

МИНИСТЕРСТВО ОБРАЗОВАНИЯ И НАУКИ УКРАИНЫ

ХАРЬКОВСКИЙ НАЦИОНАЛЬНЫЙ УНИВЕРСИТЕТ
РАДИОЭЛЕКТРОНИКИ

ISSN 0555-2656

БИОНИКА
ИНТЕЛЛЕКТА
ИНФОРМАЦИЯ, ЯЗЫК, ИНТЕЛЛЕКТ

№ 3 (74)

2010

НАУЧНО-ТЕХНИЧЕСКИЙ ЖУРНАЛ

Основан в 1967 г.

Свидетельство о государственной регистрации КВ № 12072-943 ПР от 07.12.2006

Журнал включен в список специальных изданий ВАК Украины
по техническим наукам
(постановление президиума ВАК Украины № 1-05/6 от 16.12.2009)

СОДЕРЖАНИЕ

<i>Бондаренко М. Ф., Кругликова Н. П., Лещинская И. А., Русакова Н. Е., Шабанов-Кушнарченко Ю. П.</i> Об алгебре предикатов	3
<i>Бондаренко М. Ф., Кругликова Н. П., Лещинская И. А., Русакова Н. Е., Шабанов-Кушнарченко Ю. П.</i> О реляционных сетях	8
<i>Вечирская И. Д., Валенда Н. А.</i> Логическое преобразование как средство реализации сетевых технологий	14
<i>Маслянюк П. П., Рябушенко А. В., Козленко М. В.</i> Компонентна модель оцінювання вартості похідних фінансових інструментів на фондовому ринку	19
<i>Петров Э. Г., Губаренко Е. В.</i> Методы и инструментальные средства систем поддержки принятия решений при организационном управлении социально-экономическими системами	26
<i>Бескоровайный В. В., Зыкова О. В.</i> Комплексирование средств поддержки принятия решений при синтезе территориально распределенных объектов	37
<i>Калмыков А. В.</i> Метод исследования процессов принятия решений в госрегулировании телекоммуникационной отрасли	42
<i>Алексієв О. П., Алексієв В. О., Хабаров В. О., Четвериков Г. Г.</i> Інформатизація транспортної інфраструктури, машин та систем	52
<i>Гороховатский В. А.</i> Структурноиерархические методы распознавания изображений при влиянии пространственных помех	58
<i>Полякова М. В., Крылов В. Н., Гуляева Н. А.</i> Оценка качества подчеркивающих преобразований для задачи сегментации изображений	62
<i>Зайцев С. А., Субботин С. А.</i> Кластерный анализ с использованием гибридной модели на основе искусственной иммунной сети	70
<i>Котов А. С.</i> Разработка комплексной технологической модели изделия в виде семантической сети	76
<i>Олейник Ан. А., Ключкова А. А., Олейник Ал. А.</i> Агентные технологии в системах групповой видеокommunikации	83
<i>Шергин В. Л., Мирошниченко Т. А.</i> Робастные методы оценивания числовых характеристик выборки	90
<i>Новосельцев И. В., Аксак Н. Г.</i> Система первичной диагностики меланомы кожи	94
<i>Аврунин О. Г., Жемчужкина Т. В., Носова Т. В.</i> Диагностические возможности электромиографического метода при исследовании функции носового клапана	99
<i>Половенко К. Г.</i> Применение метода вейвлет-анализа для определения аномальных участков на электроэнцефалограмме человека	105
<i>Губин В. А.</i> Слабоструктурированные текстовые документы как источники данных	109
<i>Танянский С. С.</i> Поддержка целостности баз данных при динамических структурных изменениях	112
<i>Кашубин С. П.</i> Логика секвенциальных схем алгоритмов	120
<i>Довбиш А. С., Мартиненко С. С.</i> Оптимізація параметрів плану навчання системи розпізнавання магнітокардіограм	129
<i>Лесна Н. С., Келеберда І. М., Шрестха С. М., Кирпунь Ю. А.</i> Сервіс-орієнтована модель персоналізованого навчання з використанням агентних технологій	133
<i>Лесна Н. С., Найдено А. В., Сокол В. В.</i> Інформаційна технологія підтримки консультування студентів дистанційної форми навчання	138
<i>Кузмин А. Я., Даюб Я.</i> Муравьиный алгоритм и учет риска в транспортной задаче	144
ОБ АВТОРАХ	149
Правила оформления рукописей для авторов научно-технического журнала «Бионика интеллекта»	152

УДК 519.7



ОБ АЛГЕБРЕ ПРЕДИКАТОВ

М. Ф. Бондаренко¹, Н. П. Кругликова², И. А. Лещинская³,
Н. Е. Русакова⁴, Ю. П. Шабанов-Кушнаренко⁵

¹⁻⁵ ХНУРЭ, г. Харьков, Украина

В статье рассматривается алгебра логических элементов, называемая алгеброй логики. Обобщение этой алгебры приводит к понятию алгебры Буля. Рассматривается алгебра конечных предикатов как частный случай алгебры Буля. В алгебре конечных предикатов вводится понятие скобочной формы, которое рекомендуется для практического применения в качестве средства экономного представления конечных отношений.

АЛГЕБРА ЛОГИКИ, АЛГЕБРА БУЛЯ, АЛГЕБРА ПРЕДИКАТОВ, СКОБОЧНАЯ ФОРМА

Введение

Аксиомам алгебры Буля удовлетворяют некоторые математические структуры, имеющие большое значение для теории интеллекта. К таким структурам относятся, например, алгебра логики, алгебра множеств, алгебра конечных предикатов. Обобщив алгебру логики, можно получить алгебру Буля. Разновидностью алгебры Буля является алгебра конечных предикатов. Целью статьи является введение в алгебре конечных предикатов понятия скобочной формы предиката, которую можно получить, применяя всюду, где только это возможно, операцию группировки предикатов узнавания предмета.

1. Алгебра логики

Рассмотрим множество $\Sigma = \{0,1\}$, состоящее из элементов 0 и 1, называемых *логическими*. Введем *логические переменные* $\xi, \eta \in \Sigma$. На множестве Σ определяем *логические операции дизъюнкции* $\xi \vee \eta$, *конъюнкции* $\xi \wedge \eta = \xi \cdot \eta = \xi \eta$ и *отрицания* $\neg \xi = \bar{\xi}$ следующими равенствами: $0 \vee 0 = 0, 0 \vee 1 = 1 \vee 0 = 1 \vee 1 = 1; 0 \wedge 0 = 0 \wedge 1 = 1 \wedge 0 = 0, 1 \wedge 1 = 1; \neg 0 = 1, \neg 1 = 0$. Система, состоящая из множества Σ и заданных на нем операций \vee, \wedge, \neg , называется *алгеброй логики*. Множество Σ играет роль *носителя* алгебры логики, а логические операции \vee, \wedge, \neg — роль ее *базисных операций*.

Мы определили алгебру логики *конструктивно*, задав ее базисные операции \vee, \wedge, \neg на носителе Σ *прямым определением*. Теперь введем эту же алгебру еще и *абстрактно косвенным определением*, задав ее базисные операции их свойствами с помощью перечисленных ниже *аксиом*, называемых *законами логики*. Свойства операций \vee, \wedge, \neg записываем в виде равенств, связывающих логические переменные ξ, η, ζ . Имеется в виду, что каждое равенство должно выполняться при любых $\xi, \eta, \zeta \in \Sigma$.

Формулируем *законы идемпотентности*

$$\xi \vee \xi = \xi, \xi \wedge \xi = \xi;$$

коммутативности

$$\xi \vee \eta = \eta \vee \xi, \xi \wedge \eta = \eta \wedge \xi;$$

ассоциативности

$$(\xi \vee \eta) \vee \zeta = \xi \vee (\eta \vee \zeta), (\xi \wedge \eta) \wedge \zeta = \xi \wedge (\eta \wedge \zeta);$$

дистрибутивности

$$(\xi \vee \eta) \zeta = \xi \zeta \vee \eta \zeta, \xi \eta \vee \zeta = (\xi \vee \zeta)(\eta \vee \zeta);$$

элиминации

$$\xi \vee \xi \eta = \xi, \xi(\xi \vee \eta) = \xi;$$

свертывания

$$\xi \vee \eta \bar{\eta} = \xi, \xi(\eta \vee \bar{\eta}) = \xi;$$

двойного отрицания

$$\bar{\bar{\xi}} = \xi;$$

де Моргана

$$\overline{\xi \vee \eta} = \bar{\xi} \bar{\eta}, \overline{\xi \eta} = \bar{\xi} \vee \bar{\eta};$$

исключенного третьего

$$\xi \vee \bar{\xi} = 1;$$

противоречия

$$\xi \bar{\xi} = 0;$$

сохранения нуля и единицы

$$\xi \wedge 0 = 0, \xi \vee 1 = 1;$$

исключения нуля и единицы

$$\xi \vee 0 = \xi, \xi \wedge 1 = \xi.$$

Логические операции \vee, \wedge, \neg подчиняются законам логики. Чтобы удостовериться в этом, достаточно проверить, что все эти законы превращаются в равенства для каждого из восьми возможных наборов значений переменных $\xi, \eta, \zeta \in \Sigma$. Эту проверку легко осуществить практически. К примеру, берем набор $\xi=1, \eta=0, \zeta=1$ и убеждаемся, что он обращает в равенство второй закон дистрибутивности $\xi \eta \vee \zeta = (\xi \vee \zeta)(\eta \vee \zeta)$:

$$1 \cdot 0 = 1 = 0 \vee 1 = 1, (1 \vee 1)(0 \vee 1) = 1 \cdot 1 = 1.$$

Вместе с тем, свойства операций \vee, \wedge, \neg , выраженные законами логики, однозначно определяют эти операции. Действительно, из законов $P \vee 1 = 1, P \vee 0 = P$ выводим

$$0 \vee 0 = 0, 0 \vee 1 = 1 \vee 0 = 1 \vee 1 = 1;$$

из законов $P \wedge 0 = 0$, $P \wedge 1 = P$ выводим

$$0 \wedge 0 = 0 \wedge 1 = 1 \wedge 0 = 0, 1 \wedge 1 = 1;$$

из законов $P \vee \bar{P} = 1$, $P\bar{P} = 0$, $\overline{P \vee Q} = \bar{P}\bar{Q}$, $\overline{PQ} = \bar{P} \vee \bar{Q}$, $\bar{\bar{P}} = P$, $P \vee Q = Q \vee P$, $PQ = QP$ выводим

$$\bar{0} = \overline{\bar{P}\bar{P}} = \bar{P} \vee \bar{\bar{P}} = \bar{P} \vee P = P \vee \bar{P} = 1,$$

$$\bar{1} = \overline{P \vee \bar{P}} = \bar{P}\bar{\bar{P}} = \bar{P}P = P\bar{P} = 0.$$

Мы фактически построили аксиоматическую теорию алгебры логики. Как видим, она имеет единственную модель, совпадающую с алгеброй логики. Следовательно, эта теория *непротиворечива* и *содержательно полна*. Итак, мы получили полноценное абстрактное определение понятия алгебры логики.

Далее, опишем на языке алгебры предикатов аксиоматическую теорию алгебры логики. Вводим четыре исходных предиката этой теории $\Sigma(\xi)$, ДИЗ(ξ, η, ζ), КОН(ξ, η, ζ), ОТР(ξ, η). Переменные ξ, η, ζ определяем на произвольно выбираемом универсуме U . Предикат $\Sigma(\xi) = \xi^0 \vee \xi^1$ формально выражает понятие логического элемента. Если $\Sigma(\xi) = 1$, то ξ есть логический элемент, если $\Sigma(\xi) = 0$, то ξ не есть логический элемент. Предикату $\Sigma(\xi)$ соответствует множество $\Sigma = \{0, 1\}$, где 0 и 1 — имена логических элементов. Предикаты ДИЗ(ξ, η, ζ), КОН(ξ, η, ζ) и ОТР(ξ, η) формально выражают понятия дизъюнкции $\xi \vee \eta = \zeta$, конъюнкции $\xi \wedge \eta = \zeta$ и отрицания $\neg \xi = \eta$ логических элементов. Следующие три аксиомы формально выражают тот факт, что отношения ДИЗ(ξ, η, ζ), КОН(ξ, η, ζ) и ОТР(ξ, η) являются функциями:

$$\forall \xi, \eta \in \Sigma \exists! \zeta \in \Sigma \text{КОН}(\xi, \eta, \zeta);$$

$$\forall \xi, \eta \in \Sigma \exists! \zeta \in \Sigma \text{ДИЗ}(\xi, \eta, \zeta);$$

$$\forall \xi \in \Sigma \exists! \eta \in \Sigma \text{ОТР}(\xi, \eta).$$

Способ перехода к формальной записи остальных законов продемонстрируем на примере первого закона ассоциативности. Соотношение $(\xi \vee \eta) \vee \zeta = \xi \vee (\eta \vee \zeta)$ выражается высказыванием:

$$\forall \xi, \eta, \zeta \in \Sigma ((\xi \vee \eta) \vee \zeta = \xi \vee (\eta \vee \zeta)).$$

Это — сокращенная запись. Переход к полной записи приводит к следующей формулировке этого закона:

$$\forall \xi, \eta, \zeta \in \Sigma \exists \alpha, \beta, \gamma, \delta \in \Sigma (\text{ДИЗ}(\xi, \eta, \alpha) \wedge$$

$$\wedge \text{ДИЗ}(\alpha, \zeta, \beta) \wedge \text{ДИЗ}(\eta, \zeta, \gamma) \wedge \text{ДИЗ}(\zeta, \gamma, \delta) \wedge D(\beta, \delta)).$$

Здесь D — предикат равенства на универсуме $U \times U$; $\alpha = \xi \vee \eta$, $\beta = \alpha \vee \zeta$, $\gamma = \eta \vee \zeta$, $\delta = \xi \vee \gamma$.

