots, v_{N_v}$  и соответствующая им матрица связей  $R = \{r_{ij}\}$ ,  $i, j = 1, \dots, N_v$ ,  $i \neq j$ , все элементы которой неотрицательные числа.

Пользуясь матрицей  $R$ , требуется разбить элементы на  $N'$  подмножеств (классов)  $q'_1, \dots, q'_{N'}$ , где число  $N'$  считается заданным заранее.

В результате ряда теоретических и экспериментальных исследований установлено [2, 3], что в достаточно широком классе задач удовлетворительные результаты классификации достигаются при максимизации функционала

$$F = \sum_{k=1}^{N'} \frac{m_k}{N_v} \left[ \frac{1}{m_k(m_k-1)} \sum_{i,j \in q_k, i \neq j} r_{ij} \right] = \frac{1}{N_v} \sum_{k=1}^{N'} \frac{1}{m_k-1} \sum_{i,j \in q_k, i \neq j} r_{ij}, \quad (1)$$

где  $m_k$  — число элементов в соответствующем классе. Условие  $i \neq j$  введено в функционал (1) для того, чтобы степень связи элемента с самим собой (если таковая имеется) не влияла на результат классификации. Для определенности при  $m_k = 1$

будем полагать  $\frac{1}{m_k - 1} \sum_{i, j \in q'_k, i \neq j} r_{ij} = 0$ . Сумма

$\sum_{i, j \in q'_k, i \neq j} r_{ij}$  есть сумма всех степеней связи между

разными элементами, попавшими в один класс  $q'_k$ .

Величина  $m_k(m_k - 1)$  — общее число таких величин, а число  $m_k/N_v$  — доля элементов, попавших в класс  $q'_k$ , т. е. число, характеризующее «размер» этого класса. Поэтому функционал  $F$  имеет смысл суммы взвешенных средних степеней связи внутри каждого класса, причем коэффициенты взвешивания пропорциональны размерам класса. В связи с этим максимизация функционала  $F$  приводит к более «плотным» (с большей средней степенью связи между элементами) классам большого размера за счет меньшей плотности классов малого размера.

Вопрос о выборе функционала при реализации вариационного подхода совсем не тривиален. Оказывается, что многие функционалы, по смыслу близкие к введенному выше функционалу  $F$ , приводят к результатам, явно противоречащим нашим представлениям о «хорошей» классификации. В качестве такого рода примеров рас-

смотрим функционалы  $F_1 = \sum_{k=1}^N \sum_{i, j \in q'_k, i \neq j} r_{ij}$ ,  $F_2 =$

$$= \sum_{k=1}^N \frac{1}{m_k(m_k - 1)} \sum_{i, j \in q'_k, i \neq j} r_{ij}.$$

Функционал  $F_1$ , на первый взгляд, кажется достаточно разумным, так как его максимизация означает такое разделение элементов на классы, при котором сумма всех связей внутри классов будет максимальной. Вместе с тем, величина  $F_1$  при прочих равных условиях будет тем большей, чем больше элементов матрицы  $A$  попадет в блоки. Поэтому, если бы все элементы матрицы  $A$  были равны между собой, то при  $N = 2$  максимум функционала  $F_1$  достигался бы при таком разбиении элементов на два (непустых) класса, при котором в один класс попадет  $N_v - 1$  элементов, а в другой — только один элемент (рис. 2).

Функционал  $F_2$  имеет смысл суммы средних степеней связей внутри классов, т. е. смысл, весьма близкий к смыслу функционала  $F$ . Вместе с тем, специальные исследования показали [2], что максимизация функционала  $F_2$  в сложных случаях приводит к плохим результатам, хотя до сих пор нет убедительного содержательного объяснения этого факта.

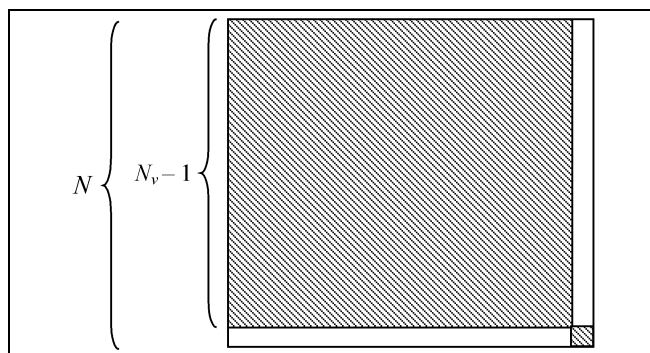


Рис. 2. Неравномерность классификации с функционалом  $F_1$

При достаточно большом числе классов максимуму  $F_2$  часто соответствует разбиение, в котором все классы, кроме одного, состоят из пары наиболее близких элементов. Одна из причин этого заключается в том, что любой  $k$ -й член

$$\frac{1}{m_k(m_k - 1)} \sum_{i, j \in q'_k, i \neq j} r_{ij}$$

функционала  $F_2$  не зависит от числа классов, но быстро уменьшается с ростом  $m_k$ .

**Алгоритм поиска локального экстремума функционала  $F$**  строится следующим образом. Пусть имеется некоторое начальное разбиение множества объектов на классы. На каждом шаге алгоритма осуществляется пробный перенос некоторого очередного элемента из того класса, в котором он находится к данному шагу, последовательно во все остальные классы, начиная с первого. При каждом таком переносе подсчитывается новое значение функционала  $F$  и сравнивается со значением этого функционала до переноса. Если при очередном пробном переносе данного элемента значение функционала возросло, то рассматриваемый элемент остается в новом классе (т. е. фактически переносится из того класса, в котором находился, в новый класс). На этом выполнение данного шага алгоритма заканчивается. Если же после пробных переносов во все другие классы значение функционала  $F$  ни разу не возросло, то рассматриваемый элемент остается в том же классе, в котором он находился до данного шага. Затем алгоритм переходит к следующему шагу, на котором осуществляются пробные переносы следующего элемента. Алгоритм останавливается после того, как просмотр всех элементов не приводит к изменению ни одного из классов. Таким образом, если считать, что «окрестностью» некоторого разбиения является совокупность всех разбиений, отличающихся от данного принадлежностью только одного элемента, то рассмотренный алгоритм доставляет функционалу экстремум, локальный по отношению к такому определению окрестности.



В качестве начального разбиения в этом алгоритме может использоваться любое разбиение элементов на заданное число  $L$  классов.

### 2.3. Эвристический подход к решению задачи классификации

Эффективность эвристических алгоритмов решения задачи классификации зависит от ее сложности. Если элементы действительно группируются в «плотные» классы, а связь между любыми элементами из разных классов существенно меньше, чем между элементами из одного класса, то такие алгоритмы дают хорошее решение задачи. Однако встречающаяся в реальных задачах ситуация редко бывает столь идеальной, так что лишь «в среднем», «как правило» элементы из одного класса сильнее связаны, чем элементы из разных классов. И чем сложнее задача, т. е. чем сильнее она отличается от «идеальной», тем труднее выделить классы и тем сложнее для этого должен быть алгоритм. Рассмотрим один из наиболее распространенных эвристических алгоритмов — иерархический алгоритм классификации «Объединение» [2].

Пусть два подмножества  $q_p$  и  $q_s$  элементов включают в себя, соответственно,  $m_p$  и  $m_s$  элементов. Будем измерять степень связи, или степень «близости» между этими двумя подмножествами величиной

$$K(q_p, q_s) = \frac{1}{m_p m_s} \sum_{i \in q_p} \sum_{j \in q_s} r_{ij}. \quad (2)$$

Каждый шаг алгоритма заключается в объединении в один класс двух наиболее «близких» (в смысле (2)) друг к другу классов, полученных в результате предыдущих шагов алгоритма, так что на каждом шаге число построенных алгоритмом классов уменьшается на единицу. Работа алгоритма продолжается до тех пор, пока не будет получено заранее заданное число  $N$  классов.

В процессе выполнения каждого шага в связи с изменением классов следует также пересчитывать величины  $K(q_r, q_t)$ . Пусть, например, на некотором шаге объединяются классы  $q_p$  и  $q_s$  в один класс, который обозначим через  $q_u$ . Если  $q_r$  и  $q_t$  не есть  $q_u$ , то соответствующая величина  $K(q_r, q_t)$  на данном шаге не изменяется. Если же один из классов, например  $q_r$ , это и есть новый класс, то, как легко видеть,

$$\begin{aligned} K(q_u, q_t) &= \frac{1}{(m_p + m_s)m_t} \sum_{i \in q_p \cup q_s} \sum_{j \in q_t} r_{ij} = \\ &= \frac{m_p K(q_p, q_t) + m_s K(q_s, q_t)}{m_p + m_s}. \end{aligned}$$

В качестве начального разбиения для работы алгоритма «Объединение» можно взять  $N_v$  классов, содержащих каждый по одному элементу. В свою очередь классы, получаемые в результате работы алгоритма «Объединение», можно использовать в качестве начального разбиения для работы описанного ранее вариационного алгоритма.

### 3. КЛАССИФИКАЦИЯ ОБЪЕКТОВ ПРОФЕССИОНАЛЬНОЙ ДЕЯТЕЛЬНОСТИ

В настоящей работе описанные алгоритмы применялись следующим образом. Классификация ОПД проводилась при разных значениях  $N$ : множество  $V$  разбивалось на 8, 9, 10, 11 и 12 классов с помощью алгоритма «Объединение». При этом каждое из полученных разбиений задавалось в качестве начального разбиения для вариационного алгоритма, который и строил окончательное разбиение на данное число классов. Построенные классификации предъявлялись экспертам пользователя, которые выбирали наилучшую из пяти классификаций, т. е. окончательное число классов  $N$ , и уточняли составы классов. Элементы матрицы связей  $\phi_{ij}$  между классами  $q_1, \dots, q_n$  выбранной пользователем классификации определялись по формуле (2):  $\phi_{ij} = \frac{1}{m_i m_j} \sum_{s \in q_i} \sum_{p \in q_j} r_{sp}$ .

### ЗАКЛЮЧЕНИЕ

Предложенные методы выделения классов объектов профессиональной деятельности специалистов позволяют перейти к автоматизированным процедурам проектирования профессиональных и образовательных стандартов. Получение границ профессиональной деятельности специалиста повышает объективность результатов разработки профессиональных и образовательных стандартов.

### ЛИТЕРАТУРА

1. Никитин В.В. Информационно-методические обеспечение формирования перечня направлений и специальностей в области информационно-коммуникационных технологий. — М.: МАКС Пресс, 2006. — 272 с.
2. Браверман Э.М., Мучник И.Б. Структурные методы обработки эмпирических данных. — М.: Наука, 1983. — 302 с.
3. Бауман Е.В., Дорофеев А.А. Классификационный анализ данных // Труды Междунар. конф. по проблемам управления. — М., 1999. — Т. 1. — С. 62—77.

☎ (495) 771-32-38, (495) 334-90-70,  
e-mail: vnikitin@hse.ru, adorof@ipu.ru

Статья представлена к публикации членом редколлегии Ф.Т. Алескеровым. □

# ПОДХОД К ПРОГНОЗИРОВАНИЮ УСПЕШНОСТИ ИННОВАЦИОННОГО ПРОЕКТА

В.М. Винокур, Л.А. Мыльников, Н.В. Перминова

Пермский государственный технический университет

Рассмотрен гибридный подход к построению прогнозов инновационных проектов, находящихся на разных стадиях реализации, позволяющий повысить точность прогнозирования как на близкую, так и на долгосрочную перспективы при условии компетентного управления изучаемым инновационным проектом.

## ВВЕДЕНИЕ

Экономика развитых стран основывается на знаниях, научных достижениях и передовых технологиях. Традиционные источники экономического роста — новые сырьевые ресурсы, неосвоенные территории, дешевая рабочая сила и др. — практически исчерпаны. В условиях глобальной экономики капитал мобилен, технологии распространяются быстро, товары производятся в странах с низкими издержками и поставляются на рынки развитых стран. Основным источником конкурентных преимуществ и развития становятся инновации во всех сферах, что, в свою очередь, делает актуальной проблему принятия решений при управлении инновационными проектами. Однако, наблюдая за развивающимися в мировой экономике процессами, ученые и политики задаются вопросом, по какому пути пойдет развитие человечества в ближайшем будущем и в отдаленной перспективе. Вставая на инновационный путь развития, наиболее важным становится выбор ключевых направлений и проектов, которые смогут принести финансовую прибыль и дать толчок в развитии экономики.

Для решения этой задачи наиболее широко применяются методы технологического прогнозирования. Среди них можно выделить следующие методы: экспертные (метод мозгового штурма, метод Дельфи и др.), описательные (морфологический, аналогий, сценариев, дерева целей и др.), статистические и математического моделирования [1]. Однако достоверность прогноза полученного каждым из этих методов в отдельности, невысока, верификация полученных результатов затруднена, особенно при небольших периодах ретроспективы и сильно удаленном горизонте прогнозирования, что как раз и свойственно инновационным проектам. Известно, что наиболее объективные и точные

прогнозы могут быть получены лишь на качественной основе, но при изучении инновационных проектов такой основе просто неоткуда взяться.

Это связано с тем, что существующие модели инновационных проектов не позволяют рассматривать: все ресурсы (материальные, финансовые, кадровые, интеллектуальные, инфраструктурные) инноваций во взаимосвязи и взаимозависимости; разные способы продвижения инноваций — горизонтальный (способ партнерства и кооперации) и вертикальный (в рамках одной организации); инновации всех типов — создающие новые рынки, новые продукты, новые способы производства, структурные инновации; работу с большим объемом информации разной размерности. Существующие модели не пригодны для создания на их основах информационных систем и, следовательно, не позволяют контролировать и управлять инновационным проектом на любой стадии и работать с несколькими инновационными проектами одновременно [2].

Поэтому наиболее эффективный способ прогнозирования — применение гибридного подхода и оценка не отдельных показателей, а суммарной

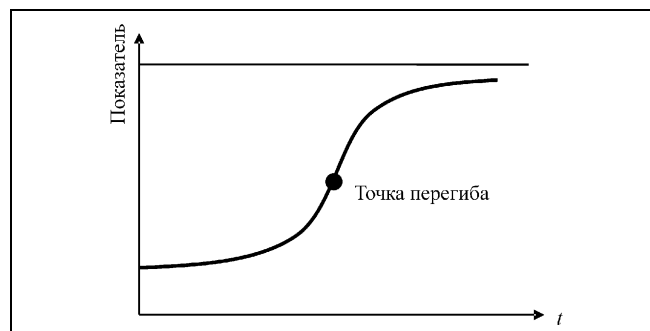


Рис. 1. Кривая смешанной динамики показателя развития инновационного проекта





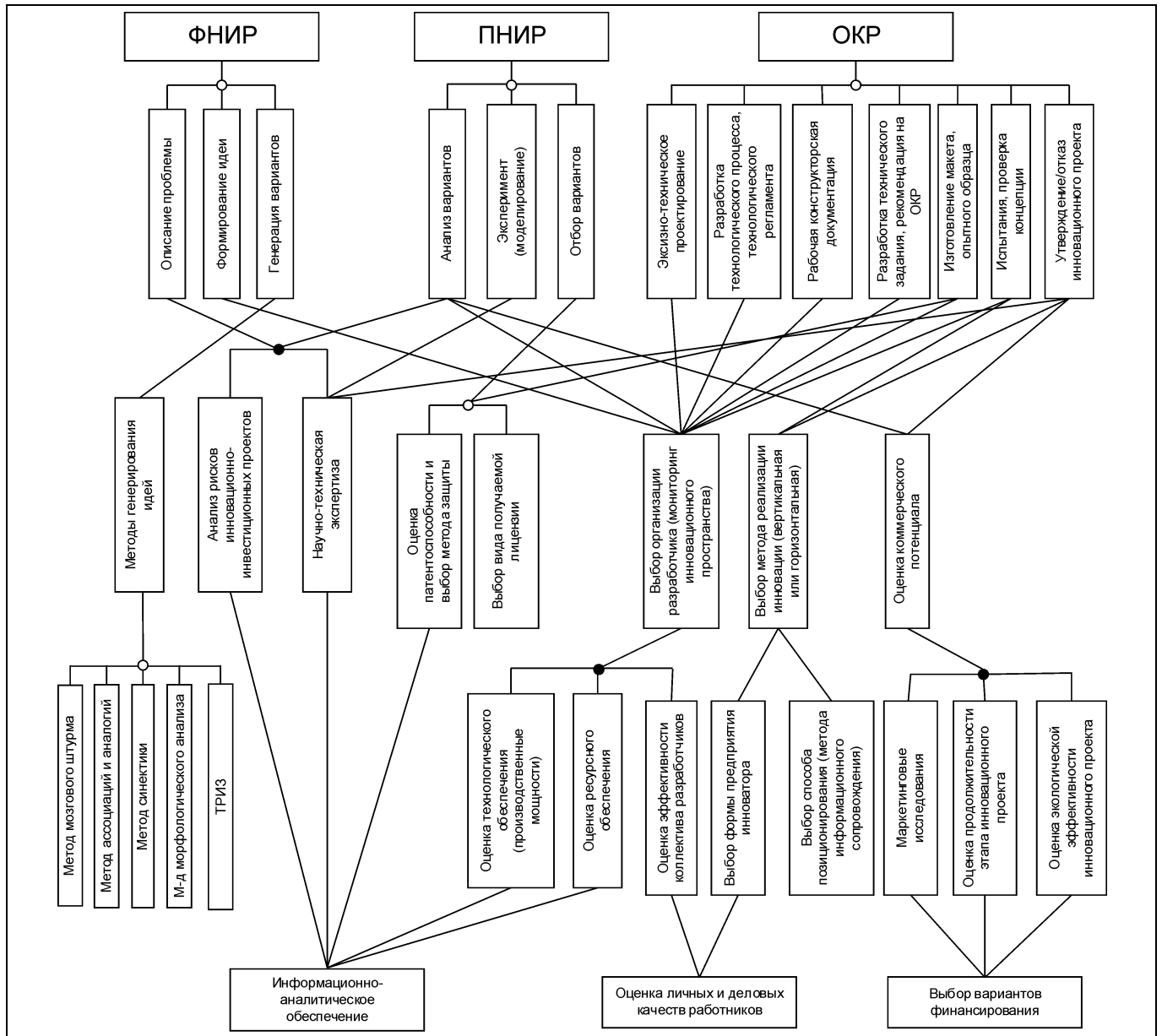
оценки. Известно, что отдельные показатели инновационного проекта описываются кривой, вид которой приведен на рис. 1.

### 1. МЕТОДИКА РЕШЕНИЯ

Можно предположить, что если отдельные показатели описываются такими кривыми, то и обобщенный критерий можно описать подобной кривой. Такая кривая может быть описана полиномами, экспоненциальной функцией, S-образной кривой (кривой Гомперца, логистической кривой) или гиперболической кривой. Тот или иной спо-

соб математического описания выбирается, исходя из специфики конкретного проекта (обычно известно, в каких областях и условиях какие математические описания ведут себя лучше).

Для построения обобщенного показателя инновационного проекта применим модифицированный метод морфологического синтеза. Для этапов инновационного проекта выберем конкретные факторы и методики, которые могут его адекватно описать, и, таким образом, получим дерево факторов и оценок, влияющих на осуществление конкретного инновационного проекта (рис. 2).



**Рис. 2. Пример морфологического дерева для стадии научной разработки инновационного проекта:**

ФНИР — фундаментальные научно-исследовательские работы; ПНИР — прикладные научно-исследовательские работы; ОКР — опытно-конструкторские работы; ТРИЗ — теория рационализаторства и изобретательства

На приведенном «И-ИЛИ» дереве черными кружками обозначены узлы «И», белыми — «ИЛИ». После создания морфологической таблицы для поиска оптимального решения требуется сформировать список требований, состоящий из обязательных (ограничений) и дополнительных (критериев). Следующий этап — свертывание показателей. Свертыванию подлежат только требования «И». Свертывать можно различными способами: суммированием (например, комплексная оценка — сумма всех показателей); определением минимакса (например, производительность вычисляется как минимум из максимумов производительности этапов); средневзвешенностью (обобщенный аддитивный критерий с ранжированными показателями); классификацией — присвоением 0 или 1 в зависимости от выполнения требования.

Осуществив операцию свертки, получим граф, описывающий исследуемый инновационный проект. Вершинами или узлами графа будут стадии, фазы или этапы инновационного проекта; ребрами — точки принятия решений или точки оценки набора показателей инновационного проекта. Направления ребер задают последовательность работы с моделью.