Наконец, запишем на языке отношений и на языке алгебры предикатов модель теории алгебры логики:

$$\Sigma = \{0, 1\}, \text{ДИЗ} = \{(0, 0, 0), (0, 1, 1), (1, 0, 1), (1, 1, 1)\},$$

$$\text{КОН} = \{(0, 0, 0), (0, 1, 0), (1, 0, 0), (1, 1, 1)\},$$

$$\text{ОТР} = \{(0, 1), (1, 0)\};$$

$$\Sigma(\xi) = \xi^0 \vee \xi^1,$$

$$\text{ДИЗ}(\xi, \eta, \zeta) = \xi^0 \eta^0 \zeta^0 \vee \xi^0 \eta^1 \zeta^1 \vee \xi^1 \eta^0 \zeta^1 \vee \xi^1 \eta^1 \zeta^1,$$

$$\text{КОН}(\xi, \eta, \zeta) = \xi^0 \eta^0 \zeta^0 \vee \xi^0 \eta^1 \zeta^0 \vee \xi^1 \eta^0 \zeta^0 \vee \xi^1 \eta^1 \zeta^1,$$

$$\text{ОТР}(\xi, \eta) = \xi^0 \eta^1 \vee \xi^1 \eta^0.$$

Для теории алгебры логики пустой и одноэлементный универсумы не годятся, т.к. эта теория обращается в противоречие. Таким образом, универсум U должен содержать, по крайней мере, два элемента.

2. Алгебра Буля

Понятие алгебры логики можно обобщить. В результате такого обобщения получаем алгебру Буля. Множество $\Sigma = \{0, 1\}$ пополняем элементами произвольным образом. В результате получаем множество M . По определению $\Sigma \subseteq M$, так что любое множество M содержит элементы 0 и 1. Элементы множества M называются булевыми. Множество M играет роль носителя алгебры Буля. Вводим булевы переменные $P, Q, R \in M$ и булевы операции \vee, \wedge, \neg над элементами множества M , удовлетворяющие законам логики: $P \vee P = P$, $P \wedge P = P$; $P \vee Q = Q \vee P$, $P \wedge Q = Q \wedge P$; $(P \vee Q) \vee R = P \vee (Q \vee R)$, $(P \wedge Q) \wedge R = P \wedge (Q \wedge R)$; $(P \vee Q)R = P R \vee Q R$, $PQ \vee R = (P \vee R)(Q \vee R)$; $P \vee PQ = Q$, $P(P \vee Q) = P$; $P \vee Q\bar{Q} = P$, $P(Q \vee \bar{Q}) = P$; $\bar{\bar{P}} = P$; $\overline{P \vee Q} = \bar{P}\bar{Q}$, $\overline{PQ} = \bar{P} \vee \bar{Q}$; $P \vee \bar{P} = 1$, $P \bar{P} = 0$; $P \wedge 0 = 0$, $P \vee 1 = 1$; $P \vee 0 = P$, $P \wedge 1 = P$.

Система, состоящая из множества M и заданных на нем булевых операций \vee, \wedge, \neg , которые удовлетворяют законам логики, называется алгеброй Буля. Булевы операции \vee, \wedge, \neg выполняют в алгебре Буля роль базисных операций. Если принять $M = \Sigma$, то алгебра Буля превращается в ее частный случай — алгебру логики. Алгебра логики является подалгеброй алгебры Буля с минимальным числом элементов в ее носителе M . Исходная система аксиом алгебры Буля избыточна. Известна [1] более экономная система аксиом алгебры Буля, состоящая всего из семи аксиом: $P \vee P = P$, $P \vee Q = Q \vee P$, $(P \vee Q) \vee R = P \vee (Q \vee R)$, $(P \vee Q)R = P R \vee Q R$, $P \vee Q\bar{Q} = P$, $\bar{\bar{P}} = P$, $\overline{P \vee Q} = \bar{P}\bar{Q}$. В этой системе элементы 0 и 1 явно не введены. Но их можно ввести прямым определением: $P \wedge \bar{P} = 0$ и $P \vee \bar{P} = 1$, где P — произвольно выбранный элемент множества M . Если операции \vee и \wedge уже определены, то отрицание $Q = \bar{P}$ определяется однозначно: оно удовлетворяет следующей системе уравнений: $P \vee Q = 1$ и $P \wedge Q = 0$.

Все законы в исходной аксиоматической системе алгебры Буля — парные, они называются двойствен-

ными друг другу. Каждый закон в паре превращается в двойственный ему после замены в нем 0 на 1, 1 на 0, \vee на \wedge и \wedge на \vee . Исключение представляет закон двойного отрицания, который двойствен самому себе, т.е. *самодвойствен*. Второй закон в каждой паре выводится из первого и из закона двойного отрицания. Выведем, к примеру, второй закон де Моргана $\overline{PQ} = \overline{P} \vee \overline{Q}$ из первого $\overline{P \vee Q} = \overline{P} \overline{Q}$ и из закона двойного отрицания $\overline{\overline{P}} = P$:

$$\overline{PQ} = \overline{\overline{\overline{PQ}}} = \overline{\overline{(P)(Q)}} = \overline{\overline{P \vee Q}} = \overline{\overline{P}} \vee \overline{\overline{Q}}.$$

Первый закон элиминации $P \vee PQ = Q$ следует из законов $P \wedge 1 = P$, $(P \vee Q)R = P R \vee Q R$, $P \vee Q = Q \vee P$, $P \vee 1 = 1$:

$$P \vee PQ = P \wedge 1 \vee PQ = P(1 \vee Q) = P(Q \vee 1) = P \wedge 1 = P.$$

Можно доказать, что и законы $P \wedge 0 = 0$, $P \vee 0 = P$ исходной системы алгебры Буля выводятся из ее экономной системы, так что обе аксиоматические системы логически равносильны. Вторая система аксиом несократима. Оказывается, что алгебры Буля существуют не для любых носителей M . Можно доказать, что в роли носителя алгебры Буля годятся лишь такие множества M , мощность $|M|$ которых выражается формулой:

$$|M| = 2^{|N|},$$

где N – какое-нибудь непустое множество.

3. Алгебра предикатов

Разновидностью булевой алгебры является алгебра конечных предикатов, описываемая ниже. Пусть A_1, A_2, \dots, A_m – какие-нибудь непустые конечные множества. Переменные x_1, x_2, \dots, x_m , заданные на этих множествах, называются предметными. Значения предметных переменных называются предметами. Множество $U = A_1 \cup A_2 \cup \dots \cup A_m$ всех предметов называется универсумом. Предикатом P типа $\tau = (x_1 \in A_1, x_2 \in A_2, \dots, x_m \in A_m)$ называется любая функция $P(x_1, x_2, \dots, x_m) = \xi$, отображающая декартово произведение $A_1 \times A_2 \times \dots \times A_m$ множеств A_1, A_2, \dots, A_m в множество Σ . Будем говорить, что предикаты P типа τ заданы на множестве A . Формируем множество M_τ всех предикатов типа τ . Любая булева алгебра, у которой в роли множества M выступает множество M_τ , называется алгеброй конечных предикатов типа τ . В роли базисных операций в алгебре конечных предикатов выступают операции дизъюнкции $P \vee Q$, конъюнкции $P \wedge Q$ предикатов $P, Q \in M_\tau$ и отрицания $\neg P$ предиката $P \in M_\tau$, которые определяются следующими равенствами:

$$\begin{aligned} (P \vee Q)(x_1, x_2, \dots, x_m) &= \\ &= P(x_1, x_2, \dots, x_m) \vee Q(x_1, x_2, \dots, x_m); \\ (P \wedge Q)(x_1, x_2, \dots, x_m) &= \\ &= P(x_1, x_2, \dots, x_m) \wedge Q(x_1, x_2, \dots, x_m); \end{aligned}$$

$$(\neg P)(x_1, x_2, \dots, x_m) = \neg(P(x_1, x_2, \dots, x_m)).$$

Имеется ввиду, что эти равенства справедливы при любых $x_1 \in A_1, x_2 \in A_2, \dots, x_m \in A_m$. Непосредственной проверкой можно доказать, что алгебра конечных предикатов любого типа τ удовлетворяет всем аксиомам булевой алгебры. В роли элементов $0, 1 \in M_\tau$ в булевой алгебре конечных предикатов используются тождественно ложный и тождественно истинный предикаты, определяемые следующим образом:

$$0(x_1, x_2, \dots, x_m) = 0,$$

$$1(x_1, x_2, \dots, x_m) = 1$$

при любых $x_1 \in A_1, x_2 \in A_2, \dots, x_m \in A_m$.

Любая булева алгебра, заданная на конечном носителе M , называется конечной. Важно то, что каждому варианту конечной булевой алгебры можно поставить во взаимно однозначное соответствие изоморфный ему вариант алгебры конечных предикатов. Это означает, что любая булева алгебра, скрытая в абстрактном определении конечной булевой алгебры, может быть сконструирована явно в виде конкретной алгебры конечных предикатов, пригодной для практического использования.

Канонической алгеброй предикатов над M_τ называется любая булева алгебра конечных предикатов над M_τ , базис элементов которой образован из предикатов $0, 1$ и всех предикатов узнавания предмета

$$P_{i,a} = x_i^a,$$

где $i = \overline{1, m}, a \in A_i$. Каноническую алгебру предикатов можно определить абстрактно для любого носителя M_τ , если к системе законов булевой алгебры добавить закон истинности

$$\bigvee_{a \in A_i} P_{i,a} = 1, \text{ где } i = \overline{1, m}$$

и закон ложности

$$P_{i,a_1} \wedge P_{i,a_2} = 0, \text{ если } a_1 \neq a_2; a_1, a_2 \in A_i, i = \overline{1, m}.$$

Дизъюнктивно-конъюнктивной алгеброй предикатов над M_τ называется такая алгебра предикатов над M_τ , базис операций которой образован из операций дизъюнкции и конъюнкции, а базис элементов – из предикатов $0, 1$ и всех предикатов узнавания предмета x_i^a , где $i = \overline{1, m}, a \in A_i$. При любом τ дизъюнктивно-конъюнктивная алгебра предикатов полна. Каждый предикат $P \in M_\tau$ в ней выражается формулой:

$$\begin{aligned} P(x_1, x_2, \dots, x_m) &= \\ &= \bigvee_{(a_1, a_2, \dots, a_m) \in A} P(a_1, a_2, \dots, a_m) x_1^{a_1} x_2^{a_2} \dots x_m^{a_m}. \end{aligned} \quad (1)$$

Произведения вида $x_1^{a_1} x_2^{a_2} \dots x_m^{a_m}$ называются конституэнтами единицы предиката P . Формула, стоящая в правой части равенства (1), называется

совершенной дизъюнктивной нормальной формой (сокращенно *СДНФ*) предиката P .

Рассмотрим пример *СДНФ* предиката. Пусть $A_1 = A_2 = A_3 = A = \{a, b, c\}$, $x_1 \in A_1, x_2 \in A_2, x_3 \in A_3$, предикат $P(x_1, x_2, x_3)$ задан табл.1.

Таблица 1

		$x_2 x_3$								
		aa	ab	ac	ba	bb	bc	ca	cb	cc
x_1	a	0	1	1	0	1	0	0	0	0
	b	0	1	1	1	1	0	0	0	0
	c	0	0	0	0	1	0	0	0	1

СДНФ предиката P , заданного табл. 1, выражается формулой:

$$P(x_1, x_2, x_3) = x_1^a x_2^a x_3^b \vee x_1^a x_2^a x_3^c \vee x_1^a x_2^b x_3^b \vee x_1^b x_2^a x_3^b \vee x_1^b x_2^a x_3^c \vee x_1^b x_2^b x_3^a \vee x_1^b x_2^b x_3^b \vee x_1^c x_2^b x_3^b \vee x_1^c x_2^c x_3^c. \quad (2)$$

СДНФ предиката ценна тем, что по ней легко отыскивается множество всех корней уравнения

$$P(x_1, x_2, \dots, x_m) = 1. \quad (3)$$

Каждой конституенте единицы $x_1^{a_1} x_2^{a_2} \dots x_m^{a_m}$ предиката P соответствует свое решение (a_1, a_2, \dots, a_m) , обращающее уравнение (3) в тождество

$$P(a_1, a_2, \dots, a_m) = 1.$$

Множество всех корней уравнения (3) называется *отношением P* , соответствующим предикату P . Предикату (2) соответствует отношение:

$$P = \{(a, a, b), (a, a, c), (a, b, b), (b, a, b), (b, a, c), (b, b, a), (b, b, b), (c, b, b), (c, c, c)\}.$$

4. Скобочная форма предиката

Все решения уравнения (3) можно записать в более компактном виде, если воспользоваться *скобочной формой предиката P* . Ее можно получить, применяя всюду, где только это возможно. Операцию группировки предикатов узнавания предмета:

$$\Phi x_i^{\sigma_1} \vee \Phi x_i^{\sigma_2} \vee \dots \vee \Phi x_i^{\sigma_r} = \Phi(x_i^{\sigma_1} \vee x_i^{\sigma_2} \vee \dots \vee x_i^{\sigma_r}). \quad (4)$$

Эту операцию используем при любом $i = \overline{1, m}$. Символом Φ обозначен общий множитель всех дизъюнктивных членов, стоящих в левой части равенства (4). Индекс r обозначает число всех дизъюнктивных членов в формуле предиката P , в которых встречается множитель Φ . Рассмотрим пример применения операции (4). Берем $i=1$ и множитель $\Phi = x_2^a x_3^b$. Выбираем из формулы (2) все подходящие дизъюнктивные члены. Полученную формулу подвергаем операции (4):

$$x_1^a x_2^a x_3^b \vee x_1^b x_2^a x_3^b = (x_1^a \vee x_1^b) x_2^a x_3^b.$$

Затем переходим к другому возможному множителю $\Phi = x_2^a x_3^c$:

$$x_1^a x_2^a x_3^c \vee x_1^b x_2^a x_3^c = (x_1^a \vee x_1^b) x_2^a x_3^c.$$

Остается множитель $\Phi = x_2^b x_3^b$:

$$x_1^a x_2^b x_3^b \vee x_1^b x_2^b x_3^b \vee x_1^c x_2^b x_3^b = (x_1^a \vee x_1^b \vee x_1^c) x_2^b x_3^b.$$