Выбор свертки показателей необходим, прежде всего, потому, что инновационный проект характеризуется всеми признаками системы, работающей в условиях неопределенности и недостатка информации, которые проявляются по-разному: при определении целей и подцелей, при появлении случайных факторов, при недостатке ресурсов. Источниками неопределенности в системе являются случайные факторы (они взаимосвязаны и оказывают косвенное воздействие на все ранее выявленные параметры системы) [3]. Основные из них:

- неопределенность взаимодействия элементов системы при отсутствии обратной связи от результатов к затратам ресурсов в течение продолжительного времени;
- неопределенность условий, в которых действует система, что проявляется в неопределенности описания факторов среды и процессов;
- непредсказуемость результатов, неоднозначность целей и ориентиров, что отличает инновационный проект от любой организационной системы, являющейся целеориентированной и ценностно-ориентированной;
- динамическая неустойчивость или нелинейная динамика и взаимодействие частей этапа инновационной разработки;
- наличие активного субъективного элемента системы — человека.

Особенность существующих показателей оценки инновационных проектов состоит в их слабой формализации. Принятие решений в системе разноразмерных оценок возможно путем приведения

их в единообразное представление в формате нечетких переменных с помощью функции Гаусса и задания нечетких множеств [4].

## 2. ТЕХНОЛОГИЯ РЕШЕНИЯ

В каждой из точек принятия решений после окончания каждого этапа или стадии инновационного проекта оценивается множество показателей. Пусть некоторое нечеткое значение  $x_i$  является результатом оценки или расчета по одной из методик. Это позволяет построить обобщенный показатель в виде:

$$M = A^T \begin{bmatrix} \sum_{i=1}^{n_1-1} x_{i1} c_{i1} \\ \vdots \\ \sum_{i=1}^{n_l-1} x_{il} c_{il} \end{bmatrix},$$

где  $A$  — матрица инцидентности (показывает последовательность применения методик и взаимосвязь стадий и этапов инновационного проекта, схематично представленных на рис. 3);  $c_i$  — весовые коэффициенты, показывающие значимость методики;  $n$  — число оценок на каждой стадии или этапе;  $l$  — число стадий и этапов (число связей между методиками, стадиями и этапами).

Сведем полученную многокритериальную задачу к однокритериальной методом линейной свертки, введя весовые коэффициенты  $k_j$ , которые являются экспертной оценкой значимости показателя в некоторой точке принятия решения:

$\sum_{j=1}^l k_j M_j \Rightarrow \text{extr}$ , где  $M_j$  — компонента вектора показателей стадии или этапа.

Таким образом, при управлении инновационным проектом осуществляется поиск оптимального решения из конечного числа альтернативных вариантов в каждой из точек на графе.

Часть переменных — показателей оценки, вычисляемых на первых стадиях инновационного проекта, являются входными для последующих стадий.

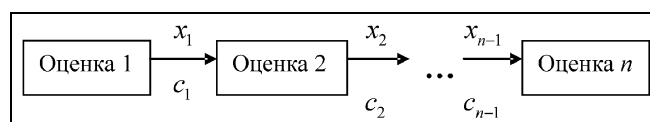


Рис. 3. Схема оценки обобщенного показателя в точке принятия решений

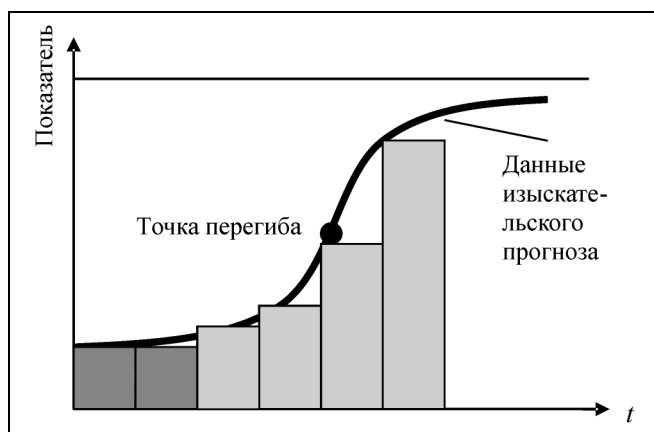


Рис. 4. Прогнозируемая кривая смешанной динамики описания обобщенного показателя развития инновационного проекта:

■ — практически данные, полученные во время реализации проекта; □ — данные нормативного прогноза

Задача принятия оптимального решения записывается в виде:

$$\sum_{j=1}^l (H_j - k_j M_j) \Rightarrow \min,$$

где  $H_j$  — желаемое значение.

Ограничения задачи зависят от типа управленческого решения. При поиске оптимальной последовательности стадий и этапов инновационного проекта при известных значениях показателей оценки в каждой точке принятия решений ограничение накладывается на взаимосвязи показателей (наличие и направление). В этом случае:

$$a_{ij} \in [1, 0, -1], \quad i, j = \overline{1, l},$$

где  $a_{ij}$  — элемент матрицы инцидентности.

При поиске оптимальных показателей при известной последовательности стадий проекта ограничение накладывается на альтернативные значения показателя в каждой точке принятия решений:

$$M_j \in D, \quad j = \overline{1, l},$$

где  $D$  — множество альтернативных значений показателя оценки.

Находя и применяя оптимальные решения для управления инновационным проектом на каждом шаге или последовательности шагов, можно построить нормативный прогноз. Причем исходными данными для расчета могут оказаться полученные реальные данные на предыдущих шагах. Используя полученные данные, можно применять методы математического моделирования и прогнозирования при расчете изыскательских прогнозов (рис. 4).

## ЗАКЛЮЧЕНИЕ

Применение описанного гибридного подхода позволяет осуществлять различные виды прогнозирования. Кроме этого, такой подход позволяет учитывать быстро протекающие изменения в изучаемой среде и мнения экспертов (которые заложены в применяемых методиках). Применение предложенного подхода может быть полезно для анализа группы проектов при формировании долгосрочных отраслевых и технологических прогнозов.

## ЛИТЕРАТУРА

1. *Инновационный менеджмент: Концепции, многоуровневые стратегии и механизмы инновационного развития* / под ред. В.М. Аньшина, А.А. Дагаева. — М.: Дело, 2006. — 584 с.
2. *Винокур В.М., Трусов А.В.* Интеллектуальная собственность как основа инновационной деятельности. — Пермь: ПГТУ, 2004. — 271 с.
3. *Радван Алькдиру, Мыльников Л.А.* Разработка модели инновационного проекта // Информационные управляющие системы: Сб. науч. тр. — Пермь, 2006. — С. 67–71.
4. *Мыльников Л.А., Перминова Н.В., Винокур В.М.* Информационная поддержка управлением инновационной деятельностью // Тр. 7-й междунар. науч.-практ. конф. «Современные информационные и электронные технологии». — Одесса, 2006. — С. 22.

e-mail: leonid@pstu.ru

Статья представлена к публикации членом редколлегии Р.М. Нижегородцевым. □

## Новая книга

Кочкаров Р.А. *Целевые программы: инструментальная поддержка.* — М.: Экономика, 2007. — 223 с.

Монография посвящена развитию методологической базы и формированию инструментальной поддержки программно-целевого планирования. Исследован жизненный цикл целевых программ, предложены подходы к их формализованному представлению. Детально описан процесс построения дерева целей, задач и мероприятий программ. Приведен пример использования компьютерного инструментария выработки управленческих решений по итогам мониторинга и контроля реализации целевых программ. Сформулированы и обоснованы предложения по расширению сферы применения методов программно-целевого планирования и повышению результативности целевых программ.

Для студентов и аспирантов экономических специальностей, а также специалистов, применяющих в своей работе программно-целевые методы.

# ЭКОНОМИЧЕСКАЯ ЭФФЕКТИВНОСТЬ УПРАВЛЯЮЩИХ СИСТЕМ<sup>1</sup>

Б.Г. Волик

Институт проблем управления им. В.А. Трапезникова РАН, г. Москва

Предложены критерии выбора предпочтительного варианта управляющей системы на основе анализа экономической эффективности конкурирующих вариантов системы.

## ВВЕДЕНИЕ

Работоспособность — важнейшее качество управляющих систем. На разных этапах жизненного цикла систем их работоспособность определяют разные свойства. В этом смысле жизненный цикл систем складывается из двух этапов: начального, до ввода в эксплуатацию, и эксплуатации, на заданном интервале времени или до предельного состояния.

На начальном этапе работоспособность систем определяется такими их свойствами, от которых зависит способность систем выполнять предписанные функции в идеальных условиях эксплуатации. Эту способность систем будем называть *функциональной работоспособностью*. Функциональная работоспособность управляющей системы определяется ее свойствами, отражающими точность и динамику управления (например, устойчивость, достижимость, параметрическая грубость, астатизм). На этапе эксплуатации решается задача сохранения функциональной работоспособности на уровне, достаточном для удовлетворения требований применения системы по её назначению. Способность систем удовлетворительно решать эту задачу будем называть *эксплуатационной работоспособностью*. Она определяется надёжностью, живучестью, техногенной безопасностью и эффективностью системы [1]. Эффективность занимает особое место среди названных свойств. Для многих классов объектов, например, технологических процессов, эффективность выступает как сводное свойство нескольких свойств, определяющих и функциональную, и эксплуатационную работоспособности.

<sup>1</sup> Работа выполнена при поддержке РФФИ, проект № 05-08-33535.

## 1. ОПРЕДЕЛЕНИЯ ЭФФЕКТИВНОСТИ

Эффективность характеризует полезность применения объекта по назначению. В зависимости от назначения объекта определение понятия *эффективность* трансформируется. Приведём ряд определений, отражающих характер эксплуатации управляющих систем.

*Эффективность* — свойство системы обеспечивать реализацию на заданном интервале времени эксплуатации заданного объёма своих проектных возможностей. Например, для вычислительной системы АСУ это отношение правильно решённых задач к числу предъявленных для решения задач на заданном интервале времени эксплуатации. Это определение отражает требования, предъявляемые, в частности, ко многим системам военного назначения.

*Эффективность* — свойство системы обеспечивать на заданном интервале времени эксплуатации целевую отдачу не ниже заданного уровня. Например, для управляющей системы энергоблоком атомной электростанции это предельный объём недовыдачи в течение года генерируемой электроэнергии из-за нарушений работоспособности управляющей системы.

*Эффективность* — свойство системы обеспечивать превышение целевой отдачи над затратами с некоторого заданного момента времени эксплуатации. Например, для АСУ технологическими процессами это может быть обеспечение превышения дохода от применения принятого варианта системы (объёма автоматизации) над затратами на обеспечение работоспособности системы с некоторого заданного момента времени.

## 2. ПОКАЗАТЕЛИ ЭФФЕКТИВНОСТИ

Основная особенность управляющих систем заключается в том, что результаты их применения (целевая отдача) выявляются только через измене-



ние объёма и (или) качества выходной продукции (выполнения предписанных функций) управляемого объекта. Затраты, связанные с созданием и эксплуатацией управляющих систем, зависят от режима эксплуатации и управляющей и управляемой систем.

Учитывая эти особенности эксплуатации, сформулируем требования к показателям эффективности управляющих систем.

Прежде всего, должны учитываться затраты и потери на всех этапах жизненного цикла [2] управляющей системы, начиная с этапа составления технического задания и заканчивая этапом демонстража и утилизации аппаратуры.

Далее, должна учитываться зависимость стоимости целевой отдачи и затрат от времени свершения, т. е. дисконтирование. Оно учитывает влияние экономического оборота средств на оценку их истинной стоимости и, вместе с тем, упрощает количественный анализ эффективности, так как позволяет приводить к единому моменту времени свершаемые в разные моменты времени компоненты целевой отдачи и затрат. Дисконтирование производится по формуле:

$$B(t_0) = B(1 + E)^{t_0 - t}, \quad (1)$$

где  $B$  — отдача (затраты), совершенные в момент времени  $t$ ;  $B(t_0)$  — отдача (затраты), приведенные к моменту времени  $t_0$ ;  $E$  — норма дисконтирования (обычно в пределах 0,05—0,1).

Принимая за момент времени приведения целевой отдачи и затрат начало эксплуатации системы  $t_0 = 0$ , получаем удобную аппроксимацию формулы (1):

$$B(0) = Be^{-Et}.$$

Наконец, показатель должен отражать существо задачи, решаемой разработчиками управляющей системы.

Типичными можно считать следующие три постановки задачи.

*Первая постановка задачи:* необходимо оценить экономическую целесообразность замены исходной (базовой) системы некоторой технически реализуемой  $i$ -й системой.

Эту задачу будем называть задачей оценки сравнительной эффективности.

В соответствии с основополагающим принципом оценки эффективности «затраты—результаты» показателем эффективности здесь следует принять экономический эффект (прибыль) от замены базового варианта системы на  $i$ -й вариант. Итак,  $\mathcal{E}_{6i}(t) = Q_{6i}(t) - Z_i(t)$ , где  $Q_{6i}$  — целевая отдача на интервале времени  $(0, t)$ , приведенная к моменту  $t = 0$ , от замены базового варианта системы на  $i$ -й;  $Z_i$  — приведенные затраты и потери, связанные с переходом на  $i$ -й вариант системы.

*Вторая постановка задачи:* необходимо оценить допустимость вложения средств в замену системы (или модернизацию) при ограничении вложения средств уровнем нормативной эффективности, принятом на фирме-собственнике объекта.

Эта задача названа задачей оценки абсолютной экономической эффективности. Она решается путём сравнения коэффициента рентабельности  $E_{6i}$  вложений в  $i$ -й вариант с нормативным коэффициентом рентабельности  $E_n$ . Итак:  $E_{6i}(t) = Q_{6i}(t)/Z_i(t)$ .

*Третья постановка задачи:* необходимо выбрать лучший вариант системы из числа технически реализуемых, одновременно удовлетворяющий ограничению по рентабельности вложения средств. Для такой постановки задачи необходима оценка по показателям задач первых двух постановок.

*Замечание.* Для некоторых классов управляемых объектов (систем), например военного назначения, целевую отдачу оказывается возможным выразить только в натуральных единицах измерения через оценку приращения выходных эффектов объекта. Однако затраты на замену (модернизацию) управляющей системы, как правило, можно выразить в денежных единицах. В таких случаях оценивать варианты систем можно посредством показателей технической эффективности, определяемых как отношение целевой отдачи (в натуральных единицах) к общим затратам (в денежных единицах).

### 3. СХЕМА АНАЛИЗА ЭФФЕКТИВНОСТИ

Основная проблема анализа эффективности заключается в получении количественных оценок целевой отдачи  $Q_{6i}$  и суммарных затрат  $Z_i$ , приведенных к моменту ввода системы в эксплуатацию. Трудности получения оценок связаны с распределённостью во времени и различием физической природы факторов, влияющих на оцениваемые характеристики. Удобный инструмент анализа — модель жизненного цикла системы [2]. Она отражает полную совокупность положительных и отрицательных эффектов, сопровождающих все этапы создания и эксплуатации анализируемого объекта. Построение такой модели следует начинать с выявления источников целевой отдачи и видов затрат и потерь, зависящих от состояния управляющей системы. Для примера назовём типичные источники целевой отдачи управляющих систем промышленными объектами:

- увеличение объёма предотвращённого ущерба, возможного при возникновении аварийных ситуаций на управляемом объекте;
- увеличение объёма выпуска продукции;
- повышение качества и, соответственно, цены продукции;
- сокращение объёма брака;

- сокращение численности оперативного персонала;
- повышение техногенной безопасности [3];
- сокращение потребления энергетических и сырьевых ресурсов;
- сокращение длительности вынужденных простоев оборудования.

Основные виды затрат и потерь:

- капитальные затраты на проектирование, изготовление и наладку системы до момента ввода её в эксплуатацию;
- недовыдача продукции на интервалах времени восстановления работоспособности системы после отказов;
- потери продукции (брак), обусловленные случайными нештатными остановками производства при отказах управляющей системы;
- текущие эксплуатационные затраты, которые складываются из условно-постоянных (зарплата персонала, платы за помещение, потребляемую энергию, каналы связи и др.) и условно-переменных (на ремонтные работы, запасные изделия и приборы и другие затраты, зависящие от режима функционирования системы).

Обратимся к количественным моделям оценки экономической эффективности управляющих систем.

Соотношения, привлекаемые для оценки целевой отдачи, определяются режимом выпуска выходной продукции (выходом) управляемым объектом.

*Выход объекта  $Q_i(t)$  дискретный (интервал времени производства продукции значительно меньше интервала ожидания).*

Тогда

$$Q_i(t) = \sum_{j=1}^N q_i(t_j) K_{ri}(t_j) e^{-Et_j}, \quad (2)$$

где  $t$  — интервал времени  $(0, t)$  оценки эффективности;  $N$  — число съёмов выходной продукции на интервале  $(0, t)$ ;  $t_j$  — моменты съёма продукции;  $K_{ri}(t_j)$  — значение коэффициента готовности управляющей системы в момент времени  $t_j$ ;  $q_i(t_j)$  — объём (в денежных единицах измерения) выходной продукции в момент времени  $t_j$ ;  $e^{-Et_j}$  — коэффициент приведения стоимости продукции полученной в момент  $t_j$  к моменту  $t = 0$ .

Выход объекта непрерывный:

$$Q_i(t) = \int_0^t q_i(\tau) e^{-E\tau} d\tau - \omega_i q_{i0} t e^{-Et/2}, \quad (3)$$

где  $q_i(\tau)$  — интенсивность прибыли (в денежных единицах в единицу времени) управляемого объекта с  $i$ -й управляющей системой;  $q_{i0}$  — среднее значение интенсивности прибыли на интервале  $(0, t)$ ;  $\omega_i$  — параметр потока отказов  $i$ -й управляющей

системы на стационарном этапе эксплуатации;  $t_b$  — среднее значение времени восстановления работоспособности управляющей системы после отказов;  $\tau$  — текущее время.

Для непрерывных производств интенсивность прибыли определяется как

$$q_i(\tau) = (u_i(\tau) - c_i(\tau)) P_p(\tau),$$

где  $u$  и  $c$  — цена и себестоимость выходной продукции объекта,  $P_p$  — производительность объекта.

В формуле (3) второй член зависимости суть среднее значение объёма недовыдачи продукции управляемым объектом на интервалах восстановления работоспособности управляющей системы на интервале времени эксплуатации  $(0, t)$ .

Аналогично, как и по формулам (2) и (3) вычисляется выход объекта  $Q_6(t)$  с исходным (базовым) вариантом управляющей системы.

Целевая отдача от замены базового варианта управляющей системы на  $i$ -й вариант составит величину:  $Q_{6i}(t) = Q_i(t) - Q_6(t)$ .

Приведем соотношения, необходимые для количественной оценки приведенных суммарных затрат на создание и эксплуатацию управляющей системы, а также потерь, обусловленных возможными её отказами.

*Капитальные затраты на проектирование и создание  $i$ -й управляющей системы, приведенные к моменту времени  $\tau = 0$  ввода системы в эксплуатацию:*

$$K_i = \sum_{r=1}^R K_{ir} e^{-Et_r} + \int_{t_0}^0 k_i(\tau) e^{-E\tau} d\tau, \quad t_r, t_0 < 0,$$

где  $K_{ri}$  — разовые вложения (например, покупки оборудования) в создание системы, производимые в моменты времени  $t_r$ ;  $R$  — общее число разовых вложений;  $k_i(\tau)$  — непрерывные затраты на этапе  $(t_0, 0)$  проектирования и создания управляющей системы.

Эксплуатационные затраты  $i$ -й управляющей системы:

$$Z_{3i}(t) = \int_0^t z_{poi}(\tau) e^{-E\tau} d\tau + \int_0^t z_{pei}(\tau) e^{-E\tau} d\tau + \omega_i t n_i e^{-Et/2} \quad (4)$$

где  $z_{poi}$  — интенсивность условно-постоянных затрат, связанных с эксплуатацией  $i$ -й управляющей системы;  $z_{pei}$  — интенсивность усреднённых на интервале  $(0, t)$  условно-переменных затрат;  $n_i$  — усреднённые потери (брак) продукции управляемого объекта, штраф за возможное нарушение экологических ограничений, затраты на восстановление работоспособности системы и объекта, связанные с одним событием отказа управляющей системы;  $e^{-E\tau}$  — коэффициент приведения затрат к моменту времени  $\tau = 0$ .



В формулах (3) и (4) невыдача продукции и потери (брак) приведены с момента времени  $t/2$  для упрощения расчётов.

Полные приведенные затраты и потери, связанные с  $i$ -м вариантом управляющей системы, составляют величину:

$$Z_i(t) = K_i + Z_{\text{э}i}(t).$$

#### 4. КРИТЕРИИ ПРИНЯТИЯ РЕШЕНИЙ

Для первой постановки задачи — выбора предпочтительного варианта управляющей системы.

Критерий выбора:  $i$ -й вариант системы предпочтительней базового (исходного) варианта при условии  $\mathcal{E}_{\text{б}i}(t) = Q_{\text{б}i}(t) - Z_i(t) > 0$ , где  $t$  — заданная продолжительность эксплуатации,  $\mathcal{E}_{\text{б}i}$  — экономический эффект перехода на  $i$ -й вариант системы.

Для объектов с изменяющимися во времени параметрами, например, параметром потока отказов, ценой продукции, затратами и др., в качестве показателя экономического эффекта следует применять показатель «время окупаемости  $T_{\text{ок}}$ ». Для таких объектов критерий выбора  $T_{\text{ок}} < t$ , где  $T_{\text{ок}} \subset \{\tau: \mathcal{E}_{\text{б}i}(\tau) = 0\}$ .

Для второй постановки задачи.

Критерий принятия решения: намечаемый к внедрению  $i$ -й вариант системы экономически допустим при условии:  $E_{\text{б}i}(T_{\text{н}}) = Q_{\text{б}i}(T_{\text{н}})/Z_i(T_{\text{н}}) > 1$ ,  $T_{\text{н}} = 1/E_{\text{н}}$  — нормативное время достижения рентабельной эксплуатации, где  $E_{\text{н}}$  — коэффициент нормативной рентабельности вложения средств в создание или модернизацию объектов, принятый в фирме-собственнике анализируемого объекта.

Для объектов с изменяющимися во времени параметрами следует применять временной критерий:  $T_{\text{pi}} < T_{\text{н}}$ , где  $T_{\text{pi}} = \{\tau: E_{\text{б}i}(\tau) = 1\}$  — момент времени достижения рентабельной эксплуатации с  $i$ -й управляющей системой.

Для третьей постановки задачи.

Критерий принятия решения: намечаемый к реализации  $i$ -й вариант системы приносит экономическую выгоду и допустим, по ограничению фирмы на вложения средств в новые разработки, при условиях:

$$T_{\text{pi}} = T_{\text{оки}} < T_{\text{н}} \text{ и } T_{\text{pi}} = T_{\text{оки}} < t.$$

#### ЗАКЛЮЧЕНИЕ

В статье приведены упрощенные зависимости целевой отдачи и эксплуатационных затрат. Упрощение заключается в том, что принято среднее значение параметра потока отказов. Для получения более точных оценок эффективности следует потери и затраты в формулах (3) и (4) распределить по видам отказов (простым, аварийным и

цепным [4]), так как значения могут существенно различаться.

Зависимости целевой отдачи (3), потерь и затрат (4) показывают пути влияния на уровень эффективности эксплуатации систем. Так, факторы, определяющие функциональную работоспособность систем, влияют на цену и себестоимость продукции, а также производительность управляемого объекта, а факторы, определяющие эксплуатационную работоспособность систем, влияют на потери и затраты на восстановление работоспособности, и управляющей системы, и управляемого объекта, иными словами, на размер предотвращённого ущерба. Для многих типов управляемых объектов, особенно технологических процессов, повышение надёжности и совершенствование подсистемы противоаварийной защиты управляющей системы служат основными рычагами повышения эффективности комплекса «объект—управляющая система».

Основная трудность анализа эффективности систем сложной структуры в получении оценок коэффициента оперативной готовности  $K_r$  и параметра потока отказов  $\omega$  анализируемых вариантов системы. Описание методов анализа для систем сложной структуры можно найти, например, в работах [4, 5]. Для практических расчётов следует применять современные программные комплексы [6].

Для восстанавливаемых систем с неограниченным временем эксплуатации мы неизбежно сталкиваемся со старением аппаратуры систем и, соответственно, с ростом интенсивности их отказов. Таким образом, возникает необходимость исследования динамики эффективности [2]. Такое исследование позволяет дополнительно получить оценки оптимального по эффективности времени эксплуатации системы и времени её полной экономической деградации.

#### ЛИТЕРАТУРА

1. Волик Б.Г. Эксплуатационная работоспособность систем // Третья международная конференция по проблемам управления (20—22 июня 2006 г.): Пленарные доклады и избранные труды. — М.: Институт проблем управления, 2006.
2. Волик Б.Г. Анализ влияния надёжности на экономическую эффективность АСУТП // Приборы и системы управления. — 1976. — № 4.
3. Волик Б.Г. Проблемы анализа техногенной безопасности // Автоматика и телемеханика. — 2002. — № 12.
4. Методы анализа структур управляющих систем / Б.Г. Волик и др. — М.: Энергоатомиздат, 1988.
5. Дж. Хенли Э., Куамато Х. Надёжность технических систем и оценка риска. — М.: Машиностроение, 1984.
6. Relex — программа анализа надёжности, безопасности, рисков / В.С. Викторова и др. // Надёжность. — 2003. — № 4.

☎ (495) 334-79-61, e-mail: lfvk@ipu.ru

Статья представлена к публикации членом редколлегии А.С. Манделем. □

# МАТЕМАТИЧЕСКАЯ ФОРМАЛИЗАЦИЯ МЕТОДОВ ИЕРАРХИЧЕСКОГО УПРАВЛЕНИЯ ЭКОЛОГО-ЭКОНОМИЧЕСКИМИ СИСТЕМАМИ

Г.А. Угольницкий, А.Б. Усов

*Южный федеральный университет, г. Ростов-на-Дону*

Сформулированы общие подходы к управлению эколого-экономическими системами. Предложены методы иерархического управления (принуждение, побуждение, убеждение), позволяющие субъекту управления верхнего уровня добиться выполнения требований устойчивого развития системы. Приведены примеры применения указанных методов в системе контроля качества речных вод и дан их сравнительный анализ.

## ВВЕДЕНИЕ

В последние десятилетия ускоренными темпами происходят концентрация производства, организация крупных промышленных объединений, глобализация экономики. Это приводит к усилению воздействия техногенных процессов на окружающую природу, создает реальную опасность истощения как невозобновляемых, так и возобновляемых ресурсов.

Рациональное использование природных ресурсов, внедрение новых ресурсосберегающих технологий требуют перехода на принципиально иные эколого-экономические отношения, в основе которых лежат сопряженность, соразмерность, сбалансированность природных и производственных потенциалов, эколого-экономический баланс. Окружающая среда должна быть включена в систему социально-экономических отношений как один из ее важнейших элементов.

Решение этих задач невозможно без комплексного подхода, предполагающего создание концепции управления эколого-экономическими системами, в которой должны быть сформулированы главные задачи управления, общая структура систем управления и принципы их организации. Актуальной проблемой является разработка новых подходов к управлению, которые исходят из требования устойчивого развития эколого-экономических систем и предполагают применение методов иерархического управления.

## 1. ИЕРАРХИЧЕСКИ УПРАВЛЯЕМАЯ ДИНАМИЧЕСКАЯ СИСТЕМА

Устойчивое развитие определяется как процесс, отвечающий потребностям настоящего, но не лишаящий будущие поколения возможности удовлетворять свои потребности. В основе стратегии устойчивого развития лежит идея равновесия между окружающей средой, экономикой и населением Земли [1–3].

При анализе управления устойчивым развитием целесообразно трактовать эколого-экономическую систему как иерархически управляемую динамическую систему [4]. Простейшая иерархически управляемая динамическая система имеет два уровня и включает в себя:

- источник воздействия верхнего уровня (Ведущий);
- источник воздействия нижнего уровня (Ведомый);
- собственно управляемую динамическую систему (УДС).

Взаимоотношения внутри такой иерархической системы устроены следующим образом: Ведущий воздействует на Ведомого и на УДС, Ведомый — только на УДС. Ведущего и Ведомого вместе можно рассматривать как совокупный источник воздействия на УДС, имеющий иерархическую структуру. В качестве Ведущего чаще всего выступают органы государственной власти, в роли ведомых — промышленные предприятия, а в качестве УДС может выступать как целиком вся экологическая





система некоторого региона, так и отдельные ее составляющие.

Предприятия забирают из окружающей среды природные ресурсы, перерабатывая которые они производят необходимый обществу конечный продукт. В окружающую среду при этом попадают продукты технологического процесса — различного вида отходы.

Основной смысл введения понятия *иерархической управляемой динамической системы* заключается в следующем. Воздействуя на эколого-экономическую систему, предприятие преследует собственные цели, в общем случае не отвечающие требованиям устойчивого развития. Сама эколого-экономическая система, будучи неодушевленным объектом, не может целенаправленно «отстаивать свои интересы»; ее ответные реакции носят спонтанный, запаздывающий характер. Поэтому нужен Ведущий, способный, используя различные механизмы управления, воздействовать на Ведомого для достижения целей устойчивого развития всей системы.

В России в настоящее время формируется экономический механизм управления природопользованием, включающий в себя как поощрительные, так и принудительные элементы регулирования. Его основу составляют платежи за пользование природными ресурсами и экономические санкции за экологические правонарушения. Одна из основных научных задач при этом заключается в разработке методов регулирования (административных, экономических), в наибольшей степени отвечающих целям концепции устойчивого развития.

Рассмотрим математическую формализацию основных методов иерархического управления на примере модели контроля качества речных вод.

## 2. МАТЕМАТИЧЕСКАЯ МОДЕЛЬ КОНТРОЛЯ КАЧЕСТВА РЕЧНЫХ ВОД

Для простоты в предлагаемой далее модели не учитываются платежи за водозабор и водосброс, в рассмотрение принимаются только платежи за сброс загрязнений. Исследуется случай только одного вида загрязнений, например, азотосодержащих загрязняющих веществ.

Пусть вдоль реки расположено  $N$  предприятий, которые сбрасывают загрязняющие вещества в реку вместе со сточными водами. Деятельность предприятий контролируется правительственным органом (центром). Центр, регулирующий качество речной и сточных вод, взимает с предприятий плату за сброс загрязнений. Основная цель центра состоит в поддержании речной системы в устойчивом состоянии. Добиться этого центр может не единственным образом, поэтому помимо основной цели он стремится к максимизации остающихся в его распоряжении средств, поступающих от предприятий.

Целевая функция центра имеет вид

$$J_0 = \int_0^{\Delta} \left\{ -C_A(y(t)) + \sum_{i=1}^N [v_i(R_i(\Phi_i) - VK_i(\Phi_i) - H_i(\Phi_i) - VS_i(W_i, P_i) - FN_i) + FN_i + FP_i + FS_i] \right\} dt \rightarrow \max \{ KN_i(t), KS_i(t), q_i(t) \}_{i=1}^N;$$

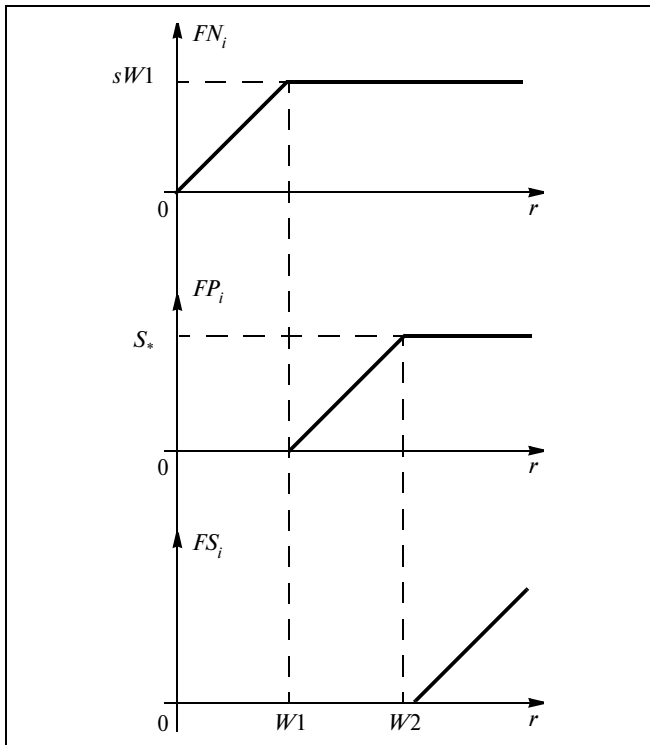
$$y(t) = \sum_{i=1}^N [1 - P_i(t)] W_i(t);$$

$$FN_i = \begin{cases} s(1 - P_i(t)) W_i(t), & \text{если } (1 - P_i(t)) W_i(t) \leq W1, \\ sW1, & \text{если } (1 - P_i(t)) W_i(t) > W1, \end{cases}$$

$$FP_i = \begin{cases} 0, & \text{если } (1 - P_i(t)) W_i(t) \leq W1, \\ sKN_i(t)((1 - P_i(t)) W_i(t) - W1), & \text{если } W1 < (1 - P_i(t)) W_i(t) \leq W2, \\ sKN_i(t)(W2 - W1), & \text{если } W2 < (1 - P_i(t)) W_i(t); \end{cases}$$

$$FS_i = \begin{cases} 0, & \text{если } (1 - P_i(t)) W_i(t) \leq W2, \\ sKS_i(t)((1 - P_i(t)) W_i(t) - W2), & \text{если } (1 - P_i(t)) W_i(t) > W2. \end{cases}$$

Здесь  $t$  — временная координата;  $W_i(t)$  и  $(1 - P_i(t)) W_i(t)$  — количество загрязнений, сбрасываемых в реку до и после очистки сточных вод в момент времени  $t$ ;  $P_i(t)$  — доля загрязнений, удаляемых на  $i$ -м предприятии в процессе очистки сточных вод;  $FS_i$ ,  $FN_i$  и  $FP_i$  — функции платежей (см. рисунок) предприятий за сброс загрязнений в пределах установленного норматива, а также за сверхнормативный и сверхлимитный сбросы соответственно, эти функции зависят от общего количества загрязнений, сбрасываемых в реку после очистки сточных вод, т. е. от величины  $r(1 - P_i(t)) W_i(t)$ ;  $s$ ,  $KN_i$  и  $KS_i$  — размеры платы за единицу сброшенных загрязнений на  $i$ -м предприятии в момент времени  $t$  при сбросе в установленных пределах, сверхнормативном и сверхлимитном сбросах соответственно;  $W1$  и  $W2$  — установленные законодательством нормативы сброса загрязнений;  $C_A(y(t))$  — функция, отражающая материальные потери общества из-за загрязненной воды (затраты на устройство новых мест отдыха в других регионах, дополнительные расходы по очистке речной воды для потребительских нужд населения и т. п.);  $y(t)$  — общее количество сброшенных в реку загрязнений;  $\Delta$  — момент времени, до которого ведется рассмотрение;  $v_i$  — ставка налога на прибыль на  $i$ -м предприятии;  $R_i(\Phi_i)$  — доход  $i$ -го предприятия от реализации произведен-


 Графики функций платежей (штрафов);  $S_* = sKN_i(W2 - W1)$ 

ной продукции при размере производственных фондов  $\Phi_i$ ;  $VK_i(\Phi_i)$  — включаемые в себестоимость издержки основного производства;  $H_i(\Phi_i)$  — суммарная заработная плата основного и природоохранного производств;  $VS_i(W_i, P_i)$  — издержки природоохранной деятельности, зависящие от объема сбрасываемых загрязнений и степени очистки сточных вод;  $q_i$  — минимально допустимая степень очистки сточных вод на  $i$ -м предприятии.

Функции  $VK_i$  и  $H_i$  зависят от объема производства, т. е. от размера производственных фондов. Функция  $VS_i(W_i, P_i)$  отражает затраты  $i$ -го предприятия на очистку сточных вод. Предполагается, что выполнены следующие соотношения:

$$\begin{aligned} VK_i(\Phi_i) + H_i(\Phi_i) &= \mu_i R_i(\Phi_i); \quad \mu_i = \text{const}; \\ VS_i(W_i, P_i) &= W_i C_p(P_i), \end{aligned}$$

где  $C_p(P_i)$  — функции затрат  $i$ -го предприятия на очистку единицы сбрасываемых в водоток загрязнений.

Цель предприятий состоит в максимизации прибыли, полученной в ходе производственной деятельности, т. е.

$$J_i = \int_0^{\Delta} \{(1 - v_i)(R_i(\Phi_i) - VK_i(\Phi_i) - H_i(\Phi_i) - VS_i(W_i, P_i) - FN_i) - FP_i - FS_i\} dt \rightarrow \max\{P_i\}. \quad (2)$$

Динамика производственных фондов предприятия описывается уравнением

$$\frac{d\Phi_i}{dt} = -k_i \Phi_i + Y_i, \quad i = 1, 2, \dots, N, \quad (3)$$

где  $k_i$  — коэффициент амортизации производственных фондов,  $Y_i$  — инвестиции в производство.

Пусть количество сбрасываемых загрязнений (до очистки) зависит от количества произведенной продукции линейно, а производственные функции имеют вид

$$\begin{aligned} W_i &= \beta_i R_i(\Phi_i), \quad R_i(\Phi_i) = \gamma_i \Phi_i^{0,5}, \\ i &= 1, 2, \dots, N, \quad \gamma_i, \beta_i = \text{const}. \end{aligned} \quad (4)$$

Основные характеристики качества речной воды — концентрация загрязняющих веществ  $B(x, t)$  и концентрация растворенного в воде кислорода  $B^{O_2}(x, t)$  описываются в случае пространственной неоднородности только вдоль русла реки уравнениями:

$$\begin{aligned} \frac{\partial B(x, t)}{\partial t} + v_x \frac{\partial B(x, t)}{\partial x} &= \frac{1}{A} \frac{\partial}{\partial x} \left[ EA \frac{\partial B(x, t)}{\partial x} \right] - \\ &k_c B(x, t) + W^0(x, t)(1 - P^0(x, t))/A; \end{aligned}$$

$$\begin{aligned} \frac{\partial B^{O_2}(x, t)}{\partial t} + v_x \frac{\partial B^{O_2}(x, t)}{\partial x} &= \frac{1}{A} \frac{\partial}{\partial t} \left[ EA \frac{\partial B^{O_2}(x, t)}{\partial x} \right] - \\ &- k_c B(x, t) + [B_{\text{нас}}^{O_2} - B^{O_2}(x, t)] + F_0 - F_1 - F_2, \end{aligned} \quad (6)$$

где  $x$  — пространственная координата;  $0 \leq x \leq L$ ;  $L$  — длина реки;  $E$  — коэффициент дисперсии;  $A$  — площадь поперечного сечения реки;  $v_x$  — скорость воды в реке;  $k_c B(x, t)$  — изменение во времени биохимического потребления кислорода из-за распада;  $K_{O_2}(B_{\text{нас}}^{O_2} - B^{O_2}(x, t))$  — добавка растворенного кислорода вследствие реэрации;  $B_{\text{нас}}^{O_2}$  — концентрация насыщения кислорода;  $F_0$  — добавка кислорода вследствие фотосинтеза;  $F_1$  — потребление растворенного кислорода на дыхание;  $F_2$  — придонное потребление растворенного кислорода.

Функции  $W^0$  и  $P^0$  отражают наличие источников загрязнений и определяются формулами

$$P^0(x, t) = \begin{cases} 0, & \text{если } x \neq x_i, \\ P_i(t) & \text{в противном случае;} \end{cases}$$

$$W^0(x, t) = \begin{cases} 0, & \text{если } x \neq x_i, \\ W_i(t) & \text{в противном случае.} \end{cases}$$

В точках  $x_i$ ,  $i = 1, 2, \dots, N$ , расположены предприятия.



Оптимизационные задачи (1) и (2) решаются при следующих ограничениях на управления

$$q_i \leq P_i(t) \leq 1 - \varepsilon, \quad 0 \leq t \leq \Delta, \quad (7)$$

$$0 \leq KN_i(t) \leq KN_{\max}, \quad 0 \leq KS_i(t) \leq KS_{\max}, \\ 0 \leq t \leq \Delta, \quad (8)$$

где значения величин  $KS_{\max}$  и  $KN_{\max}$  заданы;  $0 < \varepsilon < 1$  — постоянная, определяемая технологическими возможностями очистки сточных вод на предприятиях.