Переходим к $i = 2$. При $\Phi = x_1^a x_3^b$, имеем:

$$x_1^a x_2^a x_3^b \vee x_1^a x_2^b x_3^b = x_1^a (x_2^a \vee x_2^b) x_3^b.$$

При $\Phi = x_1^b x_3^b$:

$$x_1^b x_2^a x_3^b \vee x_1^b x_2^b x_3^b = x_1^b (x_2^a \vee x_2^b) x_3^b.$$

Переходим к $i = 3$. При $\Phi = x_1^a x_2^a$, имеем:

$$x_1^a x_2^a x_3^b \vee x_1^a x_2^a x_3^c = x_1^a x_2^a (x_3^b \vee x_3^c).$$

При $\Phi = x_1^b x_2^a$:

$$x_1^b x_2^a x_3^b \vee x_1^b x_2^a x_3^c = x_1^b x_2^a (x_3^b \vee x_3^c).$$

При $\Phi = x_1^b x_2^b$:

$$x_1^b x_2^b x_3^a \vee x_1^b x_2^b x_3^b = x_1^b x_2^b (x_3^a \vee x_3^b).$$

Заменяя в формуле (2) использованные дизъюнктивные члены вновь полученными, приходим к следующей формуле предиката P :

$$P(x_1, x_2, x_3) = (x_1^a \vee x_1^b) x_2^a x_3^b \vee (x_1^a \vee x_1^b) x_2^a x_3^c \vee (x_1^a \vee x_1^b \vee x_1^c) x_2^b x_3^b \vee x_1^a (x_2^a \vee x_2^b) x_3^b \vee x_1^b (x_2^a \vee x_2^b) x_3^b \vee x_1^a x_2^a (x_3^b \vee x_3^c) \vee x_1^b x_2^a (x_3^b \vee x_3^c) \vee x_1^b x_2^b (x_3^a \vee x_3^b) \vee x_1^c x_2^c x_3^c.$$

Процесс группировки еще не закончен. Принимая $i = 3$ и $\Phi = (x_1^a \vee x_1^b) x_2^a$, получаем:

$$(x_1^a \vee x_1^b) x_2^a x_3^b \vee (x_1^a \vee x_1^b) x_2^a x_3^c = (x_1^a \vee x_1^b) x_2^a (x_3^b \vee x_3^c).$$

Точно так же при $i = 1$ и $\Phi = (x_2^a \vee x_2^b) x_3^b$ имеем:

$$x_1^a (x_2^a \vee x_2^b) x_3^b \vee x_1^b (x_2^a \vee x_2^b) x_3^b = (x_1^a \vee x_1^b) (x_2^a \vee x_2^b) x_3^b.$$

Кроме того, при $i = 2$ и $\Phi = x_1^b (x_3^b \vee x_3^c)$ находим:

$$x_1^a x_2^a (x_3^b \vee x_3^c) \vee x_1^b x_2^a (x_3^b \vee x_3^c) = (x_1^a \vee x_1^b) x_2^a (x_3^b \vee x_3^c).$$

Окончательно получаем следующую скобочную форму предиката P

$$P(x_1, x_2, x_3) = (x_1^a \vee x_1^b) x_2^a (x_3^b \vee x_3^c) \vee (x_1^a \vee x_1^b \vee x_1^c) x_2^b x_3^b \vee (x_1^a \vee x_1^b) (x_2^a \vee x_2^b) x_3^b \vee x_1^b x_2^b (x_3^a \vee x_3^b) \vee x_1^c x_2^c x_3^c \quad (5)$$

Формула (5) состоит из пяти дизъюнктивных членов, называемых *простыми импликантами*. Каждая из них представляет собой наибольшее декартово произве-

дение, включенное в отношение P . Дизъюнкция всех простых импликант предиката P соответствует множеству всех решений уравнения (3).

Простые импликанты предиката P можно наглядно изобразить графически на таблице предиката. Каждой простой импликанте в m -мерном предметном пространстве соответствует свое нерасширяемое декартово произведение, включенное в отношение P . В табл. 2 в качестве примера представлены все простые импликанты, фигурирующие в формуле (5).

Таблица 2

		x_2x_3								
		aa	ab	ac	ba	bb	bc	ca	cb	cc
x_1	a	0	1	1	0	1	0	0	0	0
	b	0	1	1	1	1	0	0	0	0
	c	0	0	0	0	1	0	0	0	1
			1	3	4	2				5
		$P(x_1, x_2, x_3)$								

Говорят, что каждая простая импликанта предиката P накрывает своими единицами единицы предиката P . Система простых импликант предиката P называется *полной*, если единицами всех простых импликант этой системы накрываются все единицы предиката P . Система всех простых импликант 1 – 5, представленная в табл. 2, полна. Она даже избыточна, т. к. система остается полной после исключения из нее импликанты 3.

Система простых импликант предиката P называется *несократимой*, если она полна и исключение любой импликанты из системы приводит к ее неполноте. Пример несократимой системы простых импликант для предиката (2) приведен в табл. 3.

Таблица 3

		x_2x_3								
		aa	ab	ac	ba	bb	bc	ca	cb	cc
x_1	a	0	1	1	0	1	0	0	0	0
	b	0	1	1	1	1	0	0	0	0
	c	0	0	0	0	1	0	0	0	1
		$P(x_1, x_2, x_3)$								

Несократимой системе простых импликант предиката P соответствует тупиковая скобочная форма предиката P . В нашем примере существует единственная тупиковая скобочная форма

$$P(x_1, x_2, x_3) = (x_1^a \vee x_1^b)x_2^a(x_3^b \vee x_3^c) \vee (x_1^a \vee x_1^b \vee x_1^c)x_2^b x_3^b \vee x_1^b x_2^b(x_3^a \vee x_3^b) \vee x_1^c x_2^c x_3^c. \quad (6)$$

Выводы

Тупиковых форм у предиката P может быть много. Выбирая из них форму с наименьшим числом вхождений предикатов узнавания предмета, получаем минимальную скобочную форму предиката P , которая может быть использована в качестве экономного представления всех решений уравнения (3). Следующим шагом может быть описание на языке алгебры конечных предикатов процесс минимизации формул алгебры булевых функций. Делается это с той целью, чтобы начать разработку теории реляционных сетей, решающих уравнения алгебры булевых функций.

Список литературы: 1. Бондаренко, М.Ф. Теория интеллекта [Текст] / М. Ф. Бондаренко, Ю. П. Шабанов-Кушнаренко. – Х.: Изд-во «СМИТ», 2007. – 576 с.

Поступила в редколлегию 11.05.2010.

УДК 519.7

Про алгебру предикатів / М.Ф. Бондаренко, Н.П. Круглікова, І.О. Лешинська, Н.Є. Русакова, Ю.П. Шабанов-Кушнаренко // Біоніка інтелекту: наук.-техн. журнал. – 2010. – № 3 (74). – С. 3–7.

У статті розглядається алгебра логіки, яка зводиться до алгебри Буля. Вводиться поняття скобочної форми в алгебрі скінчених предикатів.

Бібліогр.: 1 найм.

UDC 519.7

About algebra of predicates / M.F. Bondarenko, N.P. Kruglikova, I.O. Leshchynska, N.E. Rusakova, Yu.P. Shabanov-Kushnarenko // Bionics of Intelligence: Sci. Mag. – 2010. – № 3 (74). – P. 3–7.

Boolean algebra which is taken to algebra Bulja is examined in the article. A concept is entered will accomplish an eye form in algebra of eventual predicates.

Ref.: 1 items.

УДК 519.7



О РЕЛЯЦИОННЫХ СЕТЯХ

М. Ф. Бондаренко¹, Н. П. Кругликова², И. А. Лещинская³,
Н. Е. Русакова⁴, Ю. П. Шабанов-Кушнаренко⁵

¹⁻⁵ ХНУРЭ, г. Харьков, Украина

В статье рассматривается алгебра предикатов и алгебра предикатных операций как математический фундамент для создания реляционных сетей. Рассматриваются также линейные логические операторы, являющиеся основным решающим средством в реляционных сетях. Реляционные сети рекомендуются на роль универсального решателя высокопроизводительных мозгоподобных ЭВМ.

АЛГЕБРА ПРЕДИКАТОВ, АЛГЕБРА ПРЕДИКАТНЫХ ОПЕРАЦИЙ, ЛИНЕЙНЫЕ ЛОГИЧЕСКИЕ ОПЕРАТОРЫ, КВАНТОРНАЯ АЛГЕБРА, РЕЛЯЦИОННАЯ СЕТЬ

Введение

Предикаты являются основным математическим инструментом, предназначенным для формального описания объектов бионики интеллекта. Язык алгебры предикатов представляет собой универсальное средство формального описания любых механизмов интеллекта человека и машин. Разработчики, проектирующие средства искусственного интеллекта используют алгебру предикатов для начального формального описания моделей. Следующим этапом является алгебра предикатных операций, на которой выражаются любые действия над отношениями. Отношения выражают свойства предметов и связи между ними. Они представляют собой универсальное средство формального описания любых объектов. За две с половиной тысячи лет науке не удалось обнаружить в мире ни одного объекта, о котором можно было бы с уверенностью сказать, что он, в принципе, не поддается формальному описанию с помощью отношений.

Целью данной статьи является использование алгебры предикатов и алгебры предикатных операций для создания реляционных сетей и рассмотрение линейных логических операторов как решающего средства в реляционных сетях.

1. Предметы и предикаты

Введем какое-нибудь непустое конечное множество $U = \{a_1, a_2, \dots, a_k\}$ различных элементов a_1, a_2, \dots, a_k . Символом k обозначено число всех элементов множества U . Элементы множества U называются *предметами*. Множество U называется *универсумом предметов*. Кроме того, введем непустое конечное множество $V = \{x_1, x_2, \dots, x_m\}$ различных предметных переменных x_1, x_2, \dots, x_m . Символом m обозначено число всех предметных переменных. Множество V называется *универсумом предметных переменных*. Каждой переменной x_i ($i = \overline{1, m}$) соответствует свое непустое множество A_i всех ее возможных значений. Оно называется *областью значений* переменной x_i . Таким образом, $x_1 \in A_1, x_2 \in A_2, \dots, x_m \in A_m$. Множества A_1, A_2, \dots, A_m можно выбирать произволь-

но, но лишь из числа подмножеств универсума U , так что $A_1, A_2, \dots, A_m \subseteq U$. Если принять $x_1 = a_1, x_2 = a_2, \dots, x_m = a_m$, то при этом образуется *набор* (a_1, a_2, \dots, a_m) предметов a_1, a_2, \dots, a_m , которые находятся на *местах* x_1, x_2, \dots, x_m . Элементы a_1, a_2, \dots, a_m называются *первым, вторым, т-ым компонентами набора* (a_1, a_2, \dots, a_m) .

Образует множество A всех возможных наборов (a_1, a_2, \dots, a_m) , составленных из предметов $x_1 = a_1, x_2 = a_2, \dots, x_m = a_m$ ($a_1 \in A_1, a_2 \in A_2, \dots, a_m \in A_m$). Оно называется *декартовым произведением* множеств A_1, A_2, \dots, A_m и записывается в виде $A = A_1 \times A_2 \times \dots \times A_m$. Множество A называется *предметным пространством типа* (A_1, A_2, \dots, A_m) . Число m называется *размерностью* пространства A . Множества A_1, A_2, \dots, A_m называются *координатными осями* пространства A . Введем логическую переменную $\xi = \{0, 1\}$. Элементы 0 и 1 множества $\{0, 1\}$ называются *логическими*. Предикатом типа (A_1, A_2, \dots, A_m) называется любая функция $P(x_1, x_2, \dots, x_m) = \xi$, отображающая множество $A = A_1 \times A_2 \times \dots \times A_m$ в множество $\{0, 1\}$. Образует множество M всех предикатов типа (A_1, A_2, \dots, A_m) . *Алгеброй предикатов* над M называется любая алгебра, носителем которой является множество всех предикатов на $A_1 \times A_2 \times \dots \times A_m$. Рассмотрим пример предиката. Пусть $A_1 = \{1, 2, 3, 4\}, A_2 = \{3, 4, 5, 6\}$. Предикат типа (A_1, A_2) задаем табл. 1:

		x_2		A_2	
		3	4	5	6
A_1	1	0	1	0	0
	2	1	0	1	0
	3	0	1	0	1
	4	0	0	1	0
		$P(x_1, x_2)$			

Дизъюнкцией $\xi \vee \eta$, *конъюнкцией* $\xi \wedge \eta = \xi \cdot \eta = \xi \eta$ и *отрицанием* $\neg \xi = \bar{\xi}$ логических элементов $\xi, \eta \in \{0, 1\}$ называются операции: $0 \vee 0 = 0, 0 \vee 1 = 1 \vee 0 = 1 \vee 1 = 1; 0 \wedge 0 = 0 \wedge 1 = 1 \wedge 0 = 0, 1 \wedge 1 = 1; \neg 0 = 1, \neg 1 = 0$. Операции \vee, \wedge, \neg называются *булевыми*. *Дизъюнкцией* $P \vee Q$, *конъюнкцией* $P \wedge Q$ и

отрицаем $\neg P$ предикатов $P, Q \in M$ называются операции:

$$\begin{aligned} (P \vee Q)(x_1, x_2, \dots, x_m) &= \\ &= P(x_1, x_2, \dots, x_m) \vee Q(x_1, x_2, \dots, x_m); \\ (P \wedge Q)(x_1, x_2, \dots, x_m) &= \\ &= P(x_1, x_2, \dots, x_m) \wedge Q(x_1, x_2, \dots, x_m); \\ (\neg P)(x_1, x_2, \dots, x_m) &= \neg(P(x_1, x_2, \dots, x_m)). \end{aligned}$$

Приводим примеры булевых операций над предикатами:

Таблица 2

	x_2	A_2			
	x_1	3	4	5	6
1	1	0	1	0	
2	0	1	0	1	
3	0	0	0	0	
4	0	0	1	0	
A_1					
					$P(x_1, x_2)$

	x_2	A_2			
	x_1	3	4	5	6
1	1	0	1	0	
2	1	1	0	0	
3	1	0	0	0	
4	1	0	1	0	
A_1					
					$Q(x_1, x_2)$

	x_2	A_2			
	x_1	3	4	5	6
1	1	0	1	0	
2	1	1	0	1	
3	1	0	0	0	
4	1	0	1	0	
A_1					
					$(P \vee Q)(x_1, x_2)$

	x_2	A_2			
	x_1	3	4	5	6
1	1	0	1	0	
2	0	1	0	0	
3	0	0	0	0	
4	0	0	1	0	
A_1					
					$(P \wedge Q)(x_1, x_2)$

	x_2	A_2			
	x_1	3	4	5	6
1	0	1	0	1	
2	1	0	1	0	
3	1	1	1	1	
4	1	1	0	1	
A_1					
					$(\neg P)(x_1, x_2)$

Булевы операции можно выполнять только над однотипными предикатами.