Требования устойчивого развития эколого-экономической системы, включающей в себя водоток, состоят в необходимости выполнения следующих неравенств:

$$0 \leq B(x, t) \leq B_{\max}, \quad B_{\min}^{O_2} \leq B^{O_2}(x, t), \\ 0 \leq t \leq \Delta, \quad 0 \leq x \leq L, \quad (9)$$

$$\frac{W_i(t)[1 - P_i(t)]}{Q_i^0(t)} \leq Q_{\max}, \quad 0 \leq t \leq \Delta, \\ i = 1, 2, \dots, N, \quad (10)$$

где  $Q_i^0(t)$  — расход воды на  $i$ -м предприятии; величины  $B_{\max}$ ,  $B_{\min}^{O_2}$  и  $Q_{\max}$  заданы.

Условия (9) связаны с государственными стандартами качества речной воды, а условия (10) — со стандартами поступления загрязнений с очищенной сточной водой для каждой точки сброса [5].

### 3. МЕТОДЫ ИЕРАРХИЧЕСКОГО УПРАВЛЕНИЯ

Устойчивое развитие эколого-экономической системы может достигаться различными методами иерархического управления. В зависимости от степени свободы индивида как объекта управления можно выделить следующие методы управления: принуждение, побуждение и убеждение.

*Метод принуждения* предполагает воздействие Ведущего на множество допустимых управленческих стратегий Ведомого. Это воздействие носит административно-законодательный характер, т. е. Ведущий запрещает Ведомому применять стратегии, не удовлетворяющие требованиям устойчивого развития. Из множества возможных стратегий принуждения Ведущий выбирает ту, которая отвечает его критерию оптимальности.

*Метод побуждения* заключается в стимулировании Ведомого путем воздействия Ведущего на его целевую функцию. Смысл метода заключается в том, чтобы сделать управления устойчивого развития экономически наиболее выгодными для Ведомого. Для этого применяется экономический механизм с обратной связью, предусматривающий поощрение Ведомого в случае выполнения им условий устойчивого развития и наказание в

противном случае. В отличие от метода принуждения, Ведомому явно не запрещается выбирать стратегии, нарушающие условия устойчивого развития. Поэтому, если возможностей Ведущего недостаточно для стимулирования Ведомого, то стратегия побуждения может оказаться неэффективной по отношению к «браконьерскому» поведению Ведомого.

*Метод убеждения* подразумевает добровольную кооперацию Ведущего и Ведомого для совместного обеспечения условий устойчивого развития и максимизации суммарной целевой функции, с последующим дележом полученной коалицией Ведущего и Ведомого суммарной прибыли между ними. Суть метода убеждения заключается в преобразовании иерархических отношений в кооперативные.

Подробная характеристика методов принуждения, побуждения и убеждения дана в работе [6]. Формализуем указанные методы управления для случаев, описываемых выражениями (1)–(10).

## 4. ФОРМАЛИЗАЦИЯ МЕТОДОВ ИЕРАРХИЧЕСКОГО УПРАВЛЕНИЯ

Основная цель центра состоит в выполнении неравенств (9) и (10), что позволяет поддерживать систему в устойчивом состоянии. Гарантировать выполнение этих условий центр может не единственным образом, поэтому, помимо основной цели, он стремится к оптимизации функции (1). При этом используются следующие методы управления.

### 4.1. Принуждение

В случае принуждения центр воздействует на область допустимых управлений предприятий и заставляет их выполнять стандарты качества речной и сточных вод, не оставляя им других возможностей, путем выбора значений  $q_i$  — минимально допустимых степеней очистки сточных вод.

Алгоритм построения равновесия принуждения состоит в следующем.

1. В результате максимизации критериев (2) при ограничениях (7) и соотношениях (3) и (4) определяются оптимальные стратегии предприятий в зависимости от управлений центра  $(P_i)^*(KN_i, KS_i, q_i)$   $i = 1, 2, \dots, N$ .

2. Метод принуждения предполагает, что размеры платы за единицы сброшенных загрязнений  $(KN_i, KS_i)$  при решении оптимизационных задач остаются неизменными и функция (1) максимизируется только по значениям  $q_i$ ,  $i = 1, 2, \dots, N$ . Оптимальными для центра являются значения  $(q_i)^*$ , которые доставляют максимум его целевой функции и позволяют выполнить стандарты качества реч-

ной и сточных вод. Таким образом, решается задача (1), (8) с учетом требований (9) и (10).

3. Равновесие принуждения имеет вид  $\{(q_i)^*, (P_i)^*\}_{i=1}^N$ , где  $(P_i)^*$  — оптимальные реакции предприятий на выбранную центром стратегию:  $(P_i)^* = (P_i)^*(q_i)^*, KN_i, KS_i$

#### 4.2. Побуждение

Метод побуждения предполагает, что центр воздействует на целевые функции предприятий. Он назначает размер платы за единицы сброшенных загрязняющих веществ (значения  $q_i, i = 1, 2, \dots, N$  остаются постоянными) и создает условия, при которых предприятиям выгодно придерживаться стандартов качества вод.

Алгоритм построения равновесия побуждения состоит в следующем.

1. Аналогично методу принуждения определяются оптимальные стратегии предприятий в зависимости от стратегии центра.

2. В методе побуждения критерий (1) центра, в котором значения  $P_i$  определяются на первом шаге алгоритма, максимизируется по  $KN_i, KS_i, i = 1, 2, \dots, N$  (значения  $q_i(t)$  не меняются) при условиях (8)–(10).

Оптимальными для центра являются размеры платы за сброс загрязнений, позволяющие выполнить стандарты качества и доставляющие максимум его целевой функции.

3. Равновесие побуждения имеет вид  $\{(KN_i)^*, (KS_i)^*, (P_i)^*\}_{i=1}^N$ , где  $(KN_i)^*, (KS_i)^*$  — оптимальная стратегия центра,  $(P_i)^* = (P_i)^*((KN_i)^*, (KS_i)^*)$  — оптимальные реакции предприятий на выбранную центром стратегию.

#### 4.3. Убеждение

Если предприятия сами стремятся поддерживать систему в устойчивом состоянии, то вместо критериев (1) и (2) у всех субъектов управления будет один общий критерий

$$J = \int_0^{\Delta} \left[ -C_A(y(t)) + \sum_{i=1}^N (R_i(\Phi_i) - VK_i(\Phi_i) - H_i(\Phi_i) - VS_i(W_i, P_i)) \right] dt \rightarrow \max(\{P_i\}_{i=1}^N),$$

рассматриваемый при ограничениях (7), (9) и (10).

### 5. МОДЕЛЬНЫЕ РАСЧЕТЫ

Все расчеты проводились для случая, когда  $C_p(Y) = D \frac{Y}{1-Y}, C_A(y) = C_1 y$ .

**Пример 1** (побуждение). Для следующего набора входных данных  $N = 1; D = 1; KN_{\max} = 8; KS_{\max} = 10; \beta_1 = 0,01$  мг/(сут·у. е.);  $\gamma_1 = 0,2$  у. е.;  $k_1 = 10^{-5}$  сут $^{-1}$ ;  $Q_1 = 10^6$  м $^3$ /сут;  $A = 700$  м $^2$ ;  $L = 100$  м;  $E = 24\,000$  м $^2$ /сут;  $\varepsilon = 0,01$ ;  $\nu_1 = 0,24$ ;  $\mu_1 = 0,5$ ;  $\Phi_0 = 5 \cdot 10^9$  у. е.;  $q_1 = 0$ ;  $C_1 = 1$ ;  $S_1 = 1$ ;  $Y_1 = 0$ ;  $W1 = 0,25\beta_1\gamma_1(\Phi_0)^{1/2}$ ;  $W2 = 0,5\beta_1\gamma_1(\Phi_0)^{1/2}$ ;  $Q_{\max} = 60$ ;  $B^{O_2} = B = 12$  мг/л (при  $t = 0$ ; при  $x = 0, L$ );  $x_1 = 30$  м;  $B_{\text{нас}}^{O_2} = 22$  мг/л;  $B_{\text{max}} = 35$  мг/л;  $B_{\text{min}}^{O_2} = 3$  мг/л;  $F_0 = F_1 = F_2 = 0$  мг/(л·сут);  $\nu_x = 100$  м/сут;  $\Delta = 360$  сут;  $K_{O_2} = 5$  сут $^{-1}$ ;  $k_c = 2,5$  сут $^{-1}$  метод побуждения реализуется, для предприятия выгодно, чтобы объемы сброса загрязнений не превышали величины  $W2$  и  $(KN_1(t))^* = 8$ ;  $(KS_1(t))^* = 1,25$ ;  $(P_1(t))^* = 0,7425$ ;  $J_0 = 5,75 \cdot 10^5$  у. е.;  $J_1 = 1,81 \cdot 10^6$  у. е.

**Пример 2** (побуждение). В случае входных данных примера 1 и  $S_1 = 0,1$  или  $D = 5$  метод побуждения не реализуется. В распоряжении центра оказываются недостаточные экономические рычаги воздействия на предприятие.

**Пример 3** (побуждение). В случае  $W1 = 0,15\beta_1\gamma_1(\Phi_0)^{1/2}$ ;  $W2 = 0,35\beta_1\gamma_1(\Phi_0)^{1/2}$  и данных примера 1 происходит уменьшение значений допустимых объемов нормативного и сверхнормативного сброса загрязнений. Предприятие своей стратегии не меняет, центр получает дополнительную прибыль по сравнению с примером 1:

$$(KN_1)^* = 8; \quad (KS_1)^* = 7,5; \quad (P_1)^* = 0,7425; \\ J_0 = 6,12 \cdot 10^5 \text{ у. е.}; \quad J_1 = 1,77 \cdot 10^6 \text{ у. е.}$$

**Пример 4** (побуждение). Для данных примера 1 и  $D = 0,001$  наблюдается уменьшение затрат на очистку сточных вод на предприятии. В результате на предприятии выбирается максимально возможная степень очистки сточных вод, и  $(KN_1)^* = 0$ ;  $(KS_1)^* = 0$ ;  $(P_1)^* = 0,99$ ;  $J_0 = 6,09 \cdot 10^5$  у. е.;  $J_1 = 1,93 \cdot 10^6$  у. е.

**Пример 5** (побуждение). При  $D = 4, S_1 = 20$  и входных данных примера 1 получим, что  $(KN_1)^* = 8$ ;  $(KS_1)^* = 0$ ;  $(P_1)^* = 0,7425$ ;  $J_0 = 7,1 \cdot 10^5$  у. е.;  $J_1 = 1,234 \cdot 10^6$  у. е.

**Пример 6** (принуждение). В случае данных примера 1 и  $S_1 = 0,1$  (значения  $q_i$  переменные и подлежат определению) и  $KN_1 = 6$ ;  $KS_1 = 8$  метод принуждения (в отличие от примера 2) реализуется и выполняются равенства  $(q_1)^* = 0,7425$ ;  $(P_1)^* = 0,7425$ ;  $J_0 = 7,63 \cdot 10^5$  у. е.;  $J_1 = 1,62 \cdot 10^6$  у. е.



**Пример 7** (принуждение). Для данных примера 1 и  $KN_1 = 6$ ;  $KS_1 = 8$  прибыль предприятия уменьшилась, центра — незначительно выросла:  $J_0 = 5,81 \cdot 10^5$  у. е.;  $J_1 = 1,57 \cdot 10^6$  у. е.

**Пример 8** (принуждение). В случае данных примера 3 и  $KN_1 = 6$ ;  $KS_1 = 8$  получим

$$(q_1)^* = 0,7425; \quad (P_1)^* = 0,7425.$$

Прибыль центра уменьшилась, а прибыль предприятия немного выросла:

$$J_0 = 6,05 \cdot 10^5 \text{ у. е.}; \quad J_1 = 1,8 \cdot 10^6 \text{ у. е.}$$

**Пример 9** (принуждение). Для данных примера 4 и  $KN_1 = 6$ ;  $KS_1 = 8$  получим, что

$$(q_1)^* = 0,866; \quad (P_1)^* = 0,9745;$$

$$J_0 = 6,5 \cdot 10^5 \text{ у. е.}; \quad J_1 = 1,65 \cdot 10^6 \text{ у. е.}$$

**Пример 10** (принуждение). Для данных примера 5 и  $KN_1 = 6$ ;  $KS_1 = 8$  получим

$$(q_1)^* = 0,2475; \quad (P_1)^* = 0,7115;$$

$$J_0 = 8,56 \cdot 10^5 \text{ у. е.}; \quad J_1 = 9,5 \cdot 10^5 \text{ у. е.}$$

**Пример 11** (убеждение). В случае данных примеров 1–3, получим, что  $(P_1)^* = 0,812$ . Совместная прибыль центра и предприятия ( $J_c$ ) значительно возросла, по сравнению с побуждением и принуждением;  $J_c = 1,84 \cdot 10^7$  у. е.

**Пример 12** (убеждение). Для данных примера 4 также наблюдается значительный рост совместной прибыли субъектов управления, а именно  $(P_1)^* = 0,866$ ;  $J_c = 6,278 \cdot 10^6$  у. е.

**Пример 13** (убеждение). В случае данных примера 5 имеем  $(P_1)^* = 0,6834$ ;  $J_c = 4,85 \cdot 10^6$  у. е.

Все примеры исследовались методом сценариев путем имитационного моделирования [7]. Уравнения (5) и (6) решались численно по явной схеме метода конечных разностей с первым порядком аппроксимации по пространственной переменной и по времени по аналогии с работой [7].

## ЗАКЛЮЧЕНИЕ

Обеспечение устойчивого развития эколого-экономических систем требует применения механизмов иерархического управления, содержащих административные, экономические и психологические составляющие и предполагающих применение методов иерархического управления—принуждения, побуждения и убеждения.

Метод принуждения позволяет поддерживать систему в устойчивом состоянии независимо от стратегий субъектов управления более низких уровней. Экономически такой подход к управле-

нию чаще всего оказывается невыгодным по сравнению с методом побуждения, который предоставляет всем субъектам управления большую свободу действий при принятии управленческих решений. С другой стороны, если у субъекта управления верхнего уровня недостаточно экономических рычагов воздействия на остальных субъектов управления, то метод побуждения может не работать, в то время как метод принуждения позволяет поддерживать систему в устойчивом состоянии.

Наиболее эффективным методом управления устойчивым развитием системы является метод убеждения, подразумевающий добровольную кооперацию субъектов управления для достижения цели, объективно отвечающей их общим интересам. В этом случае устойчивое развитие эколого-экономической системы достигается наиболее экономически выгодным для обоих субъектов управления способом.

Выбор метода управления в конкретной системе зависит от объективных (имеющихся у Ведущего возможностей воздействия на Ведомого, преследуемых им целей) и субъективных условий (уровня культуры субъектов управления).

Математические модели, разработанные в соответствии с приведенными принципами организации систем управления, применяются в системе Ростовского областного комитета по охране окружающей среды и природным ресурсам для решения задач, связанных с контролем качества речных вод.

## ЛИТЕРАТУРА

1. Рюмина Е.В. Анализ эколого-экономических взаимодействий. — М.: Наука, 2000. — 158 с.
2. Данилов-Данильян В.И., Лосев К.С. Экологический вызов и устойчивое развитие. — М.: Прогресс-Традиция, 2000. — 415 с.
3. Новая парадигма развития России. Комплексные исследования проблем устойчивого развития / Под ред. В.А. Коптюга, В.М. Матросова, В.К. Левашова. — М.: Academia; Изд-во МГУК, 1999. — 459 с.
4. Угольницкий Г.А. Управление эколого-экономическими системами. — М.: Вузовская книга, 1999. — 132 с.
5. Лаукс Д., Стединжер Дж., Хейт Д. Планирование и анализ водохозяйственных систем. — М.: Энергоатомиздат, 1984. — 400 с.
6. Фатхутдинов Р., Сивкова Л. Принуждение, побуждение, убеждение: новый подход к методам управления // Управление персоналом. — 1999. — № 2. — С. 32–40.
7. Угольницкий Г.А., Усов А.Б. Методы иерархического управления качеством воды с учетом манипуляции центра и триггеры предприятий // Водные ресурсы. — 2004. — Т. 3, № 3. — С. 375–382.

e-mail: ougoln@mail.ru, usov@math.rsu.ru

Статья представлена к публикации членом редколлегии Д.А. Новиковым. □

# АЛГОРИТМ ОПЕРАТИВНОЙ ВЕРИФИКАЦИИ ОПАСНОГО ЭПИДЕМИЧЕСКОГО ЗАБОЛЕВАНИЯ ВЕРОЯТНОСТНЫМ МЕТОДОМ И ЕГО РЕАЛИЗАЦИЯ

О.В. Михеев<sup>(1)</sup>, А.С. Самохина<sup>(2)</sup>

<sup>(1)</sup>ЗАО НВК «ВИСТ», г. Москва;

<sup>(2)</sup>Институт проблем управления им. В.А. Трапезникова РАН, г. Москва

Разработан алгоритм для системы поддержки принятия решений в целях оперативной верификации опасного эпидемического заболевания при управлении предотвращением и ликвидацией последствий биотеррактов.

## ВВЕДЕНИЕ

В последнее десятилетие активно обсуждаются вопросы предупреждения распространения опасных эпидемических заболеваний (ОЭЗ). Эти вопросы имеют комплексный характер и связаны не только с организацией работы санитарно-эпидемиологических служб, но и с проблемами биологической безопасности [1, 2].

Как правило, сценарии биологических чрезвычайных ситуаций (БЧС) крайне не оптимистичны как в смысле человеческих потерь, так и в смысле затрат на ликвидацию их прямых последствий, а также потерь вследствие дезорганизации экономики региона и последствий психологического воздействия на население. Применение компьютерной техники для принятия решений по предупреждению, противодействию и ликвидации последствий БЧС крайне необходимо, так как повышает управляемость ситуаций в режиме реального времени, увеличивает информированность управляющих органов, сокращает время принятия решений. Одна из важнейших мер предупреждения распространения опасных заболеваний состоит в создании и дальнейшем использовании системы поддержки принятия решений (СППР) по преодолению и ликвидации последствий БЧС. Такая система, работающая практически в реальном масштабе времени, помогает организовать поддержку процесса принятия решений в условиях острой нехватки времени, противоречивой информации об оперативной обстановке и противоречивых мнениях руководителей различных служб [3].

В СППР должно быть предусмотрено два режима (этапа) работы — подготовительный и оперативный [4]. В подготовительном режиме, в частности, осуществляется классификация различных объектов, с которыми затем работают в оперативном режиме. Такими объектами могут быть, например, особо опасные эпидемические заболевания, которые могут быть применены при биологическом терроризме. В оперативном режиме некоторый инцидент (событие, зафиксированное диспетчерской службой), должен быть сначала верифицирован как БЧС, т. е. должно быть оперативно определено, относится ли заболевание к опасным эпидемическим. Если наблюдается ОЭЗ, то проводится идентификация заболевания, масштаба и способа его распространения, т. е. определяется сценарий развития события [5, 6]. Рассмотрим алгоритм оперативной верификации заболевания, включенного в базу знаний, созданную в подготовительном режиме работы СППР.

## 1. ОБОБЩЕННЫЙ АЛГОРИТМ ОПЕРАТИВНОЙ ВЕРИФИКАЦИИ ОПАСНЫХ ЭПИДЕМИЧЕСКИХ ЗАБОЛЕВАНИЙ

Сформулируем алгоритм решения диагностической задачи, позволяющий верифицировать известные типы особо опасных инфекционных заболеваний.

Пусть  $O(i, j)$  — вероятность  $i$ -го общего признака для  $j$ -го заболевания,  $i = \overline{1, N}$ ,  $j = \overline{1, M}$ . Общими считаются признаки, характерные для ряда заболеваний.



Пусть  $C(i, j)$  — вероятность  $i$ -го специфического признака для  $j$ -го заболевания,  $i = \overline{1, K}$ ,  $j = \overline{1, M}$ . Специфическими считаются признаки, характерные для данного конкретного заболевания.

Далее, пусть  $BO(i)$  — характеристика  $i$ -го общего признака больного,  $i = \overline{1, N}$ , а  $BC(i)$  — характеристика  $i$ -го специфического признака больного,  $i = \overline{1, K}$ . Положим

$$BO(i) = \begin{cases} 1, & \text{если признак присутствует,} \\ 0,5, & \text{если признак возможен,} \\ 0, & \text{если признак отсутствует,} \end{cases}$$

и

$$BC(i) = \begin{cases} 1, & \text{если признак присутствует,} \\ 0,5, & \text{если признак возможен,} \\ 0, & \text{если признак отсутствует,} \end{cases}$$

Обозначим  $KO(j)$  — число общих признаков для  $j$ -го заболевания, вероятность которого не менее 0,5,  $KC(j)$  — число специфических признаков для  $j$ -го заболевания, вероятность которого не менее 0,5.

Тогда вероятность того, что пациент страдает  $j$ -м заболеванием, определяется выражением:

$$B(j) = \frac{\sum_{i=1}^N BO(i)O(i, j) + \sum_{i=1}^K BC(i)C(i, j)}{KO(j) + KC(j)} \times \frac{\sum_{i=1}^N BC(i)C(i, j)}{KC(j)}.$$

Если вероятность какого-либо одного или нескольких заболеваний превосходит некоторое пороговое значение, то система выдает диагностическую гипотезу о возможном наличии заболевания с определенной степенью вероятности.

## 2. ВХОДНЫЕ ДАННЫЕ СИСТЕМЫ ПОДДЕРЖКИ ПРИНЯТИЯ РЕШЕНИЙ ДЛЯ ВЕРИФИКАЦИИ ОПАСНЫХ ЭПИДЕМИЧЕСКИХ ЗАБОЛЕВАНИЙ

Входными данными системы являются:

- ответы пользователя на вопросы системы;
- база данных с описаниями признаков;
- база данных с описанием заболеваний;
- таблица соответствий между заболеваниями и признаками;
- таблица значений признаков заболеваний;
- база данных с информацией о пациентах.

Входные данные можно разбить на два больших блока:

- оперативные данные, поступающие с пользовательского интерфейса, т. е. интерфейса лечащего врача; сюда также входит и база данных о пациентах (их «больничные карточки»);
- содержимое базы знаний, заполненной экспертами на предварительном этапе построения СППР.

Структура и представление групп вопросов в опроснике, заполняемого лечащим врачом, служат эффективным инструментом для ввода данных, а также для наглядного рабочего представления данных в форме табл. 1.

База знаний по опасным эпидемическим заболеваниям заполняется экспертами на этапе классификации ОЭЗ, во время подготовительного этапа построения системы противодействия биотеррактам [7].

Таблица 1

Пример сокращенного опросника, заполняемого лечащим врачом

Фамилия, имя, отчество	Алексеев Алексей Борисович
Дата рождения	02.01.1980
Дата заболевания	08.12.2005
Дата обращения за медицинской помощью	17.12.2005
Жалобы	Озноб, жар, головная боль, головокружение, общая слабость, мышечная слабость, повышенная утомляемость, тошнота
Общее состояние больного	Тяжелое, сознание сохранено, возбуждение, нарушение речи, бред, вестибулярные нарушения, температура 38—40 °С и более, резкое повышение температуры
Состояние кожных покровов и слизистых	Гиперемия лица, отечность, одутловатость лица, гиперемия слизистых ротоглотки, дыхание затруднено, сухой кашель, кашель со стекловидной мокротой
Состояние сердечно-сосудистой системы	Тахикардия, гипотония
Симптомы поражения нервной системы	Симптомы поражения мозговых оболочек
Состояние органов гепатобилиарной системы	Увеличение печени, симптомы нарушения гемодинамики
Оценка вероятности выявления ОЭЗ	0,65
Оценка условной вероятности нозоформы	Чума: 0,64, вирусная геморрагическая лихорадка 0,51

Для построения базы знаний по особо опасным инфекционным заболеваниям привлекается группа, состоящая из  $L$  экспертов. Каждый член группы оценивает других членов по 10-балльной шкале. Для каждого эксперта вычисляется его вес

$$\rho^0(i) = \frac{\sum_{j=1, j \neq i}^L \delta(i, j)}{L-1},$$

где  $\delta(i, j)$  — оценка  $j$ -м экспертом  $i$ -го эксперта,  $\rho^0(i)$  — первоначальный вес  $i$ -го эксперта. Для уточнения веса  $i$ -го эксперта целесообразно воспользоваться формулой

$$\rho^1(i) = \frac{\sum_{j=1, j \neq i}^L \rho^0(i) \delta(i, j)}{L-1}.$$

Каждый эксперт заполняет табл. 2, проставляя значения  $O(i, j, k)$  — вероятность  $i$ -го общего симптома  $j$ -го больного, проставленная  $k$ -м экспертом и  $C(i, j, k)$  — вероятность  $i$ -го специфического симптома  $j$ -го больного, проставленная  $k$ -м экс-

пертом в соответствующих графах. Элементы итоговой таблицы строятся по формуле:

$$O(i, j) = \frac{\sum_{k=1}^L \rho^1(k) O(i, j, k)}{\sum_{k=1}^L \rho^1(k)},$$

аналогично вычисляются значения  $C(i, j)$ . Пример начала описания итоговых таблиц приведен в табл. 2.

### 3. ВЫХОДНЫЕ ДАННЫЕ СИСТЕМЫ ПОДДЕРЖКИ ПРИНЯТИЯ РЕШЕНИЙ ДЛЯ ВЕРИФИКАЦИИ ОПАСНЫХ ЭПИДЕМИЧЕСКИХ ЗАБОЛЕВАНИЙ

Выходные данные программы представляют собой диагноз, поставленный на основе наблюдаемых симптомов и базы знаний о заболеваниях. Диагноз выдается на экран в качестве окончательного ответа экспертной системы пользователю. Кроме того, информация об обнаруженном заболевании и наблюдаемых признаках заносится в электронную карточку пациента.

Таблица 2

Пример итоговой таблицы для чумы, вирусной геморрагической лихорадки (в. г. л.), натуральной оспы (н. о.)

Краткий анамнез	Личные данные пациента							
	Дата начала заболевания							
	Дата обращения за медицинской помощью							
	Начало заболевания	острое	0,9	0,7	0,9	1	1	1
		подострое	0,1	0,2	0,1	0	0	0
		вялотекущее	0,0	0,1	0,0	0	0	0
	Заболевание		Чума	В. г. л.	Н. о.	Чума	В. г. л.	Н. о.
	Жалобы	озноб	0,9	0,9	0,9	1	1	1
		жар	0,9	0,8	0,7	1	1	1
		головная боль	0,9	0,9	0,7	1	1	1
		головокружение	0,5	0,1	0,1	1	0	0
		нарушение зрения	0,0	0,0	0,0	0	0	0
		светобоязнь	0,0	0,0	0,0	0	0	0
		общая слабость	0,9	0,9	0,9	1	1	1
		мышечная слабость	0,9	0,9	0,9	1	1	1
		повышенная утомляемость	0,9	0,9	0,9	1	1	1
		сонливость	0,0	0,1	0,4	0	0	0,5
бессонница		0,3	0,1	0,1	0,5	0	0	
мышечные боли		0,8	0,7	0,1	1	1	0	
боли в суставах		0,0	0,3	0,1	0,5	0	0	
боль в грудной клетке		0,7	0,1	0,1	1	0	0	
одышка	0,5	0,1	0,1	1	0	0		
кашель	0,5	0,1	0,1	1	0	0		
озноб	0,9	0,9	0,9	1	1	1		





В процессе работы система генерирует несколько рабочих версий окончательного диагноза, и затем «отсеиваются» лишние гипотезы, которые имеют вес, меньший, чем некоторое значение, заранее заданное специалистом.

Например, в процессе работы системы сформировалось пять версий диагноза с вероятностями от 0,51 до 0,98. Пусть порог уверенности, заданный специалистом, равен 0,75. Тогда система выдаст все версии, вероятности которых больше 0,75, в порядке убывания их вероятностей, «отсеив» остальные:

0,98 — вероятность заболевания 1;  
0,93 — “ “ “ 2;  
0,87 — “ “ “ 3.

Система также ведет регистрацию пациентов, их состояний и статистику заболеваний: частоту встречаемости заболевания, сезонность эпидемических заболеваний, возрастной состав заболевших, гендерный состав заболевших и др.

Разработанная СППР обладает дружественным пользовательским интерфейсом, не требующим специальных навыков и понятным как для лечащего врача, так и для эксперта-медика.

### ЗАКЛЮЧЕНИЕ

Предложенный алгоритм позволяет определять наличие опасных эпидемических заболеваний с некоторыми вероятностями. При наличии заболевания с помощью компьютерной системы предотвращения и ликвидации последствий биотеррактов выполняется его детальная идентификация, предполагающая дальнейшие клинические анализы в специализированных учреждениях. На основании поставленного диагноза и оперативной обстановки можно моделировать сценарии развития опасных эпидемических заболеваний [5, 6, 8].

### ЛИТЕРАТУРА

1. Организация и проведение противоэпидемических мероприятий при террористических актах с применением биологических агентов / Методические рекомендации МР 2510/11646-01-34.
2. Противодействие биологическому терроризму. Практическое руководство по противоэпидемическому обеспечению / под ред. Г.Г. Онищенко. — М., 2003. — 311 с.
3. Трахтенгерц Э.А., Шершаков В.М., Камаев Д.А. Компьютерная поддержка управления ликвидацией последствий радиационного воздействия. — М.: СИНТЕГ, 2004. 456 с.
4. Самохина А.С. Общая постановка задачи классификации и идентификации в системе поддержки принятия решений при биологической чрезвычайной ситуации // Межвуз. сб. науч. тр. «Теоретические вопросы вычислительной техники, программного обеспечения и информационных технологий в муниципальном хозяйстве». — М., 2005. — С. 216—221.
5. Габусу П.А., Михеев О.В., Самохина А.С. Экспериментальные исследования имитационной математической модели распространения опасного эпидемического заболевания // Тр. междунар. конф. «Параллельные вычисления и задачи управления» РАСО'2006 / Ин-т пробл. упр. — М., 2006. С. 1494—1508.
6. Габусу П.А., Михеев О.В., Самохина А.С. Идентификация сценариев эпидемического заболевания // Проблемы безопасности и чрезвычайных ситуаций. — 2006. — № 2. — С. 92—106.
7. Габусу П.А., Михеев О.В., Самохина А.С. Анализ схем обработки первичных данных в системе предотвращения биологической чрезвычайной ситуации // Там же. — № 4. — С. 64—71.
8. Габусу П.А., Михеев О.В., Самохина А.С. Взаимодействие V-модели и имитационной математической модели распространения опасного эпидемического заболевания // Там же. — № 3. — С. 96—105.

☎ (495) 318-69-73,

e-mail: absamokhin@yandex.ru, olgm@ncgroup.ru

Статья представлена к публикации членом редколлегии В.Н. Новосельцевым. □

## Новая книга

**Бухарин С.Н., Цыганов В.В. Методы и технологии информационных войн. — М.: Академический проект, 2007. — 382 с.**

Изложены методы и технологии информационных войн, основанные на их теории и методологии, а также практические применения. На основе изучения целей информационных войн спроектированы процедуры ситуационного анализа и алгоритмы определения стратегий информационного противоборства (в том числе провокаций). Разработаны процедуры планирования информационных операций, алгоритмы принятия решений в информационных войнах и подходы к управлению информационным противоборством. Эти методы и технологии применены для анализа и проектирования комплексных систем управления противоборством в информационных войнах разной природы и масштаба — от корпоративных и отраслевых до региональных и национальных. Проанализированы экологические информационные войны в условиях глобализации. В приложениях даны необходимые сведения об экспертных и статистических методах.

Для руководителей и специалистов по информационным операциям и связям с общественностью, управлению, социологии, политологии, экономике, профессорско-преподавательского состава и студентов социально-экономических кафедр вузов, а также читателей, интересующихся информационным противоборством.

# МОДЕЛИ СИСТЕМЫ УПРАВЛЕНИЯ КАЧЕСТВОМ ПОДГОТОВКИ СПЕЦИАЛИСТОВ ПРИ ИННОВАЦИОННЫХ ТЕХНОЛОГИЯХ ОБУЧЕНИЯ

А.Ю. Беляков, С.Г. Пуйсанс, И.Д. Столбова, В.А. Харитонов

*Пермский государственный технический университет*

Рассмотрены две модели, первая позволяет оценить устойчивость образовательного процесса относительно достижения локальных целей обучения, а вторая — глобальную направленность процесса к формированию заданных компетенций специалиста.

## ВВЕДЕНИЕ

Знания и образование в современном мире приобрели несомненную экономическую ценность. По оценкам Всемирного банка, сделанным в 1994 г., 76 % национального богатства США составляет человеческий капитал, т. е. накопленные населением знания и умения, используемые ими в производственной деятельности и повседневной жизни. Физический или воспроизводимый капитал дает 19 % богатства США, а природный фактор — 5 %. В Западной Европе эти цифры выглядят, соответственно, как 74, 23 и 2. А вот какие у нас: благосостояние России лишь на 50 % держится на человеческом капитале, 10 % дает воспроизводимый капитал и 40 % обеспечивают природные ресурсы [1]. Поэтому проблема качества интеллектуальных ресурсов и, прежде всего, качества подготовки специалистов с высшим образованием, способных решать самые сложные проблемы прогресса во всех сферах науки, техники, технологии и общественной практики, выдвинулась в число первоочередных национальных проблем России.

До сих пор качество подготовки выпускников у нас в стране оценивалось как результат предоставляемых образовательных услуг уже фактически после того, как воздействие на студента закончено, по результатам итоговой аттестации. И поэтому принимались во внимание такие показатели, как число дипломов с отличием и число внедренных работ в практику. Но сегодня важнее другое — сформировать систему, которая позволяла бы управлять качеством в процессе обучения студентов

и обеспечивать качество подготовки в самом образовательном процессе. Этому способствует внедрение компетентностного подхода при формировании рабочих учебных планов по отдельной образовательной программе и применение инновационных технологий, к которым можно отнести кредитно-модульную и дистанционную технологии обучения [2].

## 1. МОДЕЛИ СИСТЕМЫ УПРАВЛЕНИЯ КАЧЕСТВОМ ОБРАЗОВАТЕЛЬНОГО ПРОЦЕССА

Предлагаемые в настоящей работе модели предназначены для совершенствования научно-методического обеспечения системы управления качеством подготовки по отдельной образовательной программе. Их разработка вызвана необходимостью эффективной поддержки решения задач управления образовательным процессом — анализа текущего состояния, прогноза, целеполагания, планирования, обеспечения ресурсами, контроля, оперативного управления и анализа изменений.

Для выработки эффективных управляющих воздействий, начиная с этапа целеполагания и заканчивая этапом оперативного управления, управляющему образованием Центру (в качестве такого Центра может выступать выпускающая кафедра или деканат факультета) необходимо обладать достаточной информацией о поведении управляемых субъектов, в частности — относительно результатов их деятельности. В сложных системах, описываемых многими критериями, в силу ограниченности возможностей управляющего органа



по переработке информации или в силу отсутствия детальной информации целесообразно применение механизмов комплексного оценивания, которые позволяют осуществлять свертку показателей, т. е. агрегировать информацию о результатах деятельности отдельных элементов.

Задачи Центра состоят в:

- разработке моделей и методов исследования эффективности образовательных систем;
- компьютерно-игровом моделировании индивидуальных образовательных траекторий с устойчивыми показателями качества подготовки специалистов;
- разработке прикладных оценочных средств на основе механизмов комплексного оценивания;
- создании методических и дидактических компонентов комплексной оценки качества образования на каждом этапе процесса обучения.

Главной целью исследования образовательных систем можно считать нахождение педагогических условий, обеспечивающих максимальную эффективность процессов обучения.

Образовательные системы, состоящие из обучаемых (учеников), обучающих (учителей) и инфраструктуры (подсистемы обслуживания учебного процесса), относятся к классу сложных систем. Они отличаются структурной сложностью, определяемой числом их состояний, разнообразием связей между ними и числом иерархических уровней восхождения по ступеням познания, сложностью функционирования (поведения), т. е. правилами перехода из состояния в состояние, и сложностью выбора поведения в многоальтернативных учебных ситуациях, характеризующихся целенаправленностью и гибкостью системы при недостаточной определенности педагогических условий. Сложные системы исследуются с позиций системного подхода, одним из основных принципов которого является принцип многомодельности, устанавливающий для каждого класса моделей свою гносеологическую нагрузку. Для образовательного процесса, как правило, модели обслуживаются экспертной (нечеткой) информацией.

Определяющими показателями эффективности (качества) образовательного процесса будем считать степень его устойчивости на этапах достижения частных целей обучения и степень соответствия сформулированным квалификационным характеристикам (профессиональным и общим компетенциям) специалиста. Первый показатель описывает локальную эффективность — эффективность использования текущего учебного времени, второй — глобальную направленность образовательного процесса к овладению обучаемыми конкретной специальностью (профессией). Совокупность этих показателей эффективности отра-

жает качество образовательного процесса. Отсюда вытекает центральная роль педагога в образовательном процессе: последовательное предложение обучаемому на выбор тех и только тех частных целей обучения, которые рационально ведут его к погружению в сформулированную предметную область и достаточно строго соответствуют подготовленности респондента к усвоению нового учебного материала. Это, в свою очередь, ставит задачу систематического контроля уровня определенных компетенций, стимулирующего ученика возможностью продолжить участие в образовательном процессе, постепенно формируя индивидуальную образовательную траекторию. Весь учебный процесс разбивается на отдельные взаимосвязанные этапы (модули), имеющие различные ценностные характеристики, выраженные в кредитах, в зависимости от важности этих модулей для формирования заявленных компетенций выпускника. Важно отметить, что кредиты в данном случае играют роль зачетных единиц, а соответствующая трудоемкость каждого модуля зависит от способностей обучаемого и влияет на устойчивость образовательного процесса [3].

Для исследования локальной устойчивости образовательного процесса построим его упрощенную модель на принципах итеративного научения [4]:

$$z_i = 1 - \exp(-kf_i(z_{i-1})t), \quad (1)$$

где  $z_i \in [0, 1)$  — полуинтервал (вследствие неисчерпаемости предметной области) уровня обучения ученика на  $i$ -м этапе;  $k$  — педагогические условия, включающие в себя индивидуальные способности ученика к обучению;  $f_i(z) \in [0, 1)$  — характеристика влияния уровня предшествующих знаний обучаемого на локальную эффективность  $i$ -го этапа обучения.

Локально устойчивый образовательный процесс обеспечивает на каждом этапе заданный уровень обучения:  $z_i \geq z_c$ .

Для тривиального случая при условии  $z_i = z_{i-1} = f_i(z) = z_c$  из формулы (1) следует выражение, устанавливающее минимальную трудоемкость этапа:  $t_c = \ln(1 - f(z_c))/kz_c$ .

Можно показать, что для фиксированной пары  $(z_c, t_c)$ , зависящей от педагогических условий  $k$  и вида функции  $f(z_c)$ , наблюдается сходимость уровня обучения  $z_{i \rightarrow \infty} \rightarrow z_c$  при начальных отклонениях (возмущениях):  $|z_0 - z_c| > 0$ .

При обычной технологии обучения трудоемкость каждого этапа устанавливается одинаковой для обучаемых с различными значениями параметра  $k$  из расчета на некоторый средний уровень



Рис. 1. Структура моделей и методы обеспечения инновационной технологии обучения

$t_c$ , в результате чего возникают неоправданные потери учебного времени либо снижается качество обучения.

При адаптированной к способностям ученика дистанционной или кредитно-модульной технологии обучения появляется возможность оптимального выбора трудоемкости этапа обучения, исходя из контекстных педагогических условий.

Адаптированность технологии [5] в данном случае может означать следующее:

- установление по методике программированного обучения [6] состава и последовательности локальных мезо- и макропроцессов обучения согласно свойствам фрактальности предметной области;
- экспертное обоснование и оценка логических связей между этапами, например, с помощью механизмов комплексного оценивания;
- управление переходом от этапа к этапу по критерию готовности согласно установленному ранее уровню педагогических условий и с учетом активности обучаемых как элементов образовательной системы.

Совокупность востребованных моделей и методов для придания адаптированности инновационной технологии обучения представлена на рис. 1.

Целенаправленное развитие инновационных технологий обучения в рамках перечисленных средств научного исследования способно раскрыть огромный потенциал качества образования, заложенный в новых возможностях индивидуального подхода, поддерживаемых современными информационными технологиями.