2. Алгебры предикатов

Булевой алгеброй предикатов над M называется любая алгебра предикатов над M , базис операций которой образован из операций дизъюнкции, конъюнкции и отрицания предикатов. Основные законы булевой алгебры предикатов: *идемпотентности* $P \vee P = P, P \wedge P = P$; *коммутативности* $P \vee Q = Q \vee P, P \wedge Q = Q \wedge P$; *ассоциативности* $(P \vee Q) \vee R = P \vee (Q \vee R), (P \wedge Q) \wedge R = P \wedge (Q \wedge R)$; *дистрибутивности* $(P \vee Q) \wedge R = P \wedge R \vee Q \wedge R, P \wedge (Q \vee R) = (P \wedge Q) \vee (P \wedge R)$; *элиминации* $P \vee P \wedge Q = Q, P \wedge P \vee Q = P$; *свертывания* $P \vee Q \wedge \bar{Q} = P, P \wedge (Q \vee \bar{Q}) = P$; *двойного отрицания* $\bar{\bar{P}} = P$; *де Моргана* $\overline{P \vee Q} = \bar{P} \wedge \bar{Q}, \overline{P \wedge Q} = \bar{P} \vee \bar{Q}$. Основные законы верны для всех $P, Q \in M$. Тождественно ложным называется предикат 0, который равен $0(x_1, x_2, \dots, x_m) = 0$ при любых $x_1 \in A_1, x_2 \in A_2, \dots, x_m \in A_m$. Тождественно истинным называется предикат 1, который равен $1(x_1, x_2, \dots, x_m) = 1$ при любых $x_1 \in A_1, x_2 \in A_2, \dots, x_m \in A_m$. Предикатом узнавания предмета $a \in A_i (i = \bar{1}, m)$ по переменной x_i называется предикат x_i^a из M , равный

$$x_i^a = \begin{cases} 1, & \text{если } x_i = a, \\ 0, & \text{если } x_i \neq a. \end{cases}$$

Канонической алгеброй предикатов над M называется булева алгебра предикатов, базис элементов которой образован из предикатов 0, 1 и всех предикатов узнавания предмета $x_i^a (i = \bar{1}, m, a \in A_i)$.

Свойства базисных элементов канонической алгебры предикатов описываются следующими законами: *исключенного третьего* $P \vee \bar{P} = 1$, *противоречия* $P \wedge \bar{P} = 0$; *сохранения 0 и 1* $P \wedge 0 = 0, P \vee 1 = 1$; *исключения 0 и 1* $P \vee 0 = P, P \wedge 1 = P$; *отрицания*

$$\overline{x_i^{a_j}} = x_i^{a_1} \vee x_i^{a_2} \vee \dots \vee x_i^{a_{j-1}} \vee x_i^{a_{j+1}} \vee \dots \vee x_i^{a_k}$$

$$(i = \bar{1}, m, A_i = \{a_1, a_2, \dots, a_j, \dots, a_k\});$$

истинности $x_i^{a_1} \vee x_i^{a_2} \vee \dots \vee x_i^{a_k} = 1$; *ложности*, если $a \neq b$, то $x_i^a x_i^b = 0$.

Дизъюнктивно-конъюнктивной алгеброй предикатов над M называется такая алгебра предикатов над M , базис операций которой образован из операций дизъюнкции и конъюнкции, а базис элементов — из предикатов 0, 1 и всех предикатов узнавания предмета $x_i^a (i = \bar{1}, m, a \in A_i)$. При любых A_1, A_2, \dots, A_m дизъюнктивно-конъюнктивная алгебра предикатов полна. На языке дизъюнктивно-конъюнктивной алгебры предикатов каждый предикат типа A_1, A_2, \dots, A_m выражается в виде:

$$P(x_1, x_2, \dots, x_m) = \bigvee_{(a_1, a_2, \dots, a_m) \in A} P(a_1, a_2, \dots, a_m) x_1^{a_1} x_2^{a_2} \dots x_m^{a_m}. \quad (1)$$

Произведения вида $x_1^{a_1} x_2^{a_2} \dots x_m^{a_m}$ называются конституэнтами единицы предиката P . Формула, стоящая в правой части равенства (1), называется совершенной дизъюнктивной нормальной формой (сокращенно СДНФ) предиката P .

В качестве примера приведем СДНФ предиката, заданного табл. 3.

Таблица 3

Предикат $P(x_1, x_2, x_3)$

		$x_2 x_3$								
		$a a$	$a b$	$a c$	$b a$	$b b$	$b c$	$c a$	$c b$	$c c$
x_1	a	0	1	1	0	1	0	0	0	0
	b	0	1	1	1	1	0	0	0	0
	c	0	0	0	0	1	0	0	0	1

$$\begin{aligned} P(x_1, x_2, x_3) &= x_1^a x_2^a x_3^b \vee x_1^a x_2^a x_3^c \vee \\ &\vee x_1^a x_2^b x_3^b \vee x_1^b x_2^a x_3^b \vee x_1^b x_2^a x_3^c \vee x_1^b x_2^b x_3^a \vee \\ &\vee x_1^b x_2^b x_3^b \vee x_1^c x_2^b x_3^b \vee x_1^c x_2^c x_3^c. \end{aligned} \quad (2)$$

По СДНФ предиката можно легко отыскать множество всех решений уравнения

$$P(x_1, x_2, \dots, x_m) = 1. \quad (3)$$

Каждой конституэнте единицы $x_1^{a_1} x_2^{a_2} \dots x_m^{a_m}$ предиката P соответствует свое решение (a_1, a_2, \dots, a_m) , обращающее уравнение (3) в тождество

$$P(a_1, a_2, \dots, a_m) = 1.$$

Множество всех корней уравнения (3) называется *отношением P*, соответствующим предикату P. Предикату (2) соответствует отношение:

$$P = \{(a, a, b), (a, a, c), (a, b, b), (b, a, b), (b, a, c), (b, b, a), (b, b, b), (c, b, b), (c, c, c)\}.$$

3. Линейные логические операторы

Отношения, задаваемые бинарными предикатами, называются *соответствиями*. Пусть $x_1 \in A_1, x_2 \in A_2$. Зададим какое-нибудь соответствие уравнением $K(x_1, x_2) = 1$ дизъюнктивно-конъюнктивной алгебры предикатов. Например $A_1 = A_2 = \{a, b, c\}$,

$$K(x_1, x_2) = x_1^a x_2^b \vee x_1^a x_2^c \vee x_1^b x_2^a.$$

Соответствие K определяем уравнением:

$$x_1^a x_2^b \vee x_1^a x_2^c \vee x_1^b x_2^a = 1. \quad (4)$$

Если $r \in A_1, s \in A_2$ и $K(r, s) = 1$, то предмет s называется *образом предмета r*, а предмет r — *прообразом предмета s относительно соответствия K*. Например, положим $x_1 = a, x_2 = c$. Тогда $x_1^a x_2^b \vee x_1^a x_2^c \vee x_1^b x_2^a = a^a c^b \vee a^a c^c \vee a^b c^a = 1 \cdot 0 \vee 1 \cdot 1 \vee 0 \cdot 0 = 1$, а значит, предмет c является образом предмета a относительно соответствия (4). Аналогично находим, что для соответствия (4) предмет a не является прообразом предмета a, поскольку в этом случае

$$r = a, s = a \text{ и } K(r, s) = x_1^a x_2^b \vee x_1^a x_2^c \vee x_1^b x_2^a = a^a a^b \vee a^a a^c \vee a^b a^a = 1 \cdot 0 \vee 1 \cdot 0 \vee 0 \cdot 1 = 0.$$

Совокупность Q всех предметов $s \in A_2$, удовлетворяющих уравнению $K(r, s) = 1$, где $r \in A_1$, называется *полным образом предмета r* относительно соответствия K. Полный образ Q предмета r может быть вычислен по формуле:

$$\exists x_1 \in A_1 (x_1^r \cdot K(x_1, x_2)) = Q(x_2). \quad (5)$$

Здесь $Q(x_2)$ — предикат, соответствующий множеству Q. Приводим пример отыскания полного образа Q предмета $r = a$ для соответствия (4):

$$\begin{aligned} Q(x_2) &= \exists x_1 \in \{a, b, c\} (x_1^a x_2^b \vee x_1^a x_2^c \vee x_1^b x_2^a) = \\ &= a^a (a^a x_2^b \vee a^a x_2^c \vee a^b x_2^a) \vee b^a (b^a x_2^b \vee b^a x_2^c \vee b^b x_2^a) \vee \\ &\vee c^a (c^a x_2^b \vee c^a x_2^c \vee c^b x_2^a) = x_2^b \vee x_2^c. \end{aligned}$$

Таким образом, $Q = \{b, c\}$. Этот же результат можно получить и по иной формуле:

$$\forall x_1 \in A_1 (x_1^r \supset K(x_1, x_2)) = Q(x_2). \quad (6)$$

Приводим пример решения той же задачи по формуле (6):

$$\begin{aligned} Q(x_2) &= \forall x_1 \in \{a, b, c\} (x_1^a \supset (x_1^a x_2^b \vee x_1^a x_2^c \vee x_1^b x_2^a)) = \\ &= (a^a \supset a^a x_2^b \vee a^a x_2^c \vee a^b x_2^a) (b^a \supset b^a x_2^b \vee b^a x_2^c \vee b^b x_2^a) \cdot \\ &\cdot (c^a \supset c^a x_2^b \vee c^a x_2^c \vee c^b x_2^a) = x_2^b \vee x_2^c. \end{aligned}$$

Совокупность P всех предметов $r \in A_1$, удовлетворяющих уравнению $K(r, s) = 1$, где $s \in A_2$, называется *полным прообразом предмета s* относительно соответствия K. Полный прообраз P предмета s может быть вычислен по формуле

$$\exists x_2 \in A_2 (x_2^s \cdot K(x_1, x_2)) = P(x_1) \quad (7)$$

или по формуле

$$\forall x_2 \in A_2 (x_2^s \supset K(x_1, x_2)) = P(x_1). \quad (8)$$

Максимальным образом множества $P \subseteq A_1$ относительно соответствия $K(r, s) = 1$ называется множество $Q_{\max} \subseteq A_2$, представляющее собой объединение образов всех предметов $r \in P$. Он вычисляется по формуле:

$$\exists x_1 \in A_1 (P(x_1) \cdot K(x_1, x_2)) = Q_{\max}(x_2). \quad (9)$$

Минимальным образом множества $P \subseteq A_1$ относительно соответствия $K(r, s) = 1$ называется множество $Q_{\min} \subseteq A_2$, представляющее собой пересечение образов всех предметов $r \in P$. Он вычисляется по формуле:

$$\forall x_1 \in A_1 (P(x_1) \supset K(x_1, x_2)) = Q_{\min}(x_2). \quad (10)$$

Преобразование $F(P) = Q_{\max}$ вида (9) обладает *аддитивностью*:

$$F(P_1 \vee P_2) = F(P_1) \vee F(P_2)$$

относительно операции дизъюнкции и однородности

$$F(\alpha P) = \alpha F(P)$$

относительно операции конъюнкции, $\alpha \in \{0, 1\}$. Оно называется *линейным логическим оператором первого рода*.

Преобразование $\Upsilon(P) = Q_{\min}$ вида (10) обладает *аддитивностью* $\Upsilon(P_1 \wedge P_2) = \Upsilon(P_1) \wedge \Upsilon(P_2)$ относительно операции конъюнкции и *однородностью* $\Upsilon(\alpha \vee P) = \alpha \vee \Upsilon(P)$ относительно операции дизъюнкции. Оно называется *линейным логическим оператором второго рода*. Можно доказать, что любой линейный логический оператор первого рода выражается в виде (9), а второго — в виде (10) при подходящем выборе предиката $K(x_1, x_2)$, который называется *ядром линейного логического оператора*.

Существует глубокая аналогия между числовой и логической математикой. На языке числовой математики записываются законы внешнего (объективного мира), а на языке логической — внутреннего (субъективного) мира человека. Эта аналогия, в частности, проявляется в существовании линейных операторов как в логической, так и в числовой математике. Числовые линейные операторы имеют вид интегральных преобразований числовых функций. Согласно теореме Радона, интегральное преобразование имеет следующий общий вид:

$$\int_{A_1} P(x_1) K(x_1, x_2) dx_1 = Q(x_2). \quad (11)$$

Здесь $K(x_1, x_2)$ – ядро линейного интегрального преобразования; $P(x_1)$ – входная функция интегрального преобразования; $Q(x_2)$ – выходная. Знаку \int числового суммирования в выражении (11) соответствует знак \exists логического суммирования в выражении (9). Числовым функциям $K(x_1, x_2)$, $P(x_1)$, $Q(x_2)$ из выражения (11) соответствуют предикаты $K(x_1, x_2)$, $P(x_1)$, $Q_{\max}(x_2)$, фигурирующие в выражении (9). Правда, аналог выражения (10) в числовой математике обычно не используется, но его легко получить с помощью операции потенцирования, переводящей сумму в произведение. Логическим аналогом операции потенцирования служит отрицание, преобразующее квантор существования в квантор общности. Интегральные операторы – это основной вид преобразований, наблюдаемых во внешнем (физическом) мире. Подобно этому, линейные логические операторы играют определяющую роль во внутреннем (психологическом) мире человека. Именно они описывают основные действия реляционной сети, реализующей процессы мышления в живой природе и в технике.

Линейные логические операторы можно наглядно изображать при помощи *двудольных графов*. Возьмем, к примеру, следующее ядро линейного логического оператора:

$$K(x_1, x_2) = x_1^1(x_2^a \vee x_2^b) \vee x_1^2 x_2^d \vee x_1^3 x_2^c \vee x_1^4(x_2^d \vee x_2^e) \quad (12)$$

Ему соответствует двудольный граф, изображенный на рис.1.

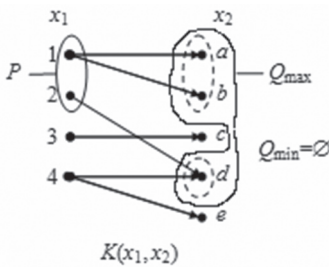


Рис. 1

Образует линейный логический оператор первого рода с ядром (12) и подадим на его вход множество $P = \{1, 2\}$, изображенное на рис.1 в виде овала. Из точек этого множества вправо простираются стрелки, оканчивающиеся элементами a, b, d , образующими максимальный образ Q_{\max} множества P . Полным образом элемента 1 служит множество $\{a, b\}$, элемента 2 – множество $\{d\}$. Эти множества не пересекаются. Поэтому выходным сигналом Q_{\min} линейного логического оператора второго рода будет пустое множество. Если поменять в двудольном графе направление всех стрелок на обратные, получаем *дуальный граф* с тем же ядром $K(x_1, x_2)$. Дуальному графу соответствуют линейный логический оператор первого рода

$$\exists x_2 \in A_2(Q(x_2)K(x_1, x_2)) = P_{\max}(x_1) \quad (13)$$

и второго рода

$$\forall x \in A_2(Q(x_2) \supset K(x_1, x_2)) = P_{\min}(x_2), \quad (14)$$

называемые *дуальными* по отношению к операторам (9) и (10). Важно отметить, что дуальные логические операторы далеко не всегда возвращают исходное множество к первоначальному виду. Так, например, множество $Q_{\max} = Q$, формируемое оператором (9) (рис. 1), возвращается дуальным оператором (13) в виде более широкого множества P_{\max} , чем исходное множество P (рис.2).

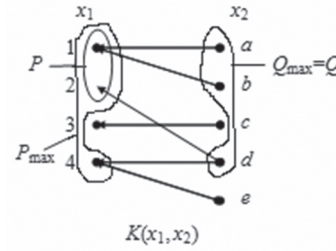


Рис. 2

4. Кванторная алгебра

Линейные логические операторы представляют собой операции над переменными предикатами. Они входят составной частью в более обширную систему, называемую *алгеброй предикатных операций*. Алгебра предикатов служит формальным средством для записи мыслей, алгебра же предикатных операций, будучи материализована в виде решающего устройства, может служить средством искусственного воспроизведения процесса мышления. Ниже описывается наиболее важная разновидность алгебры предикатных операций – *кванторная алгебра*.

Пусть U – какой-нибудь универсум предметов; $A = A_1 \times A_2 \times \dots \times A_m$ – предметное пространство на U ; $A_1, A_2, \dots, A_m \subseteq U$; x_1, x_2, \dots, x_m – предметные переменные пространства A ; $x_1 \in A_1, x_2 \in A_2, \dots, x_m \in A_m$. Обозначим буквой M множество всех предикатов типа (A_1, A_2, \dots, A_m) . Это множество называется *универсумом предикатов* типа (A_1, A_2, \dots, A_m) . Множество M^n называется *предикатным пространством* размерности n . Символами X_1, X_2, \dots, X_n обозначены предикатные переменные пространства M^n . Предикатное пространство M^n образовано из всех возможных наборов (P_1, P_2, \dots, P_n) предикатов $X_1 = P_1, X_2 = P_2, \dots, X_n = P_n$. Предикатной операцией над M^n называется каждая функция $F(X_1, X_2, \dots, X_n) = Y$, которая отображает предикатное пространство M^n в множество M .

Образует множество R всех предикатных операций над M^n . Алгеброй предикатных операций над R называется любая алгебра с носителем R . Дизъюнкцией, конъюнкцией и отрицанием предикатных операций F и G называются операции $F \vee G$, $F \wedge G$ и \bar{F} , значения которых определяются правилами:

$$\begin{aligned}(F \vee G)(X_1, X_2, \dots, X_n) &= F(X_1, X_2, \dots, X_n) \vee G(X_1, X_2, \dots, X_n); \\ (F \wedge G)(X_1, X_2, \dots, X_n) &= F(X_1, X_2, \dots, X_n) \wedge G(X_1, X_2, \dots, X_n); \\ \overline{F}(X_1, X_2, \dots, X_n) &= \overline{F(X_1, X_2, \dots, X_n)}\end{aligned}$$

для всех $F, G, F \vee G, F \wedge G, \overline{F} \in R$. Включением $F \subseteq G$ предикатных операций $F, G \in R$ называется отношение

$$F \subseteq G \Leftrightarrow F \vee G = G.$$

Булевой алгеброй предикатных операций над R называется каждая алгебра предикатных операций над R с базисом операций $F \vee G, F \wedge G, \overline{F}$, где $F, G, F \vee G, F \wedge G, \overline{F} \in R$.

Тождественной предикатной операцией по переменной X_j ($j = \overline{1, n}$) называется любая предикатная операция, имеющая значения

$$F(X_1, X_2, \dots, X_j, \dots, X_n) = X_j$$

при любых значениях переменных $X_1, X_2, \dots, X_j, \dots, X_n \in M$. Каждой предикатной переменной $X_j \in M$ соответствует своя тождественная предикатная операция $X_j \in R$. Всего существует n различных тождественных предикатных операций. Кванторной алгеброй называется булева алгебра предикатных операций, которая имеет базис операций $F \vee G, F \wedge G, \overline{F}, \exists x_i \in A_i(F), \forall x_i \in A_i(F)$ ($i = \overline{1, m}$) и базис элементов $X_j \in R$ ($j = \overline{1, n}$), и $x_i^a \in M$ ($i = \overline{1, m}, a \in A_i$). Каждая кванторная алгебра полна. Один из кванторов можно исключить из базиса кванторной алгебры без потери ею свойства полноты. Каждый из этих кванторов выражается через другой по формулам:

$$\exists x_i \in A_i(F) = \neg \forall x_i \in A_i(\neg F);$$

$$\forall x_i \in A_i(F) = \neg \exists x_i \in A_i(\neg F).$$

После такого исключения кванторная алгебра становится несократимой. Для большего удобства пользования кванторной алгеброй ее базис консервативно расширяют операциями импликации $F \supset G$, равнозначности $F \sim G$, замены $x_i / x_j(F)$, перестановки $x_i | x_j(F)$ и подстановки $x_i / a(F)$ ($i, j = \overline{1, m}, a \in A_i$), а также предикатом равенства $D(x_i, x_j)$, где $i, j = \overline{1, m}, x_i, x_j \in A_i \cap A_j$.

5. Реляционная сеть

Что такое мышление и вообще – сознательная деятельность человека? Мы предлагаем следующий ответ: это – протекающая во времени деятельность некоторой материальной системы. Весьма вероятно, что в роли этой системы (или хотя бы какой-то ее части) выступает мозг человека. Что же эта система делает? На этот вопрос предлагается такой ответ: она решает какие-то уравнения. Мышление – это процесс решения уравнений, в роли же найденных решений, т.е. продукции деятельности системы, выступают сформированные ею мысли. Если мы обратимся к внешнему миру, то обнаружим, что все в нем происходит по законам при-

роды, которые имеют вид уравнений. Любой физический процесс с математической точки зрения есть процесс решения этих уравнений. Например, траектории движения планет Солнечной системы – не что иное, как результат решения уравнений небесной механики. Какие же уравнения решает мыслящая (сознающая) система? По нашему мнению это – уравнения алгебры предикатов. Они обладают предельной общностью, поскольку способны выражать любые отношения. За две с половиной тысячи лет развития науки и техники никому еще не удалось обнаружить объект, о котором можно было бы с уверенностью сказать, что его невозможно описать на языке отношений. Каким же инструментом мыслящая система решает уравнения алгебры предикатов? Им может быть только механизм, материализующий алгебру предикатных операций, поскольку именно эта алгебра, и никакая иная, способна формально описывать любые действия над произвольными отношениями. Механизм, решающий уравнения алгебры предикатов, называется нами *реляционной сетью*. Такое название мотивировано тем, что, во-первых, мозг человека реализует нейронную сеть; во-вторых, с психологической точки зрения механизм мышления представляется как ассоциативная сеть; в-третьих, с математической точки зрения механизм мышления предстает как устройство для обработки отношений (по англ. relation). Реляционная сеть состоит из *полюсов* и *ветвей*, соединяющих полюсы. Каждому полюсу соответствует своя предметная переменная x_i с областью определения A_i ($i = \overline{1, m}$). Пара полюсов x_i и x_j , соединенных ветвью $K(x_i, x_j)$, реализует линейный логический оператор первого рода

$$\exists x_i \in A_i(P_i(x_i)K_{ij}(x_i, x_j)) = Q_{j \max}(x_j) \quad (15)$$

или второго рода

$$\forall x_i \in A_i(P_i(x_i) \supset K_{ij}(x_i, x_j)) = Q_{j \min}(x_j). \quad (16)$$

Сеть называется *первого рода*, если в ней действуют лишь операторы первого рода. Аналогично определяются сети *второго рода*. Если в сети используются операторы обоих видов, сеть называется *комбинированной*. Сеть отыскивает решение уравнения

$$K(x_1, x_2, \dots, x_m) = 1 \quad (17)$$

при ограничениях, накладываемых на области изменения переменных x_i ($i = \overline{1, m}$) $x_i \in P_i$, где $P_i \subseteq A_i$. Если решение уравнения отыскивается при более сложном ограничении

$$L(x_1, x_2, \dots, x_m) = 1,$$

тогда сеть достраивается таким образом, чтобы она соответствовала уравнению $K' = 1$, где $K' = K \wedge L$. Построению сети, реализующей предикат K , предшествует *бинаризация* предиката K , т.е. представление его в виде

$$K(x_1, x_2, \dots, x_m) = \bigwedge_{\substack{i=1 \\ j=1 \\ i \neq j}}^m K_{ij}(x_i, x_j), \quad (18)$$

которая может производиться различными способами.

Выводы

Решение уравнения (17) сетью осуществляется по тактам. В течение каждого такта одновременно срабатывают все линейные логические операторы сети. В сети первого рода после каждого такта осуществляется пересечение всех множеств $Q_{j \max}$, сходящихся со всех сторон к каждому из полюсов x_j . В сети второго рода множества $Q_{j \min}$, наоборот, объединяются. Сеть первого рода может формировать лишние решения, а второго – может не найти некоторые из действительных решений. В процессе решения уравнения (17) с увеличением номера такта работы сети множества $Q_{j \max}$ и $Q_{j \min}$ сближаются, причем всегда $Q_{j \min} \leq Q_{j \max}$. На некотором такте сближение множеств $Q_{j \min}$ и $Q_{j \max}$ прекращается. Если это достигается одновременно на всех полюсах, то на этом процесс решения уравнения (17) заканчивается. Если при всех $j = 1, m$ оказывается, что $Q_{j \min} = Q_{j \max}$, то это означает, что сеть нашла все решения, не пропустив ни одного, и не причислила к решению ни одного ошибочного. Если же такое равенство не достигнуто в конце работы сети, это значит, что сеть сработала не вполне эффективно. Этот признак может быть использован при оценке степени доброкачественности метода синтеза сети, в частности – метода

бинаризации предиката K (17). Важно отметить, что существуют такие методы синтеза сети, которые обеспечивают ее безупречную работу при решении любого уравнения вида (17).

Список литературы: 1. *Бондаренко, М.Ф.* Теория интеллекта [Текст] / М. Ф. Бондаренко, Ю. П. Шабанов-Кушнарченко. – Х.: Изд-во «СМИТ», 2007. – 576 с. 2. *Бондаренко, М. Ф.* О мозгоподобных ЭВМ [Текст] / М. Ф. Бондаренко, З. В. Дударь, И. А. Ефимова, В. А. Лещинский, С. Ю. Шабанов-Кушнарченко // Радиоэлектроника и информатика научн.-техн. журнал. – Х.: Изд-во ХНУРЭ, 2004. – № 2. – С. 89–105.

Поступила в редколлегию 11.05.2010.

УДК 519.7

Про реляційні мережі / М.Ф. Бондаренко, Н.П. Кругликова, І.О. Лещинська, Н.Є. Русакова, Ю.П. Шабанов-Кушнарченко // Біоніка інтелекту: наук.-техн. журнал. – 2010. – № 3 (74). – С. 8–13.

У статті розглядається алгебри предикатів та предикатних операцій як математичний фундамент для синтезу реляційних мереж.

Бібліогр.: 2 найм.

UDC 519.7

About relational networks / M.F. Bondarenko, N.P. Kruglikova, I.O. Leshchynska, N.E. Rusakova, Yu.P. Shabanov-Kushnarenko // Bionics of Intelligence: Sci. Mag. – 2010. – № 3 (74). – P. 8–13.

Algebra of predicates, algebra of predicate operations and linear boolean operators which are a basic decision mean in relyacionnykh networks, is examined in the article.

Ref.: 2 items.

УДК 519.7:007.52; 519.711.3



ЛОГИЧЕСКОЕ ПРЕОБРАЗОВАНИЕ КАК СРЕДСТВО РЕАЛИЗАЦИИ СЕТЕВЫХ ТЕХНОЛОГИЙ

И.Д. Вечирская¹, Н.А.Валенда²^{1,2} ХНУРЭ, г. Харьков, Украина, ira_se@list.ru

Исследованы подходы к применению линейных логических преобразований для решения проблем обработки и представления естественных языковых данных с помощью информационных технологий. Проанализированы такие понятия, как сетевая модель, структура сетевой модели, сетевой метод планирования и управления на примере логической сети глагольной флексии.