Востребованное обсуждаемой моделью управления качеством образования комплексное оценивание всегда проходит две обязательные фазы: раз-

работку (конструирование) деревьев критериев с матрицами свертки в узлах на стадии разработки учебных планов и собственно оценивание, сопровождаемое по возможности глубоким анализом полученных результатов в интересах дальнейшего развития объектов оценивания. Рассмотрим новый класс инструментальных средств комплексного оценивания, полученный в ходе построения упомянутой модели системы управления и использующий топологические свойства матриц свертки с целью расширения функциональных возможностей прикладных программных продуктов в каждой фазе их применения [7, 8].

Для комплексного оценивания качества образования с помощью реализованной модели необходимо все переменные (частные критерии качества) привести к стандартной шкале [1, 5], а назначенные экспертами их текущие значения ввести в понятной для них дефазифицированной форме  $\widehat{X}$ . Окончательный и промежуточные результаты счета представляются над каждой матрицей дерева целей образовательного процесса.

Подавая на вход механизма оценивания последовательность наборов исходных данных, можно получить информацию об изменениях (развитии, деградации) комплексной оценки качества сопоставляемых объектов в виде траекторий, составленных из последовательностей дискретных значений. При управлении развитием (изменениями) комплексной оценки качества объекта образовательной системы возникает необходимость в установлении критериев, по отношению к которым она наиболее чувствительна в рассматриваемой рабочей точке  $(\widehat{X}_{P_1}, \dots, \widehat{X}_{P_N})$ .

Для отдельной матрицы можно построить линейаризованные частные функции чувствительности графоаналитическим способом, используя топологические свойства заданной матрицы свертки вдоль прямых, параллельных одной либо другой координате и проходящих через рабочую точку. Однако данным способом сложнее построить транзитивные функции чувствительности, описывающие чувствительность комплексной (общей) оценки от частного критерия.

Достижимость значений комплексной оценки компетенций студента, требуемых для успешного освоения образовательной программы, путем изменения частных критериев может быть установлена пошагово. На каждом шаге в окрестности рабочей точки для этого строится приращение комплексной оценки

$$\Delta \widehat{X}_P = \frac{\partial \widehat{X}_P}{\partial \widehat{X}_{P_1}} \Delta \widehat{X}_1 + \dots + \frac{\partial \widehat{X}_P}{\partial \widehat{X}_{P_N}} \Delta \widehat{X}_N.$$



Как и все графоаналитические методы, данный подход приближенный. Уменьшить его вычислительную погрешность можно, переходя к прямым вычислениям функции чувствительности, которые могут быть ускорены табличными процедурами [7, 8].

## 2. ПРИМЕР

В качестве иллюстративного примера рассмотрим свертку уровней профессиональной  $X_{\text{проф}}$  и социальной  $X_{\text{соц}}$  компетенций специалиста по информационным технологиям, оценивающую качество его подготовки  $X_{\text{спец}}$  (рис. 2). Предположим, что совместно с работодателем составлен «портрет» специалиста, который содержит оптимальный набор ключевых компетенций, как профессиональных, так и общих (социальных и личностных). Разработаны типовой учебный план подготовки специалиста и возможные образовательные траектории в рамках выбранной инновационной технологии. Построено дерево целей и разработаны диагностические средства оценивания уровня каждой компетенции по четырехбалльной системе на входе и выходе каждого образовательного модуля.

Рассмотрим процесс комплексного оценивания уровня компетенций студента и выработки управляющих решений в рамках отдельного образовательного модуля. Считается, что при изучении данного модуля формируется не только определенная ключевая профессиональная компетенция (например, способность проектирования баз данных), но и одна из необходимых социальных компетенций (например, социальная коммуникативность). Поэтому при реализации учебной программы по дисциплине важно знать, как по текущей оценке уровня различных компетенций студента организовать учебный процесс, чтобы получить наибольший эффект в качестве подготовки специалиста. Для этой цели и служит предлагаемый математический инструментарий.

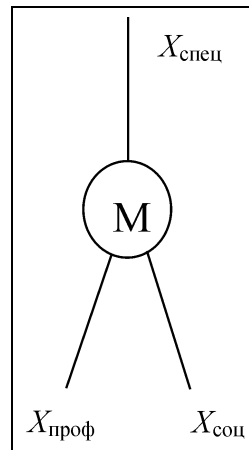


Рис. 2. Дерево свертки профессиональной и социальной компетентностей

В рассматриваемом примере по специальной методике с учетом предпочтений работодателя и на основе экспертных оценок синтезирован механизм свертки — матрицы  $M$  (см. рис. 2), наполнение и топология которой приведена на рис. 3. Содержательно данный вариант свертки стимулирует приоритетное развитие уровня профессиональной компетентности в соответствии с заказом на выпускника. Положим, текущее состояние подготовленности будущего специалиста оценивается парой значений уровня компетенций ( $X_{\text{проф}} = 2,8$ ;  $X_{\text{соц}} = 1,5$ ), образуя рабочую точку на топологичес-

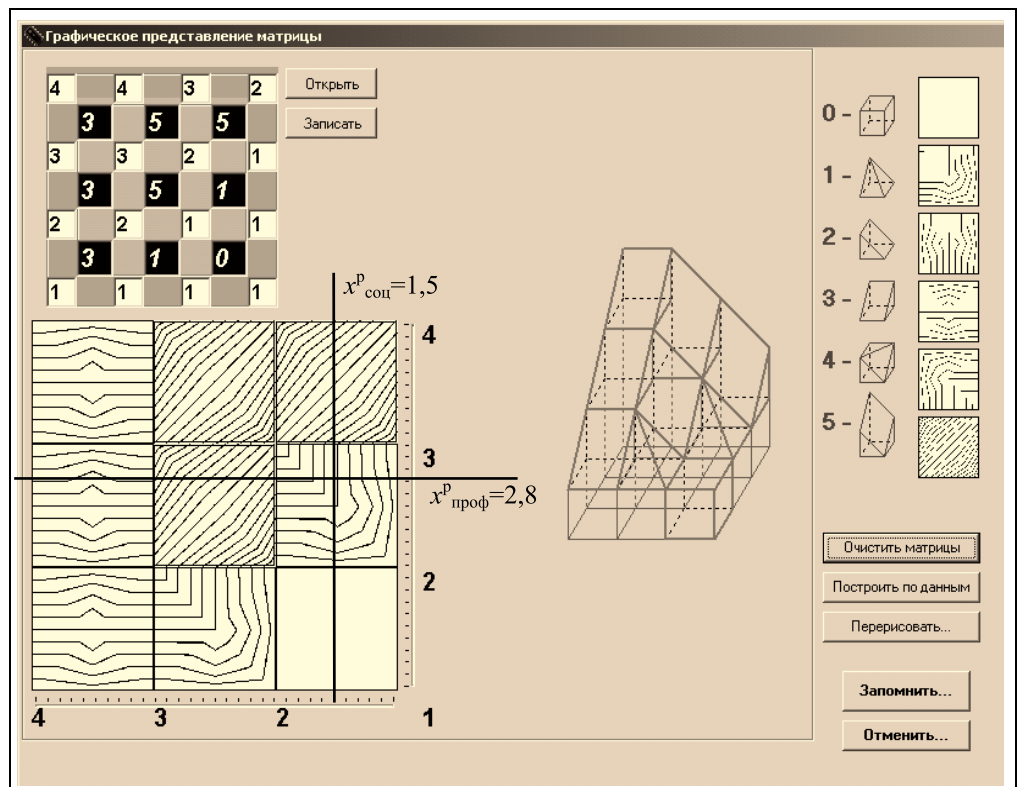
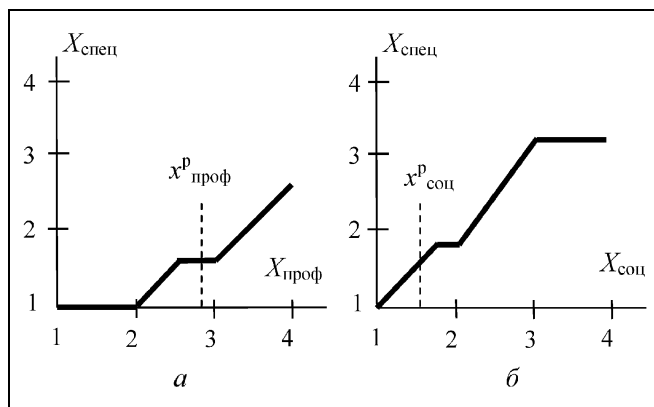


Рис. 3. Графическое представление топологии матрицы свертки уровней профессиональной и социальной компетентностей

**ЗАКЛЮЧЕНИЕ**

Результатом подготовки выпускника вуза является некоторая интегральная характеристика, определяемая набором компетенций специалиста данного направления или выпускника по отдельной образовательной программе. В условиях модернизации Российского образования очень важна задача текущей обобщенной оценки уровня сформированности заданных компетенций студента, а также подбора индивидуальных траекторий обучения с учетом особенностей конкретного обучающегося. Предложенные модели управления позволяют не только комплексно оценивать текущее состояние подготовки студента, но и индивидуализировать образовательный процесс с целью достижения требуемого уровня компетенций каждым студентом, а следовательно, повысить качество подготовки будущего специалиста.



**Рис. 4.** Функции чувствительности комплексной оценки уровня подготовленности специалиста  $X_{\text{спец}}$  к уровням профессиональной  $X_{\text{проф}}$  (а) и социальной  $X_{\text{соц}}$  (б) компетентностей

кой картине матрицы свертки (см. рис. 3). Для рассматриваемой рабочей точки построены функции чувствительности, представленные в графическом виде на рис. 4.

Из анализа поведения функций видно, что повышение общего уровня подготовленности специалиста в данной ситуации следует связывать, прежде всего, с успешным освоением социальной составляющей его подготовки (уровень социальной компетенции в данной точке образовательной траектории сильнее влияет на требуемый уровень подготовки специалиста). Поэтому для педагога в данный момент образовательного процесса целесообразно организовать дистанционные формы общения со студентом (чат-консультации, групповые форумы по актуальным проблемам, дистанционную приемку домашних заданий, интернет-тестирование и т. п.), которые существенно развивают заявленную социальную компетентность ученика и, соответственно, повышают качество выпускника.

Таким образом, предлагаемая модель позволяет не только оценивать текущее состояние в подготовке студента, но и управлять учебным процессом с целью повышения качества будущего специалиста.

**ЛИТЕРАТУРА**

1. Качалов В.А. Стандарты ИСО 9000 и проблемы управления качеством в вузах. — М., 2001.
2. Ташинов А., Лалетин В., Столбова И. Формирование общих и профессиональных компетенций при инновационных технологиях обучения // Высшее образование в России. — 2007. — № 1. — С. 128—133.
3. Сазонов Б., Максимов Н., Каравалева Е. Зачетные единицы в ГОС нового поколения // Там же. — 2006. — № 10. — С. 3—15.
4. Новиков Д.А. Закономерности итеративного научения. — М.: Институт проблем управления РАН, 1998. — 77 с.
5. Растринин Л.А., Эренштейн М.Х. Адаптивное обучение с моделью обучающегося. — Рига: Зинатне, 1988. — 160 с.
6. Беспалько В.П. Программированное обучение (дидактические основы). — М.: Высшая школа, 1970. — 300 с.
7. Системы конструирования матриц свертки в экспертных задачах комплексного оценивания / А.А. Белых и др. // Вестник УГТУ — УПИ: Строительство и образование. — Екатеринбург, 2006. — С. 24—26.
8. Бурков В.Н., Новиков Д.А. Как управлять проектами. — М.: СИНТЕГ-ГЕО, 1997. — 188 с.

☎ (342) 241-07-09, e-mail: nedstf@pstu.ac.ru

Статья представлена к публикации членом редколлегии Ф.Т. Алескеровым. □

## Новая книга

**Левин В. И. Современная математическая логика. — Изд-во Пенз. гос. технол. академии, 2007. — 96 с.**

Рассмотрены некоторые вопросы современной (неклассической) логики: непрерывная логика и ее обобщения, интервальные математика и логика, порядковая логика и логические определители, логика и множества, логическое моделирование систем, история логического моделирования в технике.

Книга подготовлена на кафедре «Научные технологии» Пензенской государственной технологической академии и предназначена для изучения курса «Математическая логика и теория алгоритмов» по специальности 220100 «Вычислительные машины, комплексы, системы и сети» и выполнения соответствующих курсовых и дипломных работ.

Для студентов, аспирантов, соискателей и инженеров-исследователей, применяющих в своей работе логические методы или интересующихся ими.

# СОЗДАНИЕ И ПРИМЕНЕНИЕ КОМПЬЮТЕРНОЙ ТЕХНОЛОГИИ ПОВЫШЕНИЯ БЕЗОПАСНОСТИ УПРАВЛЕНИЯ МОРСКИМИ ПОДВИЖНЫМИ ОБЪЕКТАМИ<sup>1</sup>

В.Г. Борисов, С.К. Данилова, В.О. Чинакал

*Институт проблем управления им. В.А. Трапезникова РАН, г. Москва*

Рассмотрены основные пути повышения безопасности управления движением морских подвижных объектов на основе единой сквозной компьютерной технологии поддержки разработки технических средств и систем управления, отработки алгоритмических структур управления на полной математической модели движения объекта и различных вариантов алгоритмов управления и использования технических средств управления в штатных и нештатных ситуациях.

## ВВЕДЕНИЕ

В современных условиях одна из актуальных проблем состоит в повышении безопасности и эффективности управления движением подвижных объектов, в частности, морских подвижных объектов, например, подводных лодок и разнообразных подводных аппаратов, использующих различные гидродинамические и гидростатические технические средства управления. При управлении такими объектами операторам различных постов управления приходится решать комплекс достаточно сложных задач. Это и оперативные задачи по прокладке или изменению заданной траектории движения МПО в зависимости от известных заранее или возникающих в процессе управления статических или динамических траекторных угроз, и задачи выбора ТСУ, и задачи координированного управления различными средствами, решаемые с учетом эффективности применения различных ТСУ и фактических ресурсных ограничений. При-

ходится учитывать не только траекторию и динамику самого движущегося управляемого объекта, состояние и предельные возможности ТСУ и системы управления, но и возможное изменение внешней обстановки и действие различных возмущений (движение других объектов, плохая видимость, ветер, течения, узкости, мели и т. п.).

Наиболее сильное влияние на безопасность и эффективность управления МПО оказывают:

- качество используемого в СУ алгоритмического обеспечения;
- качество информационного обеспечения;
- эффективность программного обеспечения;
- уровень подготовки операторов постов управления МПО.

В связи с этим в данной работе рассмотрены возможности создания и применения общей методики для отработки различных аспектов безопасного управления МПО и, в первую очередь, совершенствования АО, ИО и ПО создаваемых перспективных систем управления нового поколения и подготовки высококвалифицированных операторов постов управления МПО и средств поддержки их работы. Для обеспечения эффективной поддержки основных этапов методики была разработана компьютерная технология, реализуемая на специ-

<sup>1</sup> Работа выполнена при поддержке РФФИ (проект № 05-08-01334а) и Программы фундаментальных исследований по теме № 301-04/01 ИПУ РАН в рамках комплексного проекта РАН № 2462.

ально созданном специализированном сетевом стенде [1].

#### Принятые сокращения:

АО — алгоритмическое обеспечение;  
БД — база данных;  
ИИ — искусственный интеллект;  
ИО — информационное обеспечение;  
ИСПР — интеллектуальная система поддержки работы операторов;  
КТ — компьютерная технология;  
МПО — морской подвижный объект;  
ПО — программное обеспечение;  
ПСУ — перспективная система управления;  
ПТС — программно технические средства;  
СИМ — система имитационного моделирования;  
СОО — система обучения операторов;  
СППР — система поддержки принятия решений;  
СУ — система управления;  
ТПР — типовые проектные решения;  
ТСУ — технические средства управления.

Под перспективной системой управления нового поколения понимается модульная, параметрически настраиваемая интегрированная система навигации и управления, включающая в себя в качестве основных компонентов собственно систему управления объектом, систему навигации, встроенные в СУ систему поддержки принятия решений и систему обучения операторов постов управления МПО. Для выработки управляющих решений в такой системе применяются эффективные методы текущей адаптации параметров моделей объекта и ТСУ, получение оценок действующих возмущений и подстройки параметров алгоритмов управления. Построение и коррекция параметров алгоритмов управления осуществляется на базе методов оптимального управления, элементов и методов искусственного интеллекта, реализуемых в интеллектуальной системе поддержки работы оператора. Разработка такой перспективной СУ осуществляется на базе единого подхода и общих принципов построения АО, ИО и ПО с помощью общих моделей, типовых алгоритмов идентификации, адаптации, управления и методов ИИ, реализуемых в виде типовых проектных решений (ТПР). На базе единых ТПР разрабатывается также и система обучения операторов управления МПО, включающая в себя полную модель объекта, ТСУ, СУ и ИСПР.

Применение современных компьютерных технологий и ТПР на базе ИИ, а также реализация бортовых ПСУ на типовых программно-технических средствах в виде распределенной системы управления обеспечивают не только резкое сокращение сроков и затрат на создание ПСУ и СОО, но, самое главное, обеспечивают значительное повы-

шение эффективности и безопасности управления МПО. Существенная часть общего положительного эффекта обеспечивается путем применения ИСПР как в процессе управления МПО, так и при обучении операторов [1–5].

### 1. ОСОБЕННОСТИ РЕШЕНИЯ ПРОБЛЕМЫ БЕЗОПАСНОГО УПРАВЛЕНИЯ МПО

Анализ существующих подходов и методов создания и эксплуатации систем управления, СППР и систем обучения показывает [2, 3], что имеются значительные резервы повышения общего уровня безопасного управления МПО.

Существенное продвижение в решении данной проблемы возможно путем применения системного подхода к созданию единого комплекса ПСУ нового поколения, обеспечивающего на базе единого подхода и качественно новый уровень управления объектом и одновременно высокий уровень подготовки операторов для ПСУ. Поисковые исследования и НИР, направленные на разработку основных компонентов такого комплекса, были начаты в Институте проблем управления РАН в 1998 г. по инициативе и при постоянной поддержке И.В. Прангишвили. Результаты этих исследований, продолжающихся и в настоящее время, позволили выполнить совместно с проектно-конструкторскими организациями ряд НИР и ОКР по созданию и отработке отдельных компонентов прототипов ПСУ для конкретных объектов.

Применительно к созданию нового поколения ПСУ для МПО рассматривались различные подходы. Один из них связан с созданием методики и эффективных программно-технических средств для создания и отработки АО, ИО и ПО для ПСУ и, в частности, с разработкой и применением специальных стендов имитационного моделирования. Созданный в ИПУ РАН специализированный программно-технический комплекс [1] реализован в виде многомашинного сетевого компьютерного стенда имитационного моделирования. Стенд позволил автоматизировать отработку алгоритмического и информационного обеспечения для распределенных ПСУ и обеспечить автоматизацию полномасштабного моделирования одновременной работы основных подсистем ПСУ с использованием общей имитационной модели. Конфигурирование полной имитационной модели позволяет включать и исследовать различные варианты модели объекта в нормальных и аварийных режимах, отрабатывать варианты перспективных компонентов АО, ИО и ПО для СУ и ТСУ, варианты реализации СППР, СОО и моделировать различные внешние возмущения и траекторные угрозы.

В работах [2, 3] были рассмотрены и другие подходы, повышающие эффективность и безопас-





ность управления МПО благодаря применению методов искусственного интеллекта, ПСУ, интеллектуальных СППР и СОО, а также разработке и применению типовой подсистемы создания виртуальной реальности в дополнение к традиционным методам визуализации отдельных режимов управления МПО.

В настоящей работе рассмотрен подход к повышению эффективности и безопасности управления МПО путем применения современных компьютерных технологий, реализуемых на базе полномасштабной СИМ и обеспечивающих совместную обработку АО, ИО и ПО, предназначенных для СУ и ТСУ, для бортовых ИСПР, встроенных в СУ, а также современных тренажеров и систем обучения операторов.