ГЛАГОЛЬНАЯ ФЛЕКСИЯ, ЛЕКСИКОГРАФИЧЕСКАЯ СИСТЕМА, ЛИНЕЙНОЕ ЛОГИЧЕСКОЕ ПРЕОБРАЗОВАНИЕ, ЛОГИЧЕСКАЯ СЕТЬ, СЕТЕВАЯ МОДЕЛЬ, СЕТЕВОЙ МЕТОД ПЛАНИРОВАНИЯ И УПРАВЛЕНИЯ, СТРУКТУРА СЕТЕВОЙ МОДЕЛИ

Введение

В настоящее время понятие сети стало настолько обширным, что трудно представить себе какую-то область деятельности, где бы не применялись сетевые технологии. Сеть как некая структура с многочисленными связями обладает определенным набором свойств как связанных с предметной областью, так и более общих, выходящих за ее границы. Возникло понятие Semantic Web, которое вместило в себя слияние Web-технологий и науки о представлении знаний, направленное на создание и поддержку сложных моделей мира, которые позволяют рассуждать о себе и о связанной с ними информации. Сеть создает синергию, эффект которой заключается в том, что за счет взаимного усиления действующих факторов общий результат превосходит сумму результатов действия отдельно взятых факторов.

Активное развитие лингвистических технологий, ценность которых как ресурсов постоянно возрастает, приводит к необходимости разработки различных лингвистических ресурсов: поисковых систем, порталов, лингвистических корпусов, различных словарей, электронных библиотек и т.д. Так, в Украинском языково-информационном фонде НАН Украины был успешно реализован проект, позволяющий большим коллективам ученых в виртуальном режиме выполнять значительные лингвистические задачи. Такая система основана на единой концептуальной модели в распределенной сетевой среде лингвистических моделей и ресурсов и объединяет цели, задачи, технологии и структуры, реализованные в системе нескольких языков [1].

Основным естественным звеном машинного перевода являются словари. Как правило, они ориентированы на детальный перевод, который предполагает наличие самых разнообразных функциональных возможностей: возможность редактирования, снятие лексических многозначностей, настройку на конкретную предметную область и т.д.

Подход к архитектурным решениям для лингвистических алгоритмов (алгоритм анализа текста на естественном языке) включает в себя три процесса: анализ входного предложения в терминах структур входного языка, преобразование этой структуры в аналогичную структуру выходного языка и затем синтез выходного предложения по полученной структуре. Перспективным предполагается использование некоторого метаязыка структур [2]. Такой метаязык должен иметь возможность описания всех структур как входного, так и выходного языков в общем случае. Сложность в этом случае состоит в разработке самого метаязыка и описании естественного языка в соответствующих терминах [3]. Таким образом, назревает необходимость создания универсального языка описания естественной языковой информации.

Создание электронных словарей осуществляется, как правило, посредством парсинга — электронный текст словаря заносят в компьютерную базу, которая поддерживает заранее разработанный формат данных. Технология реализации этого процесса бывает достаточно сложной по причине отсутствия формальных средств для описания конкретной математической модели по определению соответствующих параметров словарных комплексов. Оптимальным было бы описание всех процессов в рамках одной общей алгебры.

Универсальным формальным средством для математического описания человеческого языка является алгебра конечных предикатов, которая представляет собой [4], [5] обобщение аппарата булевых функций и многозначной логики. Конечные алфавитные операторы [6], [7], вводимые алгеброй конечных предикатов, повышают качество формального представления детерминированных, дискретных и конечных информационных процессов. Алгебра конечных предикатов позволяет описывать эти процессы с помощью уравнений. Язык уравнений вполне соответствует специфике интеллектуальных процессов, чем выгодно отличается от алгоритмического описания [4]. Исчисление

предикатов, как и исчисление высказываний, дает аппарат уравнений, который может быть использован для описания естественного языка. Эффективно решать построенные уравнения позволяет алгебра логики [8].

Таким образом, цель работы состоит в развитии алгебро-логического аппарата линейных логических преобразований как средства реализации объектов естественного языка.

**1. Линейный логический функционал.
Линейное логическое преобразование**

Введем понятие линейного логического функционала [9].

Функционал — отображение f произвольного множества X в множество R действительных (в нашем случае будем рассматривать в качестве множества X множество натуральных чисел $N \in R$) или S комплексных чисел. Если X наделено структурой векторного пространства, топологического пространства, упорядоченного множества, то возникают соответственно важные классы линейных, непрерывных, монотонных функционалов [10].

Линейным логическим функционалом на L_m называется любая функция $f: L_m \rightarrow G$, обладающая свойствами аддитивности и однородности, которые можно представить в следующем виде:

- 1) свойство аддитивности $f(x \vee y) = f(x) \vee f(y)$, где $x, y \in L_m$;
- 2) свойство однородности $f(\alpha x) = \alpha f(x)$, где $x \in L_m, \alpha \in G$.

Понятие линейного логического функционала вводят и для произвольных логических пространств.

Утверждение об общем виде линейного логического функционала. Пусть $x \in L_m, x = (\xi_1, \xi_2, \dots, \xi_m)$, где $\xi_1, \xi_2, \dots, \xi_m$ — координаты вектора x . Тогда любой линейный логический функционал $f: L_m \rightarrow G$ (и только такой функционал) может быть представлен в виде:

$$f(x) = \alpha_1 \xi_1 \vee \alpha_2 \xi_2 \vee \dots \vee \alpha_m \xi_m. \quad (1)$$

Выражение (1) называется общим видом линейного логического функционала на L_m .

Преобразование — отображение u некоторого множества M (вообще говоря, наделенного некоторой структурой) в себя. Образ элемента $\alpha \in M$ при преобразовании u обозначается $u(\alpha)$, или $u\alpha$, или α^u . Совокупность всех преобразований множества M в себя образует относительно операции умножения (суперпозиции) преобразований полугруппу, называемую симметрической полугруппой на множестве M [11].

Таким образом, обобщаем понятие функционала, переходя от множества натуральных чисел к произвольному множеству любой природы, наделив его при этом структурой логического пространства [12].

Под таким множеством можно подразумевать множество произвольных символов, в частности, множество, определяющее структуру какой-либо лексикографической модели данных [1].

Линейное логическое преобразование. Логическим преобразованием, отображающим пространство L_m размерности m в пространство L'_n размерности n , называется любая функция $F: L_m \rightarrow L'_n$. Линейным логическим преобразованием называется любое логическое преобразование, обладающее следующими свойствами:

- 1) аддитивности — $F(x \vee y) = F(x) \vee F(y)$;
- 2) однородности — $F(\alpha x) = \alpha F(x)$.

Утверждение об общем виде линейного логического преобразования. Пусть $x \in L_m, y \in L'_n, x = (\xi_1, \xi_2, \dots, \xi_m), y = (\eta_1, \eta_2, \dots, \eta_n)$, L_m и L'_n — логические пространства. Тогда любое линейное логическое преобразование можно представить в виде:

$$\eta(j) = \exists i \in A \alpha(i, j) \xi(i),$$

где $A = \{1, 2, \dots, m\}, B = \{1, 2, \dots, n\}, i \in A, j \in B$.

Утверждение об общем виде линейного логического преобразования. Предикатная интерпретация. Для того чтобы функция $F: P_A \rightarrow P_B$ была линейным логическим преобразованием, необходимо и достаточно, чтобы для любого $y \in B$, где $K(x, y)$ — предикат из $P_{A \times B}$, она имела вид

$$Q(y) = \exists x \in A (K(x, y) P(x)). \quad (2)$$

Если рассматривать построение словарей, то естественно будет под преобразованием F подразумевать отображение какого-либо слова в множество значений его грамматических признаков. При этом ядро линейного логического преобразования будет отображать сами категории выделенных грамматических признаков. Так, элементами множества грамматических признаков могут служить род, число, падеж, время, залог, склонение, лицо, вид и т.д. Для решения задач формализации естественного языка на этом этапе важно правильно выделить категории грамматических признаков, разделив их на соответствующие уровни иерархии (рис. 1).

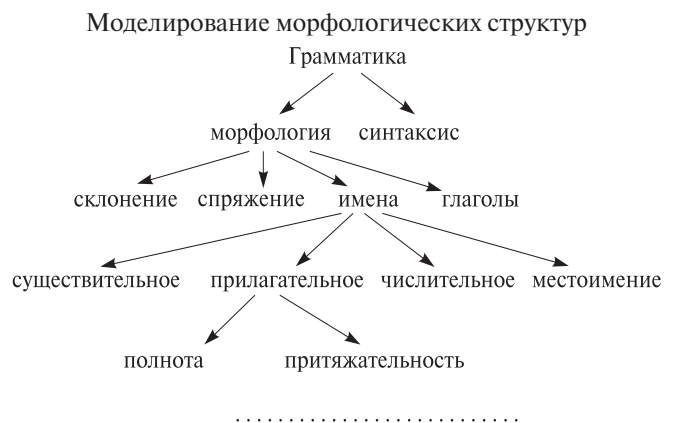


Рис. 1. Графическое представление примера моделирования морфологических структур

2. Логическая сеть как средство формального описания и реализации произвольного отношения

Далее перейдем к исследованию вопроса построения логической сети [13] для формализации естественных языковых структур.

Сеть – обобщение понятия графа. Сеть задается парой вида (V, E) , в которой V – некоторое множество, $E = (E_0, E_1, E_2, \dots)$ – семейство наборов элементов из V . В наборах $E_i \in E$ элементы могут, вообще говоря, повторяться. Элементы множества V называются вершинами сети, элементы набора E_0 – полюсами сети, наборы $E_i, i = 1, 2, \dots$, – ребрами сети.

Понятие сети используется в определении и описании управляющей системы и специальных классов управляющих систем (контактные схемы, схемы из функциональных элементов), диаграмм переходов автоматов, коммуникационных сетей и др.

Сетевая модель – интерпретация программы (плана) реализации некоторого комплекса взаимосвязанных работ в виде графа ориентированного без контуров, отражающего естественный порядок выполнения работ во времени с некоторыми дополнительными данными комплекса (стоимость, ресурсы, продолжительность и т.д.) В зависимости от условий при обработке информации сетевая модель может иметь и другие формы представления – табличную, цифровую и т.п. Все формы сетевой модели равносильны. Для логических сетей показательной является именно табличная форма представления.

Основной сетевой модели является ее структура.

Сетевая модель строится на основе своей структуры в зависимости от целей, которые ставятся при этом относительно комплекса работ.

Таким образом, логическая сеть – это неориентированный связный граф (V, E) с множеством вершин V и множеством ребер E . Множеством вершин V логической сети служит множество всех предметных переменных описываемой модели, множеством ребер E – множество всех бинарных предикатов модели, каждый из которых является функцией, зависящей от пары переменных, принадлежащих множеству V . Каждая вершина сети, называемая полюсом сети, обозначает некоторую предметную переменную модели, а каждое ребро сети, называемое ветвью сети, обозначает один из бинарных предикатов модели. Таким образом, логическая сеть – это способ представления описываемой модели. Однако представить модель логической сетью можно лишь в том случае, когда предикат модели выражается через конъюнкцию бинарных предикатов. Поэтому «логическая сеть» – это укороченный вариант термина «бинарная логическая сеть».

Логической сетью называется сеть, ребра которой отображают логическую связь (отношения) между элементами некоторого множества.

Закон, определяющий работу графа (логической сети), следующий:

$$v^{a_i} = \begin{cases} 1, & a_i \in A_i; \\ 0, & a_i \notin A_i, \end{cases}$$

где v – переменная, обозначающая некоторую вершину графа; a_i – соответствующее значение. Таким образом, вычисления происходят в полюсах логической сети. Работу ребер можно определить формулой для вычисления линейного логического преобразования (2).

Однако термин «логическая сеть» уже сейчас понимается в двух разных смыслах. С одной стороны, так можно называть предикат модели, представленный конъюнкцией бинарных предикатов. То есть это просто предикатная модель особого вида. С другой стороны, «логическая сеть» понимается как устройство (автомат – см. [14]), которое построено на основе вышеупомянутой бинарной предикатной модели и которое решает систему бинарных предикатных уравнений. Это устройство выглядит так (рис. 2):

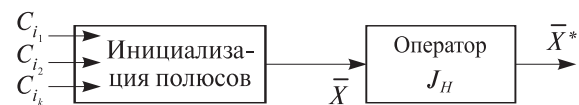


Рис. 2. Логическая сеть как устройство

Логическая сеть как устройство работает в итеративном режиме. Тогда уместно говорить о состояниях её полюсов, о движении информации по ветвям сети, о пересечении множеств при определении очередного состояния полюса и пр. На каждом такте (итерации) выполняются преобразования множеств (движение информации по ветвям) и обновляются состояния некоторых из полюсов сети.

3. Сетевая модель глагольной флексии

Рассмотрим далее сетевую модель глагольной флексии с точки зрения сетевых технологий. Модель глагольной флексии L_{21} представляет собой морфологическое отношение, описываемое следующим уравнением алгебры предикатов

$$L_{21}(X_I, X_{II}, Y) = 1, \quad (3)$$

где $X_I = \{x_2, x_3, x_{11}, x_{12}, x_{13}, x_{14}\}$ – набор грамматических признаков, характеризующих влияние дальнего текста на окончание (флексия); $X_{II} = \{x_{06}, x_{*6}, x_{07}, x_{*7}, x_{15}, x_{16}, x_{17}, x_{18}, x_{19}\}$ – набор лексических признаков, характеризующих влияние ближнего текста на окончание; $Y = \{y_1, y_2, y_3\}$ – глагольное окончание, которое состоит из первой,

второй и третьей букв – y_1, y_2 и y_3 соответственно [15].

Выполнив бинарную декомпозицию предиката (3), мы построили модель логической сети для морфологического отношения L_{21} . Полюсы сети представлены системой из двадцати двух уравнений, выражающих области определения предметных переменных модели:

$$\begin{aligned}
 &x_2^m \vee x_2^{mc} \vee x_2^c = 1, \quad x_3^e \vee x_3^m = 1, \quad x_{11}^u \vee x_{11}^a = 1, \\
 &x_{12}^u \vee x_{12}^n = 1, \quad x_{13}^u \vee x_{13}^n = 1, \quad x_{14}^1 \vee x_{14}^2 \vee x_{14}^3 = 1, \\
 &t^1 \vee t^2 \vee \dots \vee t^{12} = 1, \quad x_{06}^y \vee x_{06}^b = 1, \quad x_{*6}^y \vee x_{*6}^b = 1, \\
 &x_{07}^a \vee x_{07}^b \vee \dots \vee x_{07}^n = 1, \quad x_{*7}^e \vee x_{*7}^c = 1, \quad x_{15}^1 \vee x_{15}^2 = 1, \\
 &x_{16}^1 \vee x_{16}^2 = 1, \quad x_{17}^d \vee x_{17}^u = 1, \quad x_{18}^e \vee x_{18}^c = 1, \quad x_{19}^d \vee x_{19}^u = 1, \\
 &s^1 \vee s^2 \vee \dots \vee s^{32} = 1, \quad u^1 \vee u^2 \vee \dots \vee u^{98} = 1, \quad (4) \\
 &v^y \vee v^{yo} \vee \dots \vee v^{ub} = 1, \\
 &y_1^a \vee y_1^e \vee y_1^v \vee y_1^u \vee y_1^y \vee y_1^{yo} \vee y_1^a \vee y_1^- = 1, \\
 &y_2^u \vee y_2^v \vee y_2^m \vee y_2^y \vee y_2^u \vee y_2^u \vee y_2^u \vee y_2^b = 1, \\
 &y_3^a \vee y_3^e \vee y_3^u \vee y_3^o \vee y_3^b \vee y_3^b = 1.
 \end{aligned}$$

Ветви сети представлены системой из двадцати одного уравнения с предикатами, полученными в результате бинарной декомпозиции предиката (3); система этих уравнений имеет вид:

$$\begin{aligned}
 &Xi[x_2, t] = 1, \quad Xi[x_3, t] = 1, \quad Xi[x_{11}, t] = 1, \quad Xi[x_{12}, t] = 1, \\
 &Xi[x_{13}, t] = 1, \quad Xi[x_{14}, t] = 1, \quad Ex[x_{06}, u] = 1, \\
 &Ex[x_{*6}, u] = 1, \quad Ex[x_{07}, u] = 1, \quad Ex[x_{*7}, u] = 1, \\
 &Ex[x_{15}, u] = 1, \quad Ex[x_{16}, u] = 1, \quad Ex[x_{17}, u] = 1, \quad (5) \\
 &Ex[x_{18}, u] = 1, \quad Ex[x_{19}, u] = 1, \quad Es[u, s] = 1, \quad Phi_1[t, v] = 1, \\
 &Phi_1[s, v] = 1, \quad Phi_2[v, y_1] = 1, \quad Phi_2[v, y_2] = 1, \\
 &Phi_2[v, y_3] = 1.
 \end{aligned}$$

На рис. 3 приведена схема логической сети для морфологического отношения L_{21} . Пунктирной линией очерчены ветви, полученные в результате декомпозиции указанного отношения. Полюсы вспомогательных переменных u и v отмечены затемненными кружками.

Сетевое планирование, сетевой метод планирования и управления как метод управления при реализации некоторого комплекса работ (проекта, программы, темы и т.п.) на основе сетевой модели позволяют существенно поднять качество планирования и управления при реализации процесса работы, а также играют роль механизма в управлении, помогающего обрабатывать информацию о фактическом состоянии работ для данного момента времени и о прогнозируемых изменениях и необходимой корректировке планов для выполнения оставшихся работ.

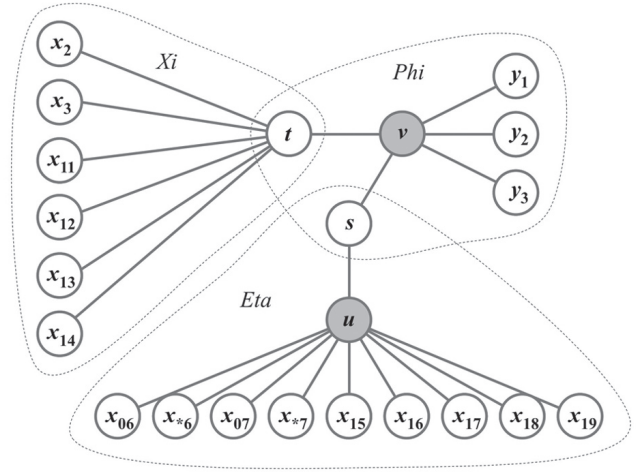


Рис. 3. Схема логической сети для модели глагольной флексии

На рис. 4, 5 изображен пример работы логической сети глагольной флексии по тактам.

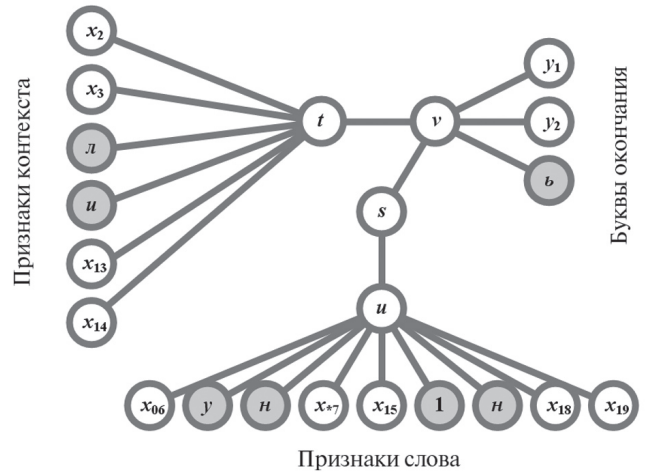


Рис. 4. Начальное состояние сети

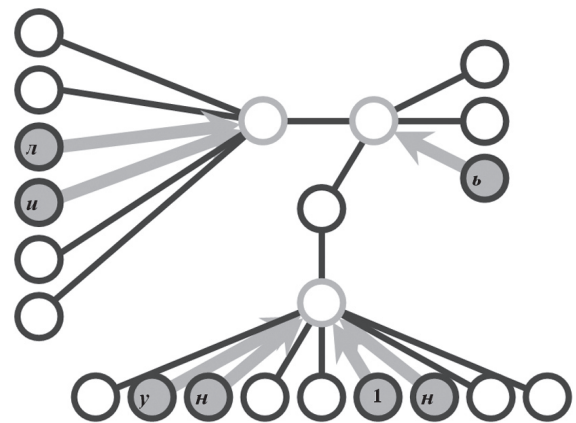


Рис. 5. Первый полутакт первого такта

В нашем примере моделирования окончания глаголов рисунки 4, 5 иллюстрируют элементы сетевого планирования (или управления).

Выводы

В статье были исследованы подходы к применению линейных логических преобразований для решения проблем обработки и представления ес-

тественноязыковых данных с помощью информационных технологий.

Актуальность сетевых технологий во всех отраслях науки и техники указывают на перспективу дальнейших исследований по применению линейных логических преобразований в самых различных областях: создание электронных словарей (уровень морфологии), машинный перевод (синтаксис, семантика).

Таким образом, на примере логической сети проанализированы такие понятия, как сетевая модель, структура сетевой модели, сетевой метод планирования и управления, определен подход к построению структур алгебры конечных предикатов для моделирования иерархии морфологических структур, элементам морфологии (глагольная флексия). В дальнейшем представляет интерес построить сеть для грамматических признаков, общих для разных частей речи (инфинитив, род, число, падеж, залог, наклонение, время, лицо вид), исследуя их взаимодействие их грамматических значений для установления соответствующих отношений.

Список литературы: 1. Широков, В.А. Элементы лексикографии [Текст] / В.А. Широков. – К.: Довіра, 2005. – 304 с. 2. Четвериков, Г.Г. Многозначные структуры (анализ, сравнение, синтез, обобщение) [Текст]. Ч.1: учеб. пособие / Г.Г. Четвериков. – К.: ІЗМН, 1997. – 192с. 3. Как переводит компьютер [Электронный ресурс] / С.В.Соколова. – Режим доступа: http://www.translationmemory.ru/technology/articles/article_Sokolova.php – 05.11.2010 г. – Загл. с экрана. 4. Шабанов-Кушнарченко, Ю.П. Теория интеллекта. Математические средства [Текст] / Ю.П. Шабанов-Кушнарченко – Х.: Вища школа, 1984. – 142 с. 5. Шабанов-Кушнарченко, Ю.П. Теория интеллекта. Технические средства [Текст] / Ю.П. Шабанов-Кушнарченко – Х.: Вища школа, 1986. – 136 с. 6. Шабанов-Кушнарченко, Ю.П. Аналитический метод явного задания конечных алфавитных операторов [Текст] / Ю.П. Шабанов-Кушнарченко // Проблемы бионики: Сб. науч. трудов. – 1980. – вып. 53 – С. 22-29. 7. Шабанов-Кушнарченко, Ю.П. Аналитические методы неявного задания конечных алфавитных операторов [Текст] / Ю.П. Шабанов-Кушнарченко // АСУ и приборы автоматики. – 1980. – Вып. 54. – С. 96-102.

8. Закревский, А.Д. Логические уравнения [Текст] / А.Д. Закревский. – Минск: Наука и техника, 1975. – 96 с. 9. Бондаренко, М.Ф. О линейных предикатах [Текст] / М.Ф. Бондаренко, С.Ю. Шабанов-Кушнарченко // Проблемы бионики: Сб. науч. трудов. – 1989. – Вып. 43. – С. 3-7. 10. Колмогоров, А.Н. Элементы теории функций и функционального анализа [Текст] / А.Н. Колмогоров, С.В. Фомин. – М., 1981. – 403 с. 11. Математическая энциклопедия [Текст]: Гл. ред. И.М. Виноградов. Т.4. Слу – Я – М., «Советская Энциклопедия», 1984. – 1248 стб., ил. 12. Вечирская, И.Д. Формальное описание логического пространства [Текст] / И.Д. Вечирская, Г.Г. Четвериков // Искусственный интеллект. – 2008. – № 3 – С. 781-789. 13. Бондаренко, М.Ф. Модели языка [Текст] / М.Ф. Бондаренко, В.А. Чикина, Ю.П. Шабанов-Кушнарченко // Бионика интеллекта: научн.-техн. журнал. – 2004. – № 1(61). – С. 27-37. 14. Толковый словарь по вычислительным системам [Текст] / Под ред. В. Иллинуорта и др.: Пер. с англ. А.К. Белоцкого и др.; Под ред. Е.К. Масловского. – М.: Машиностроение, 1990. – 560 с. 15. Логическая сеть для модели глагольной флексии русского языка [Текст] / З.В. Дударь, А.А. Иванюков, В.В. Климушев, В.И. Обризан // Восточно-европейский журнал передовых технологий – Харьков, 2006. – № 4/2. – С. 80-89.

Поступила в редколлегию 14.05.2010

УДК 519.7:007.52; 519.711.3

Логічне перетворення як засіб реалізації мережевих технологій / І.Д. Вечірська, Н.А. Валенда // Біоніка інтелекту: наук.-техн. журнал. – 2010. – № 3 (74). – С. 14-18.

Досліджено поняття мережевої моделі, структури мережевої моделі, мереженого методу планування і управління на прикладі дієслівної флексії російської мови. Визначено підхід до побудови структур алгебри скінченних предикатів для моделювання ієрархії морфологічних структур.

Л. 5. Бібліогр.: 15 найм.

UDC 519.7:007.52; 519.711.3

The linear logical transformation as network engineering / I.D. Vechirskaya, N.A.Valenda // Bionics of Intelligence: Sci. Mag. – 2010. – № 3 (74). – P. 14-18.

The model of verbal inflection on russian language as concept of network model, structure of network model, network technique of planning and management is investigated. Approach to AFP-structures building, what is models the morphological structures hierarchy is specified.

Fig. 5. Ref.: 15 items.

УДК 004.942



КОМПОНЕНТНА МОДЕЛЬ ОЦІНЮВАННЯ ВАРТОСТІ ПОХІДНИХ ФІНАНСОВИХ ІНСТРУМЕНТІВ НА ФОНДОВОМУ РИНКУ

П.П. Маслянюк¹, А.В. Рябушенко², М.В. Козленко³

¹ НТУУ «КПІ», м. Київ, Україна, mpp@amlab.ntu-kpi.kiev.ua;

² НТУУ «КПІ», м. Київ, Україна, ariabushenko@eerc.kiev.ua;

³ Ecole Polytechnique, м. Париж, Франція, mycola_k@ukr.net

В роботі представлена компонентна модель та специфікація компонента оцінювання вартості похідних фінансових інструментів системи управління фінансово-інвестиційної діяльності. Проведено аналіз існуючих рішень оцінювання вартості похідних фінансових інструментів та запропоновано вдосконалений метод оцінювання вартості похідних фінансових інструментів – модель випадкового середнього і мультифрактальної волатильності.

СИСТЕМНА ІНЖЕНЕРІЯ, КОМПОНЕНТНЕ МОДЕЛЮВАННЯ, МУЛЬТИФРАКТАЛЬНІ МОДЕЛІ, СТАХОСТИЧНІ ДИФЕРЕНЦІЙНІ РІВНЯННЯ, ПОХІДНІ ФІНАНСОВІ ІНСТРУМЕНТИ

Вступ

Похідні фінансові інструменти (ПФІ) виникли в середині XIX сторіччя. Перші контракти такого роду уклалися на Чиказькому торговому форумі, який існує з 1848 року. Прообразом сучасних ПФІ був to-arrive contract, умови якого нагадували сучасні ф'ючерси. Такі контракти користувалися попитом як альтернатива прямій торгівлі. Подібні контракти пропонувалися і на заснованій в 1919 році Чиказькій товарній біржі (Chicago Board of Trade) [1].

Першим майданчиком, який працював виключно з ПФІ, стала Чиказька опціонна біржа, заснована в 1973 році. Це була перша впорядкована структура торгівлі call-опціонами. З 1977 року ведеться торгівля put-опціонами. У 1983 році Чиказька опціонна біржа створила опціон на індекс акцій, який нині відомий як Standard and Poor's S&P 100 і досі залишається найбільш популярним опціоном у лістингу бірж. Протягом наступних років обсяги торгівлі ПФІ, передусім опціонами та ф'ючерсами, невпинно зростали, сягнувши сотень мільярдів доларів.