## 2. РАЗРАБОТКА МЕТОДИКИ ПОВЫШЕНИЯ БЕЗОПАСНОСТИ УПРАВЛЕНИЯ МПО

В основу разработанной общей методики, реализуемой на базе методов системного анализа и единой компьютерной технологии, положены общие принципы и подходы к повышению общего уровня безопасного управления МПО. Методика учитывает специфику управления движением МПО и позволяет получить сопоставительные оценки эффективности работы ТСУ, СУ и операторов при возникновении типовых аварийных ситуаций и их сочетаний. Например, заклинка рулей при высокой скорости хода или пробойна в прочном корпусе могут приводить к наступлению других аварийных ситуаций (провал за допустимую глубину, запредельный крен или дифферент и др.). Аналогичные аварийные ситуации могут вызываться и при неправильных действиях операторов постов управления движением МПО.

Основное внимание уделено согласованной разработке вариантов АО, ИО и ПО для СУ и ТСУ, а также соответствующих вариантов построения СППР и СОО. Общая методика обеспечения безопасного управления реализует:

- получение текущих оценок параметров модели движения объекта и параметров альтернативных математических моделей неизменяемых возмущений, действующих на объект, с помощью интегрированной системы навигации и управления;
- генерацию, проверку и выбор наиболее вероятных гипотез о моделях, значениях их параметров, состоянии объекта, работоспособности ТСУ и оставшихся ресурсах управления;
- определение безопасного управления МПО с помощью принципа максимума Л.С. Понтрягина и учета текущих оценок параметров выбранных моделей;

– прогноз движения объекта и определение областей безопасного управления МПО с учетом статических и динамических траекторных угроз;

– визуализацию пространственного движения МПО, работы ТСУ и областей безопасного управления МПО с помощью компьютерных методов создания виртуальной реальности для операторов постов управления;

– интеллектуальную поддержку управления МПО и принятия решений операторами в штатных и нештатных ситуациях с помощью интеллектуальной СППР;

– построение эффективной системы обучения операторов постов управления МПО на основе изложенных принципов, единых математических моделей, типовых АО, ИО и ПО, общей системы визуализации и интеллектуальной поддержки, используемых при построении СУ, СППР и СОО.

Применение комплекса компьютерных технологий позволяет не только реализовать основные компоненты ПСУ на основе указанных принципов, но и дает возможность получить численные оценки эффективности применения различных алгоритмов и средств управления, значительно увеличить число исследованных вариантов использования ТСУ, АО, ИО и ПО, обеспечить углубленное тестирование функциональности и надежности работы ПО, а также значительно сократить затраты и сроки создания СУ и СОО.

При реализации общей КТ на базе имитационного исследовательского комплекса проверялись различные варианты построения и совместного применения компьютерных моделей, методик и компьютерных технологий, позволяющих реализовать отдельные задачи и принципы. Исследовались возможности применения различных универсальных и специализированных пакетов, прикладных программ и инструментальных сред типа MATLAB-SIMULINK и др. Однако, за исключением реализации отдельных программных компонентов, при построении имитационного исследовательского комплекса реального времени потребовался значительный объем программирования на языках высокого уровня.

Научные и методические вопросы реализации отдельных компонентов ПСУ рассматривались в ряде известных работ, а также в некоторых работах авторов. В частности, рассмотрен подход к определению многомерного оптимального управления МПО и построению допустимых безопасных управлений на основе принципа максимума Л.С. Понтрягина [2], рассмотрена структура ИСПР и особенности ее построения [3]. Допустимые безопасные управления определяются с учетом имеющихся ресурсов управления, данных системы навигации и наиболее вероятной модели внешних возмущений. Заданные и фактические траектории

движения МПО, работа ТСУ, области безопасного управления и имеющиеся ресурсы управления представляются операторам постов управления с помощью трехмерной системы виртуальной реальности. Алгоритмы управления, синтезируемые для ПСУ, реализуются и в СУ, и в ИСПР. Доступ к ИСПР обеспечивается с различных постов управления МПО.

Разработанная общая методика отработки АО, ИО и ПО позволяет обеспечить компьютерную поддержку реализации следующих основных этапов создания ПСУ:

- выявление и исследование основных предпосылок и причин возникновения и развития ситуаций, угрожающих безопасному управлению МПО, уточнение математических моделей самого МПО и моделей ТСУ по результатам ходовых испытаний или данным эксплуатации МПО, построение моделей, учитывающих отказы в аппаратно-программных средствах СУ и ТСУ, адаптация параметров текущих моделей внешних возмущений;

- определение оценок предельных возможностей МПО, ТСУ и СУ по обеспечению требований безопасности управления МПО при выполнении поставленных целей и учете остальных тактико-технических требований, включая возможные аварийные ситуации, отказы ТСУ, СУ и неправильные действия операторов;

- совместное определение согласованных требований и параметров МПО, ТСУ и прототипа АО, ИО и ПО для типовой СУ на этапе автоматизированного проектирования ТСУ и СУ МПО;

- исследование на стенде СИМ [1] эффективности совместной работы МПО, ТСУ и СУ на этапе создания и отработки рабочих проектов с использованием полных математических моделей и получением численных оценок для различных вариантов и ситуаций;

- определение требований на доработку СУ и разработку ИСПР и СОО на базе использования в них единых моделей МПО, ТСУ и внешних возмущений, а также рекомендаций по обеспечению эффективной поддержки действий операторов в экстремальных ситуациях;

- автоматизированное тестирование и испытания работы аппаратно-программных средств СУ при комплексном моделировании совместной работы МПО, ТСУ и СУ на этапе ввода объекта в эксплуатацию;

- разработка ИСПР, встроенных в СУ, и создание СОО на базе единых моделей МПО, ТСУ, СУ и моделей внешних возмущений [2, 3];

- доработка ТСУ, СУ, ИСПР и СОО по результатам ходовых испытаний или модернизация этих подсистем после эксплуатации МПО (с повторением всех или некоторых перечисленных пунктов).

### 3. МАТЕМАТИЧЕСКИЕ МОДЕЛИ, ИСПОЛЪЗУЕМЫЕ ПРИ ПОСТРОЕНИИ СИМ

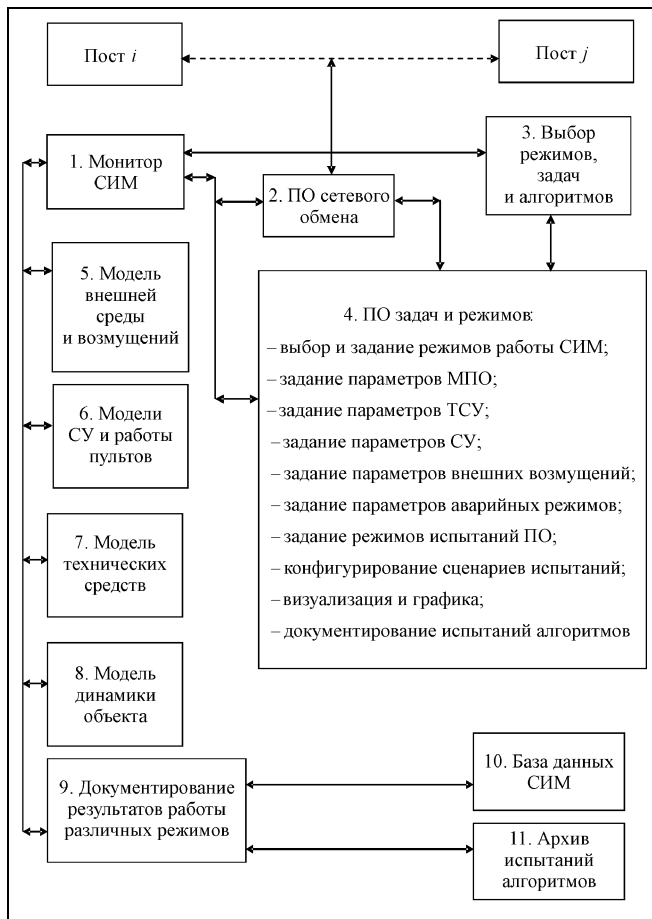
В качестве математической модели пространственного движения объекта принята система обыкновенных нелинейных дифференциальных уравнений 12-го порядка, представленная в нормальной форме. В общей модели имитации управления МПО, кроме модели движения МПО, учитываются модели работы ТСУ, модели системы управления и модели действия возмущений на объект. В общей модели учитывается сложная структура всех взаимосвязей между моделями. Модель зависимости тяги от задаваемого числа оборотов входит в уравнение движения в виде задания функции, имитирующей набор и сброс оборотов. Модели зависимостей сил и моментов учитывают влияние волнения моря, изменение плотности воды, влияние свободной поверхности и других возмущений и входят покомпонентно в уравнения сил и моментов общей имитационной модели.

Математические модели работы ТСУ описываются системами нелинейных дифференциальных уравнений до 6-го порядка для процессов, происходящих при функционировании различных исполнительных органов и видов арматуры (различных видов клапанов, кингстонов, насосов, цистерн и др.), а также включают в себя ограничения по скорости и положению приводов исполнительных органов, представляющие собой существенные нелинейности типа ограничений сверху, снизу и гистерезиса. В структуру типовой СУ входят также различные фильтры для оценивания неизмеряемых фазовых координат движения МПО. Общий порядок полной имитационной модели, реализуемой на СИМ, определяется моделью объекта, моделями СУ, моделями ТСУ и внешних возмущений и в зависимости от числа и типа исполнительных органов может быть более 40.

Рассмотрим кратко особенности реализации основных компонентов СИМ.

### 4. РЕАЛИЗАЦИЯ КОМПЬЮТЕРНОЙ ТЕХНОЛОГИИ НА СПЕЦИАЛИЗИРОВАННОЙ СИМ

На рисунке представлена общая структура основных блоков распределенной сетевой специализированной СИМ, включающей в себя шесть рабочих станций, сервер и инжиниринговую станцию. Штриховой линией условно показана связь ПО рабочих станций и ПО сервера по сети, а ниже показаны основные блоки ПО (блоки 1–11), реализующие соответствующие задачи на сервере и рабочих станциях. Имитация работы полной имитационной модели производится пошагово. Монитор (блок 1) системы имитационного моделирования синхронизирует на каждом шаге обмен данными между сервером и отдельными задачами на



Структурная схема ПО системы имитационного моделирования

рабочих станциях через дополнительное ПО обеспечения режимов сетевого обмена (блок 2). При формировании сигналов с сервера используются также следующие имитационные модели: — динамики объекта (блок 8), внешней среды и возмущений (блок 5), работы технических средств (блок 7).

В ПО задач, реализуемых на каждом шаге имитации на отдельных рабочих станциях в различных режимах, входят следующие блоки: 2 — дополнительное ПО поддержки режимов сетевого обмена; 4 — ПО конфигурирования модулей, задач и режимов; 6 — имитации работы системы управления и пультов; 9 — текущего документирования работы режимов; 10 — база исходных данных СИМ по объекту, ТСУ, возмущениям и алгоритмам систем управления; 11 — архив испытаний алгоритмов.

В прикладное ПО инжиниринговой станции входят блок 2 и блок 3 — выбора режимов, задач, сценариев испытаний и формирования необходимых параметров перед началом испытания работы заданных последовательностей алгоритмов и режимов на стенде. После проведения испытания режима это ПО формирует данные для блоков 10

и 11 (на сервере), получая и обрабатывая накопленные данные из блоков 9 рабочих станций.

Возможны различные варианты работы стенда — от имитации автономной работы отдельных задач и режимов, реализуемых на различных постах управления, до имитации комплексных режимов взаимодействия работы алгоритмов, ТСУ и различных постов управления с обменом по сети всеми необходимыми данными между ними.

## 5. ПРИМЕРЫ ДИНАМИЧЕСКОГО ПОСТРОЕНИЯ И ОТОБРАЖЕНИЯ ВИРТУАЛЬНОЙ РЕАЛЬНОСТИ ДЛЯ ОПЕРАТОРА УПРАВЛЕНИЯ ДВИЖЕНИЕМ МПО

На основе проведенных исследований и применения КТ был разработан прототип АО, ИО и ПО основных подсистем ПСУ [4, 5]. На рисунках (см. цветную вклейку) представлены примеры видеокладов виртуального представления операторам постов управления движением МПО текущей и прогнозируемой ситуации, состояния ТСУ и движения объекта [4, 5]. Для существенного улучшения восприятия и оценки операторами поступающей концентрированной оперативной информации при построении подсистемы виртуальных отображений применены принципы и методы инженерной психологии. Оператор осуществляет пространственное координированное ручное управление МПО по курсу и глубине с помощью четырех ТСУ, управляя положением горизонтальных кормовых и носовых рулей, вертикальными рулями и скоростью хода (см. верхний рисунок). На нижнем рисунке приведен пример видеоклада при автоматическом управлении МПО с представлением оператору предсказанного положения объекта, ограничений по управлению и выполнения требований обеспечения траекторной безопасности.

## ЗАКЛЮЧЕНИЕ

Применение единой компьютерной технологии при разработке систем управления, интеллектуальных систем поддержки работы операторов, тренажеров и систем обучения операторов позволяет не только создать программно-технические средства нового поколения, но и гарантировать высокий уровень подготовки операторов, необходимый для эффективной эксплуатации этих средств. В конечном итоге все это и обеспечивает безопасное, безаварийное и эффективное управление морскими подвижными объектами, особенно в нестандартных и чрезвычайных ситуациях;

Повышение качества разработки указанных систем обеспечивается благодаря типизации и стандартизации проектных решений в различных подсистемах, применению общих моделей и методов, конфигурированию и параметрической на-

стройке программного обеспечения для различных классов подвижных объектов.

Применение разработанной компьютерной технологии и типовых проектных решений обеспечивает резкое сокращение сроков разработок и затрат на создание и модификацию отдельных компонентов перспективных систем управления и систем обучения операторов.

Дальнейшие публикации о системе имитационного моделирования предполагаются в последующих номерах журнала.

## ЛИТЕРАТУРА

1. *Построение систем координированного управления движением подводных объектов на базе сетевых средств* / Борисов В.Г. и др. // Тр. междунар. конф. «Моринтех — 2001». — СПб., 2001.
2. *Борисов В.Г., Данилова С.К., Чинакал В.О.* Интеллектуальная система поддержки принятия решений для интегриро-

ванных систем управления и навигации морскими подвижными объектами // Тр. XI Санкт-Петербургской междунар. конф. по интегрированным навигационным системам. — СПб.: Электроприбор, 2004.

3. *Борисов В.Г., Данилова С.К., Чинакал В.О.* О построении интегрированных систем навигации, управления и обучения операторов на базе методов интеллектуального управления // Тр. XII Санкт-Петербургской междунар. конф. по интегрированным навигационным системам. — СПб.: Электроприбор, 2005.
4. *Международный военно-морской салон IMDS — 2003.* Компьютерный понятийный тренажер для обучения операторов управления морскими подвижными объектами. — СПб. 2003.
5. *Пятая Международная конференция и выставка по морским интеллектуальным технологиям «Моринтех — 2003».* Интеллектуальные технологии в эксплуатации кораблей и судов (системы обработки информации и управление). — СПб.: Моринтех, 2003.

☎ (495) 334-92-40, e-mail: lab57\_1@ipu.ru

Статья представлена к публикации членом редколлегии Б.Г. Воликом. □

УДК 312.1:444

# СИНТЕЗ АЛГОРИТМА ДЛЯ ОДНОЙ ЗАДАЧИ УПРАВЛЕНИЯ МОРСКИМИ ПОДВИЖНЫМИ ОБЪЕКТАМИ

Л.В. Евстигнеева, Н.Н. Тарасов, М.Г. Тахтамышев

*Институт проблем управления им. В.А. Трапезникова РАН, г. Москва*

Предложен алгоритм фильтрации и восстановления неизмеряемых координат для решения задачи синтеза алгоритма управления морскими подвижными объектами в режиме «без хода», позволяющий существенно повысить качество процессов управления.

## ВВЕДЕНИЕ

Одна из задач, стоящих перед проектировщиками систем управления движением морских подвижных объектов (МПО), заключается в обеспечении стабилизации МПО в режиме «без хода». Традиционно разработчики таких систем применяют структуры линейных регуляторов, но если в режимах движения при наличии скорости хода требуемое качество процессов обеспечивается, как

правило, с помощью ПД и ПИД-регуляторов, то для стабилизации МПО в режиме «без хода» такие структуры оказываются неэффективными, поскольку не позволяют получить необходимого запаса устойчивости системы. Для решения этой задачи в существующих системах управления МПО в законы управления вводятся высшие производные регулируемых координат с большими коэффициентами усиления. Однако при такой структуре регулятора, как показывает практика, существенно



увеличивается уровень шумов в канале управления, что, в свою очередь, приводит к ухудшению качества процессов автоматического управления.

В настоящее время при проектировании систем управления широкое применение находят подходы, основанные на применении методов фильтрации и восстановления координат. Реализация таких подходов позволяет не только снижать уровень помех измеряемых координат, но и восстанавливать неизмеряемые координаты, используемые в законах управления. Однако реализация таких фильтров на бортовых комплексах часто вызывает затруднения из-за ограниченной производительности бортовых вычислительных машин.

В настоящей статье рассматривается структура регулятора, в которой вместо высших производных используются оценки координат управляющих воздействий. Как показали результаты исследований, такой способ реализации не требует больших коэффициентов усиления, что в итоге приводит к существенному снижению уровня помех и повышению качества процессов управления.

### 1. МАТЕМАТИЧЕСКАЯ МОДЕЛЬ ОБЪЕКТА

Объект управления описывается системой нелинейных дифференциальных уравнений вида

$$\begin{aligned} \dot{V}_y &= a_1 V_y |V_y| + b_1 u_c, \\ \dot{\eta} &= V_y, \\ \dot{\omega}_z &= a_2 |V_y| \omega_z + b_2 (l_1 + l_2) u_{c1}, \\ \dot{\psi} &= \omega_z \end{aligned} \quad (1)$$

где  $a_i$  — известные параметры,  $V_y$ ,  $\omega_z$ ,  $\psi$  и  $\eta$  — координаты движения,  $u_c$  и  $u_{c1}$  — управляющие воздействия. Производные  $\dot{u}_i$  управляющих воздействий представляют собой нелинейные функции, вид которых показан на рис. 1, где  $\sigma_i$  — алгоритмы управления:  $\sigma_c = f(\eta, \dot{\eta}, u_c)$ ,  $\sigma_{c1} = f(\psi, \dot{\psi}, u_{c1})$ .

В качестве измерителей служат датчики величин  $\psi$ ,  $\eta$  и  $\dot{\eta}$  (так как для рассматриваемой задачи  $\dot{\eta} = V_y$ , в дальнейшем под измеряемой величиной подразумевается координата  $V_y$ ). Уравнения, описывающие процесс измерения, с достаточной степенью точности могут быть представлены в виде

$$\begin{aligned} \psi_{\text{изм}} &= y_\psi = \psi + \xi_\psi \\ \eta_{\text{изм}} &= y_\eta = \eta + \xi_\eta, \\ V_{\text{изм}} &= y_{V_y} = V_y + \xi_{V_y}, \end{aligned} \quad (2)$$

где  $y_\psi$ ,  $y_\eta$ ,  $y_{V_y}$  — измеренные значения параметров  $\psi$ ,  $\eta$  и  $V_y$ , а  $\xi_\psi$ ,  $\xi_\eta$  и  $\xi_{V_y}$  — случайные величины, характеризующие погрешность измерения.

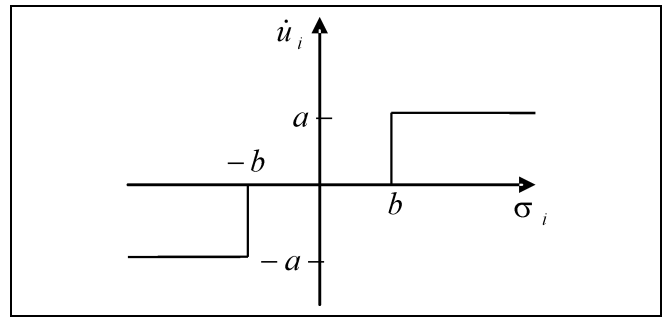


Рис. 1. Вид производных управляющих воздействий

Из системы уравнения (1) видно, что с помощью управляющего воздействия  $u_c$  осуществляется управление МПО по глубине  $\eta$ , а с помощью воздействия  $u_{c1}$  — стабилизация по дифференту  $\psi$ .

### 2. АЛГОРИТМ УПРАВЛЕНИЯ

Традиционно применяемая структура регулятора, как правило, задается в виде

$$\begin{aligned} \sigma_c &= -k_{\Delta\eta} \Delta\eta - k_{\dot{\eta}} \dot{\eta} - k_{\ddot{\eta}} \ddot{\eta}, \\ \sigma_{c1} &= -k_\psi \Delta\psi - k_{\dot{\psi}} \dot{\psi} - k_{\ddot{\psi}} \ddot{\psi}, \end{aligned} \quad (3)$$

где  $\Delta\eta = \eta - \eta_{\text{зад}}$ ,  $k_i$  — постоянные коэффициенты. Анализ показывает, что при такой структуре регулятора требуются очень большие коэффициенты, особенно при высших производных. Реализация такой структуры затруднительна даже в том случае, когда все координаты движения  $\eta$ ,  $\dot{\eta}$ ,  $\ddot{\eta}$ ,  $\psi$ ,  $\dot{\psi}$ ,  $\ddot{\psi}$ , измеряются с достаточно высокой точностью. Однако измерению доступны лишь координаты  $\eta$ ,  $\dot{\eta}$  и  $\psi$ , причем с довольно высоким уровнем шумов. Поэтому получение их оценок, а тем более оценок их производных с требуемой точностью, практически невозможно. Таким образом, использование законов управления с большими коэффициентами усиления приводит к существенным погрешностям в управлении реальными объектами.

На рис. 2 представлены результаты моделирования процессов переходов по глубине для одного МПО, где символом  $\bullet$  обозначен процесс перехода при отсутствии ошибок измерений ( $\sigma_{Q_1} = -k_{\Delta\eta} \Delta\eta - k_{\dot{\eta}} \dot{\eta} - k_{\ddot{\eta}} \ddot{\eta}$ ), а символом  $\blacktriangle$  — процесс перехода при наличии реальных помех по координатам  $\dot{\eta}$  и  $\ddot{\eta}$ :  $\sigma_{Q_1} = -k_{\Delta\eta} \Delta\eta - k_{\dot{\eta}} V_{\text{изм}} - k_{\ddot{\eta}} (\ddot{\eta} + \xi_{\ddot{\eta}})$ ; для оценки влияния помех по каждой из координат приведен также процесс перехода (символ  $\blacksquare$ ) при наличии помех только по координате  $\dot{\eta}$ :  $\sigma_{Q_1} = -k_{\Delta\eta} \Delta\eta - k_{\dot{\eta}} V_{\text{изм}} - k_{\ddot{\eta}} \ddot{\eta}$ .

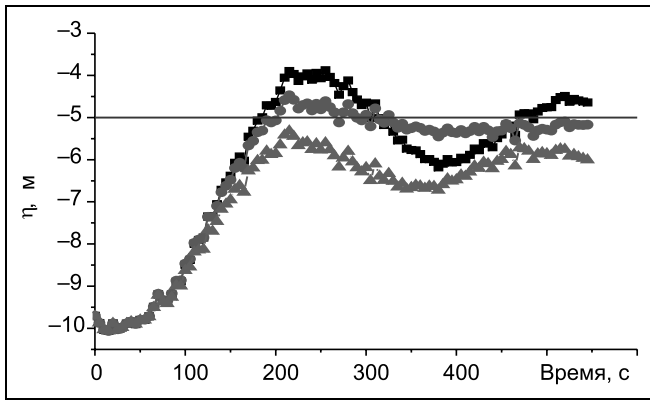


Рис. 2. Результаты моделирования процессов перехода (алгоритм (3))

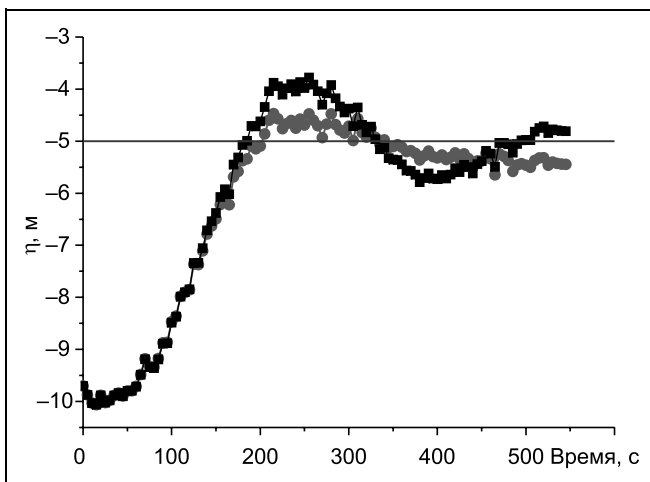


Рис. 3. Результаты моделирования процессов перехода (алгоритм (4))

Отметим, что для рассматриваемых МПО критерием качества процессов перехода по глубине служит минимум времени перехода, при условии, что перерегулирование не должно превышать 10 % от уровня перехода.

Алгоритм  $\sigma_{Q_1} = -k_{\Delta\eta}\Delta\eta - k_{\dot{\eta}}\dot{\eta} - k_{\ddot{\eta}}\ddot{\eta}$  обеспечивает требуемое качество управления только в «идеальном» случае, при отсутствии ошибок измерения. Учет реальных ошибок измерения (около 10 % от измеряемой величины) показывает, что данная структура алгоритма не позволяет обеспечить требуемое качество процессов управления.

Рассмотрим предлагаемую структуру алгоритма управления, не включающую в себя производные координат второго порядка:

$$\begin{aligned}\sigma_{Q_1} &= -k_{\Delta\eta}\Delta\eta - k_{\dot{\eta}}\dot{\eta} - k_{u_c}u_c, \\ \sigma_{Q_2} &= -k_{\Delta\psi}\Delta\psi - k_{\dot{\psi}}\dot{\psi} - k_{u_{c1}}u_{c1}.\end{aligned}\quad (4)$$

Реализация этой структуры, как показали результаты анализа для одного класса МПО, требует существенно меньших значений коэффициентов  $k_{u_c}$  по сравнению с коэффициентами  $k_{\dot{\eta}}$  и  $k_{\ddot{\eta}}$  в структуре алгоритма (3).

На рис. 3 представлены результаты моделирования переходов того же МПО по глубине в соответствии с алгоритмом (4): символом  $\bullet$  обозначен процесс перехода при отсутствии ошибок измерений ( $\sigma_{Q_1} = -k_{\Delta\eta}\Delta\eta - k_{\dot{\eta}}\dot{\eta} - k_{u_c}u_c$ ), а символом  $\blacksquare$  — процесс перехода при наличии реальных помех 10 % помех по координатам  $\dot{\eta}$  и  $u_c$ :  $\sigma_{Q_1} = -k_{\Delta\eta}\Delta\eta - k_{\dot{\eta}}V_{\text{уизм}} - k_{u_c}(u_c + \xi_{u_c})$ .

При отсутствии ошибок измерений алгоритм (4), как и алгоритм (3), обеспечивает хорошее качество управления. Ошибки измерения при реализации алгоритма (4) сказываются на качестве процесса управления значительно в меньшей степени, чем при реализации алгоритма (3).

Однако реализация алгоритма (4) связана с необходимостью восстановления неизмеряемой координаты  $u_c$ . Поэтому предложен алгоритм, использующий оценки как измеряемых, так и неизмеряемых координат движения:

$$\begin{aligned}\sigma_{Q_1} &= -k_{\Delta\eta}\Delta\hat{\eta} - k_{\dot{\eta}}\dot{\hat{\eta}} - k_{u_c}\hat{u}_c, \\ \sigma_{Q_2} &= -k_{\Delta\psi}\Delta\hat{\psi} - k_{\dot{\psi}}\dot{\hat{\psi}} - k_{u_{c1}}\hat{u}_{c1},\end{aligned}\quad (5)$$

где  $\Delta\hat{\eta} = \hat{\eta} - \eta_3$ .

Для получения оценок координат, входящих в алгоритм (5), возникает необходимость в решении задачи фильтрации и восстановления [1].

### 3. АЛГОРИТМ ФИЛЬТРАЦИИ И ВОССТАНОВЛЕНИЯ КООРДИНАТ

Реализация фильтра для нелинейной модели объекта (1) требует большого объема вычислений. Это связано с необходимостью решения в реальном времени уравнения Риккати размерности  $n(n+1)/2$ , что предъявляет высокие требования к характеристикам бортовых вычислительных комплексов.

Ниже приводится способ решения задачи фильтрации и восстановления координат, не требующий для реализации решения системы уравнений большой размерности в реальном времени. Из системы уравнений (2) можно получить  $V_y = y_{V_{\text{изм}}} - \xi_{V_{\text{изм}}}$ , или  $V_y = y_{V_{\text{изм}}} - \xi_{V_{\text{изм}}} \approx y_{V_{\text{изм}}}$ .



Подставляя это выражение в модель движения, получим:

$$\begin{aligned} \dot{V}_y &= a_1 V_y |y_{V_{\text{изм}}}| + b_1 u_c, \\ \dot{\eta} &= V_y, \\ \dot{u}_c &= q_1. \end{aligned} \quad (6)$$

Легко проверить, что система (2), (6) является наблюдаемой, и параметры, необходимые для формирования закона управления, можно получить с помощью алгоритма фильтрации вида [2]:

$$\begin{aligned} \hat{V}_y &= a_1 \hat{V}_y |y_{V_{\text{изм}}}| + b_1 \hat{u}_c + K_V (y_\eta - \hat{\eta}), \\ \hat{\eta} &= \hat{V}_y + K_V (y_\eta - \hat{\eta}), \\ \hat{u}_c &= q_1 + K_V (y_\eta - \hat{\eta}). \end{aligned} \quad (7)$$

Коэффициенты усиления  $K = PC^T R^{-1} = (K_{V_c}, K_\eta, K_{u_c})^T$  в цепи обратной связи определяются, как правило, из уравнения Риккати  $\dot{P} = AP + PA^T - PC^T R^{-1} CP + Q$ ,

$$\text{где } A = \begin{pmatrix} a_1 |y_{V_y}| & 0 & b_1 \\ 1 & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix}; C = (0 \ 1 \ 0); Q = \begin{pmatrix} q_1 & 0 & 0 \\ 0 & q_2 & 0 \\ 0 & 0 & q_3 \end{pmatrix};$$

$R = r$ .

Для реализации фильтра (7) не требуется решение уравнений Риккати в реальном времени, и коэффициенты  $K_i$  могут быть просчитаны заранее [3].

На рис. 4 приведены процессы переходов по глубине для системы управления МПО, в соответствии с алгоритмом (5) и фильтром (7) с посто-

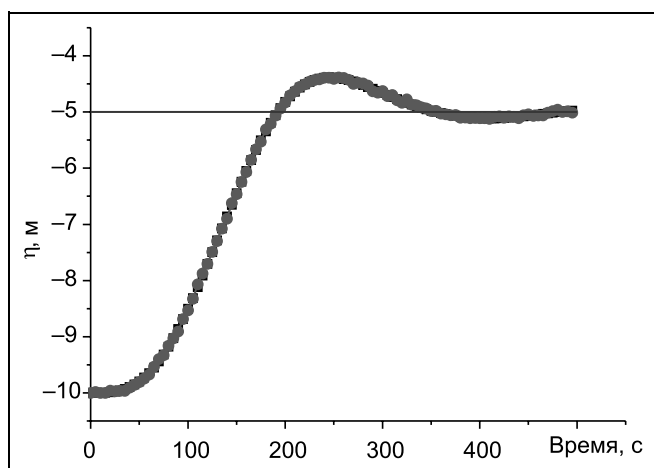


Рис. 4. Результаты моделирования процессов перехода (алгоритм (5) и диаметр (7))

янными коэффициентами  $K_i$ . Символом  $\bullet$  обозначен процесс перехода при отсутствии ошибок измерений, а символом  $\blacksquare$  — процесс перехода при наличии реальных помех измерения (10 %).

Как видно, эти процессы практически совпадают, и качество процессов в обоих случаях соответствует требуемому (минимум времени перехода при условии 10 %-й трубки по перерегулированию). Таким образом, если алгоритм (3) и алгоритм (4), в условиях реально действующих шумов измерений, не могут обеспечить требуемое качество процессов управления, то применение алгоритма (5) совместно фильтром (7) позволяет решить поставленную задачу.

## ЗАКЛЮЧЕНИЕ

Рассмотрена задача синтеза алгоритма управления морскими подвижными объектами в режиме «без хода». Для оценки текущего состояния объекта применена структура фильтра Калмана, что позволяет получить оценки неизмеряемых координат. С помощью полученных оценок реализован алгоритм управления МПО, обеспечивающий высокое качество процессов управления благодаря существенному снижению уровня шумов по сравнению с традиционно применяемым алгоритмом. Известно, что реализация фильтра Калмана в бортовых системах управления вызывает затруднения из-за необходимости решать уравнения Риккати в реальном времени. Предложенный подход позволяет избежать этих затруднений. Этот подход предполагает:

- приведение нелинейной системы дифференциальных уравнений, описывающих движение МПО, к линейной системе с переменными коэффициентами;
- определение постоянного (для данного объекта) набора коэффициентов усиления при невязках путем решения уравнений Риккати для линейной системы;
- возможность корректировки полученного набора постоянных коэффициентов с учетом реальных характеристик процессов управления.

## ЛИТЕРАТУРА

1. *Фильтрация и стохастическое управление в динамических системах* / Под ред. К.Т. Леондеса. — М.: Мир, 1980.
2. *Воронов А.А. Устойчивость, управляемость, наблюдаемость*. — М.: Наука, 1979.
3. *Венгеров А.А., Шаренский В.А. Прикладные вопросы оптимальной линейной фильтрации*. — М.: Энергоиздат, 1982.

☎ (495) 334-92-20, e-mail: takhnamv@ipu.ru

Статья представлена к публикации членом редколлегии Б.Г. Воликом. □



# CONTENTS & ABSTRACTS

## DYNAMICS MODEL IN AGING SYSTEMS CONTROL PROBLEMS . . . . . 2

**Baranov V.V., Matrosov V.M.**

A model of markovian process's transient function is built that describes a controlled aging process. The model can be used in aging systems efficiency and safety control.

## ESTIMATION OF STATE VECTOR COMPONENTS IN A NONLINEAR SINGULAR SYSTEM . . . . . 8

**Rykin S.E.**

The sufficient observability condition is formulated and proved for a class of nonlinear dynamic systems linear with respect to the estimated subvector of the full state vector. Nonlinear estimation techniques for state variable components based on slide modes are offered. The method's efficiency is shown by the example of the estimation algorithm synthesis for mechanical variables of a synchronous motor.

## DISTRIBUTION LAWS OF DYNAMIC OBJECTS' PARAMETER ESTIMATES IN THE IDENTIFICATION BY EXPONENTIAL MODULATION METHOD . . . . . 14

**Anisimov D.N., Khripkov A.V.**

A new method of dynamic objects identification named the Exponential Modulation Method is considered. The paper shows that the method can be easily realized and features high noise immunity. The problems related with the selection of optimal input signal are discussed. The distribution laws of the estimates under the disturbances such as white noise and sine wave with random phase are determined.

## INTEGRATION AND GLOBALIZATION OF DATA PROCESSING SOFTWARE . . . . . 19

**Alexeevsky M.V.**

An effective method for solving the problem of globalization and integration of separate corporate software into in unified information system is discussed. The method presumes the transition to the three-tier network architecture Client → Provider → Server using the software toolkits of COBRA/KORINF technology.

## RANKING OF WEB PAGES USING MUTUAL INFORMATION BETWEEN THE HYPERLINKS . . . . . 24

**Alguliev R.M., Alyguliev R.M.**

To improve Web pages ranking efficiency, the paper offers 3 modifications of PageRank algorithm. The first one features the measurement of hyperlink independence degree for calculating the hyperlink's weight. The second and third ones considering the thematic affinity of Web pages are the improved versions of algorithms WPR and Topic-Centric respectively.

## PS-2000 MULTIPROCESSOR ARCHITECTURE ON A CHIP . . . 30

**Zatuliveter Yu.S., Fischenko E.A.**

The contemporary problems of building multiprocessor architectures on a chip are examined. The paper suggests to develop the architecture well proven in domestic PS-2000 multiprocessor computing system for the realization on a chip.

## CLUSTER ANALYSIS APPLICATION TO AUTOMATIC FLAW DETECTION . . . . . 36

**Vasiliev I.L., Sidorov D.N.**

The problem of automatic detection and categorization of flaws that arises in machine vision system development is discussed. A method of its solution is proposed. At the first stage, an autonomous process of system learning is performed where the learning samples of the analyzed flaw symptoms are split into clusters with medians. At the second stage, the real-time inspection process is performed where each incoming pattern is categorized according to the nearest median type. The paper notes that the comparative analysis against other techniques showed the effectiveness of the method proposed.

## EVALUATION OF DEPARTMENTAL TARGET PROGRAMS CONTRIBUTION TO THE ACCOMPLISHMENT OF GOVERNMENTAL GOALS . . . . . 43

**Chebotaurov P.Yu., Mit'kin A.N., Shmerling D.S.**

A numerical-expert procedure for evaluating the contribution of departmental target programs into the accomplishment of Russian Federa-

tion government's goals is offered for developing the list of such programs financed from the national budget.

## PROFESSIONAL WORK OBJECTS' CLASSIFICATION FOR DESIGNING PROFESSIONAL AND EDUCATIONAL STANDARDS . . . . . 51

**Nikitin V.V., Mal'tseva S.B., Dorofeyuk A.A., Mandel' A.S., Chernyavsky A.L.**

The paper shows that the formation of generic notions describing the area of specialist's professional activities can be reduced to the problem of specific coupling matrix diagonalization. A human-computer solution procedure is offered.

## AN APPROACH TO INNOVATION PROJECT SUCCESS FORECASTING . . . . . 56

**Vinokur V.M., Myl'nikov L.A., Perminova N.V.**

A hybrid approach to forecasting innovative projects at different stages of their implementation is discussed. The technique enables both near-term and long-term forecast precision improvement subject to competent management of the innovative project under investigation.

## COST-EFFECTIVENESS OF CONTROL SYSTEMS . . . . . 60

**Volik B.G.**

The criteria for selecting the preferable variant of control system on the basis of the cost-effectiveness analysis of competing system variants are offered.

## MATHEMATICAL FORMALIZATION OF HIERARCHICAL CONTROL METHODS FOR ENVIRONMENTAL-ECONOMIC SYSTEMS . . . . . 64

**Ougolnitsky G.A., Usov A.B.**

The paper discusses general approaches to the control of environmental-economic systems. The methods of hierarchical control (compulsion, impetus, and persuasion) are offered that allow the upper level control subject to ensure the fulfillment of system's sustainable development requirements. Application examples for a river water quality control system are adduced, and the comparative analysis of different control methods is included.

## AN ALGORITHM FOR FAST VERIFICATION OF A DANGEROUS EPIDEMIC DISEASE USING PROBABILITY TECHNIQUE AND ITS IMPLEMENTATION . . . . . 70

**Mikheev O.V., Samokhina A.S.**

An algorithm for a decision-making support system intended for fast verification of a dangerous epidemic disease is developed. The system can be used for prevention and elimination of the consequences of biological attacks.

## THE MODELS OF TRAINING QUALITY MANAGEMENT SYSTEM USING INNOVATIVE EDUCATIONAL TECHNOLOGY . . . . . 74

**Belyakov A.Yu., Puisans S.G., Stolbova I.D., Kharitonov V.A.**

Two models of quality management system for specialist training on the level of a separate program using innovative educational technology are considered.

## DEVELOPMENT AND APPLICATION OF COMPUTER TECHNOLOGY FOR IMPROVING THE SAFETY OF MOVABLE NAUTICAL OBJECTS CONTROL . . . . . 79

**Borisov V.G., Danilova S.K., Chinakal V.O.**

The paper discusses the primary ways to improve the safety of movable nautical objects control based on the unified through computer-based technology for control hardware and systems development support, control algorithms debugging using the full mathematical model of object's dynamics and various control algorithms for both normal and abnormal operation.

## CONTROL ALGORITHM SYNTHESIS FOR A MOVABLE NAUTICAL OBJECT: A CASE STUDY . . . . . 84

**Yevstigneeva L.V., Tarasov N.N., Takhtamyshv M.G.**

The paper presents a filtering and nonmeasurable coordinates evaluation algorithm for synthesizing the control algorithm of movable nautical objects in "no motion" mode. The new algorithm enables significant improvement of control performance.