В Україні ринок ПФІ лише зароджується. Торгівля ф'ючерсами ведеться з 1996 року, проте недосконала законодавча база гальмує розвиток ринку. З появою законопроектів про ринок деривативів очікується стрімкий розвиток фінансово-інвестиційної діяльності в Україні. Для реалізації проектів інформатизації організаційних систем фінансово-інвестиційної діяльності актуальною постає проблема створення ефективних інформаційно-комунікаційних систем (ІКС), які забезпечують реалізацію бізнес-процесів фінансово-інвестиційної діяльності. Застосування фундаментальних положень прикладного системного аналізу і теорії систем для розв'язку задач інформатизації організаційних систем (Орг.С) фінансово-інвестиційної діяльності дозволяє використовувати сучасні методи системної інженерії для розробки таких ІКС [2–5].

У [6] пропонується визначити процес розробки як певним чином упорядкований, достатній і функціонально повний ряд робіт з розробки проектів інформатизації Орг.С. із заданими показниками ефективності та пропонується компонентний процес розробки проектів інформатизації. Застосування цього процесу дозволяє зберегти переваги ітеративного і водоспадного процесів розробки і при цьому значно розширює можливості проектування середніх та великих інформаційно-комунікаційних систем.

Оцінювання вартості ПФІ відіграє неабияку роль в підвищенні ефективності управління інвестиційним портфелем цінних паперів. Досягти бажаної структури виплат портфеля можна лише при використанні ПФІ, тому що тільки вони мають нелінійну структуру виплат. Оцінювання вартості ПФІ є необхідним і для оптимізації структури виплат інвестиційного портфеля і для управління ризиками. Оскільки вартість ПФІ залежить від інших ФІ, оптимізація інвестиційного портфеля таких інструментів потребує попередньої оцінки справедливої вартості базових ФІ та інших параметрів моделей.

Перспективність застосування системи оцінювання вартості похідних фінансових інструментів пояснюється тим, що протягом останніх років ринок ПФІ стрімко зростає. Зокрема, з 2004 до 2007 року обсяг торгів усіма видами опціонів подвоївся. Станом на 2009 рік обсяг торгівлі валютними опціонами на біржі становив близько 11 трильйонів доларів США, а опціонами з відсотковою ставкою – 108 трильйонів доларів. У 2006 році річний обіг ПФІ склав 1,6 квадрильйони доларів США і в кілька разів перевищив обіг ринку експорту товарів [7,8]. Відтак ефективно управління інвестиційним портфелем цінних паперів вимагає запровадження компонента оцінювання вартості. Темпи зростання ринку ПФІ свідчать про те, що проблема розробки компонента оцінювання вартості ПФІ є дуже актуальною.

Таким чином, формалізація процесів оцінювання вартості ПФІ та створення ІКС із застосування прикладного системного аналізу, системної інженерії та математичного моделювання на сьогодні є актуальною проблемою в Україні.

Формальним моделям “семантико-синтаксичних структур мови відводиться вирішальна роль у сучасній проблематиці комп’ютерної лінгвістики та системах штучного інтелекту (ШтІ). Це пов’язано з необхідністю створення програмно-апаратного комплексу генерації та аналізу речень природної мови (ПМ).

1. Постановка задачі

Метою роботи є розробка компонентної моделі оцінювання вартості ПФІ в системі управління фінансово-інвестиційною діяльністю [6]; дослідження існуючих методів моделювання вартості ПФІ та розробка вдосконаленої математичної моделі.

Об’єкт дослідження – компонент оцінювання вартості ПФІ системи управління фінансово-інвестиційною діяльністю.

Предмет дослідження – системна інженерія бізнес-процесів оцінювання вартості ПФІ.

2. Аналіз існуючих методів оцінювання вартості ПФІ

У попередніх дослідженнях [1, 9] проведено ґрунтовний аналіз існуючих методів оцінювання вартості похідних фінансових інструментів (ПФІ), оцінюється ефективність їх застосування. У [9] описуються також чисельні методи розв’язку стохастичних диференціальних рівнянь, які використовуються при оцінюванні вартості ПФІ.

До складу пакету математичних методів компонента оцінювання вартості ПФІ (рис. 1) включений ряд стандартних методів оцінювання: біноміальні дерева, триноміальні дерева, модель Блека-Шоулза. Також до компонента входять методи розв’язку стохастичних диференціальних рівнянь: аналітичні методи – стохастичне числення Іто, чисельні методи – метод скінченних елементів, скінченнорізницеві методи, метод скінченних об’ємів, метод Монте-Карло і американське Монте-Карло. Для знаходження значень параметрів моделей до складу компонента включений блок калібрування.

Одним з найпоширеніших на сьогодні та найпростіших методів оцінювання вартості ПФІ є біноміальна модель [10]. Цей метод виходить з припущення, що вартість опціону є дискретною випадковою величиною з біноміальним розподілом. Підраховуючи ймовірності збільшення або зменшення вартості інструмента на кожному кроці, маємо таку формулу для вартості:

$$c = e^{-r\Delta t} (pA_u + (1-p)A_d),$$

де c – поточна вартість опціону; p – ймовірність кроку вгору; u – величина кроку вгору; d – вели-

чина кроку вниз; S – поточна вартість базового ФІ; A_u – функція виплат опціону в разі зростання ціни базового фінансового інструмента в u разів; A_d – функція виплат опціону в разі зменшення ціни базового фінансового інструмента в d разів; r – безризикова дохідність.

За умови досить великої кількості кроків біноміальне дерево дозволяє досягти високої точності обчислень [1]. Особливо ефективним воно є для американських опціонів, оскільки дозволяє враховувати можливість раннього виконання.

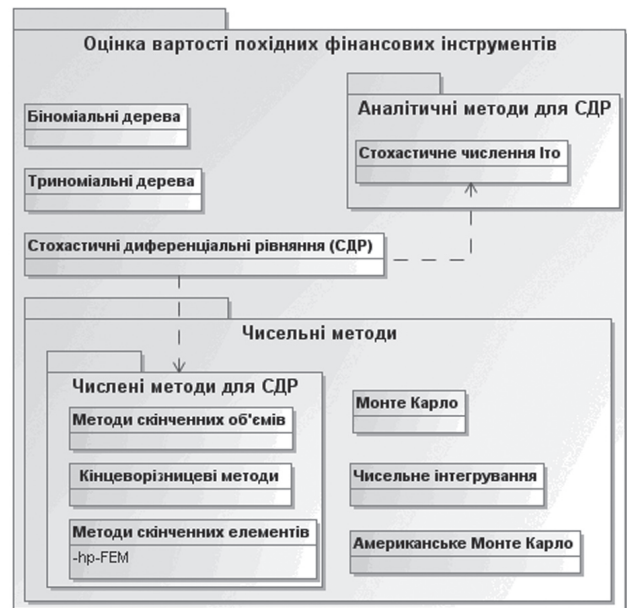


Рис. 1. Математичні методи компонента оцінювання вартості ПФІ. Пакет класів у нотації UML

Триноміальна модель є модифікацією біноміальної моделі і передбачає виділення на кожному кроці не двох, а трьох значень цін базового фінансового інструмента та опціону. Така модифікація дозволяє дещо підвищити точність обчислень [1].

Модель Блека-Шоулза, запропонована американськими економістами Фішером Блеком та Майроном Шоулзом у 1973 році, була одним з історично перших методів оцінювання вартості опціонів. Цей метод полягає в розв’язуванні стохастичного диференціального рівняння, виходячи з певних припущень про стан фінансового ринку (зокрема, про те, що вартість базових ФІ є геометричним броунівським рухом з постійною волатильністю) [11], завдяки яким запропонована модель залишається однією з найточніших та широко використовується у прогнозуванні.

До складу пакету входять також чисельні методи. Метод Монте-Карло, що використовує генерації випадкових шляхів базових ФІ, заснований на отриманні значної кількості випадкових шляхів, на основі яких оцінюється вартість ПФІ. Для оцінювання вартості американських опціонів цей метод має модифікацію, що передбачає оцінювання

вартості утримання опціону методом найменших квадратів. Методи чисельного інтегрування використовуються в тому випадку, коли оцінку вартості ПФІ можна звести до визначеного кратного інтегралу [9].

Таким чином, поєднання існуючих та нових методів оцінки справедливої вартості похідних фінансових інструментів у в одній компонентній моделі системи управління фінансово-інвестиційної діяльності дозволяє суттєво підвищити ефективність оцінювання вартості.

3. Компонентна модель оцінювання вартості ПФІ

Призначення компонента — оцінювання вартості ПФІ з метою оптимізації структури виплат інвестиційного портфеля, а також зменшення ризиків.

Компонентна модель системи управління інвестиційним портфелем та її специфікація представлена у [6].

Компонентна модель компонента оцінювання вартості ПФІ у нотатції UML [12] представлена на рис. 2.

Компонентна модель компонента оцінювання вартості ПФІ відображає внутрішню структуру компонента та дозволяє формалізувати архітектуру та топологію сутностей інформатизації бізнес-процесів фінансово-інвестиційної діяльності.

Діаграма діяльності зображена на рис. 3. Діаграма діяльності є динамічним представленням компонента оцінювання вартості ПФІ.

За допомогою графічного інтерфейсу користувач задає множину ПФІ для аналізу вартості та визначає базові ФІ для цих ПФІ, отримує необхідні дані з зовнішніх інформаційних ресурсів та передає їх через інтерфейс IDPG до компоненту оцінювання чутливостей та через інтерфейс IDPMM до блока методів обчислення вартості ПФІ. В подальшому здійснюється оцінка чутливостей у компоненті оцінки чутливостей, результати якої передаються через інтерфейс IG до компонента оцінювання ризиків. Паралельно здійснюється підрахунок вартості ПФІ одним зі стандартних методів, або за допомогою моделі випадкового середнього і мультифрактальної волатильності, Stochastic Mean Multifractal Volatility Model (SMMVM), що пред-

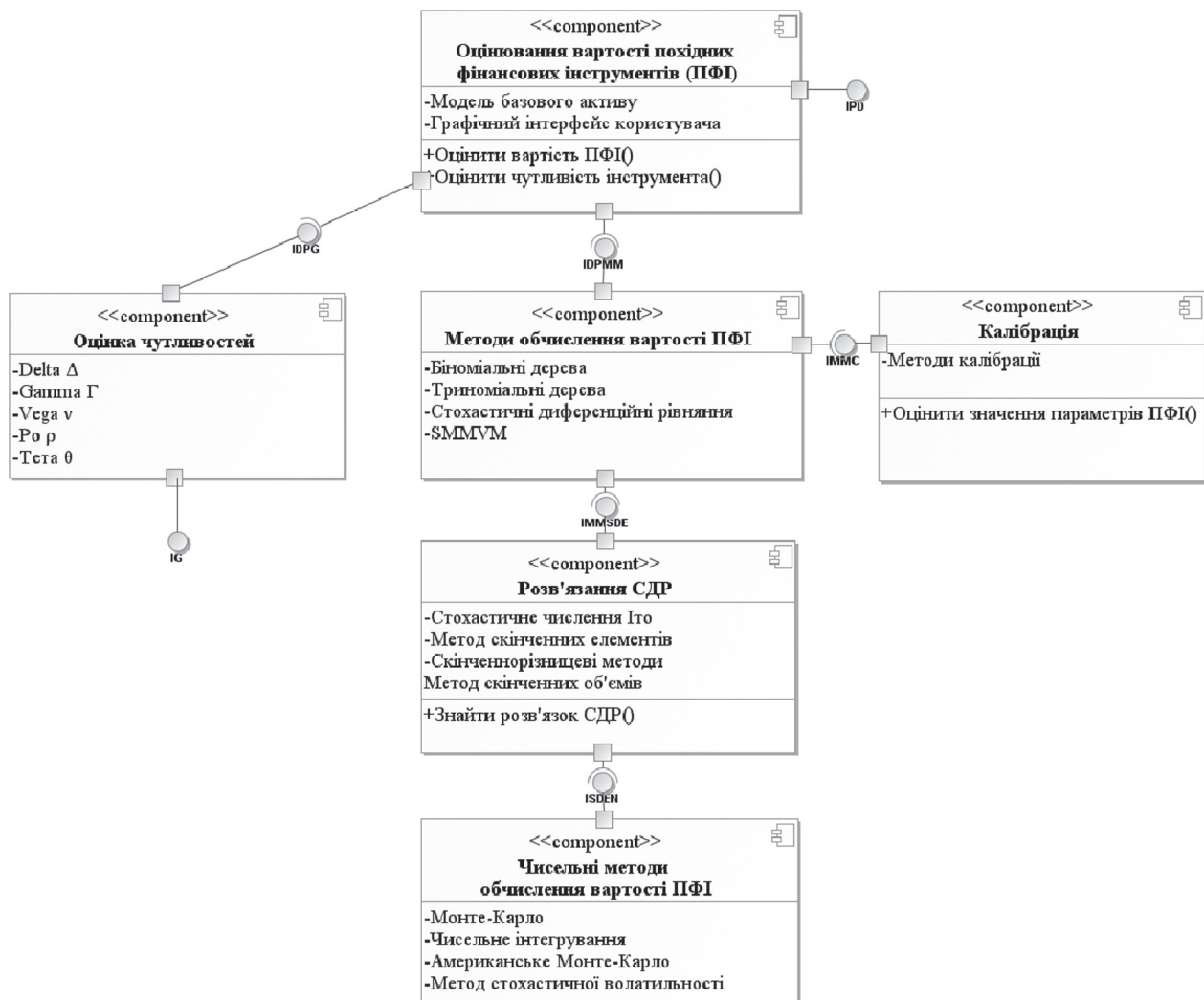


Рис. 2. Компонентна модель оцінювання вартості ПФІ. Діаграма компонентів в нотатції UML

ставлена нижче. В ході обчислення здійснюється калібрація кожної з використаних моделей, а в разі обрання моделі SMMVM або інших складних моделей, які не мають аналітичного розв’язку, додатково здійснюється розв’язання стохастичних диференційних рівнянь чисельними методами.

Компонент складається з таких складових: компонента оцінки чутливості, блоку методів обчислення вартості ПФІ, блоку чисельних методів обчислення вартості ПФІ, компонента розв’язання стохастичних диференційних рівнянь та для знаходження значень параметрів моделей до складу компонента включений блок калібрування.

Формат специфікації компонентної моделі оцінювання вартості ПФІ формується за основними загальносистемними ознаками — властивостями та операціями [5].

Назва: Компонент оцінювання вартості ПФІ системи управління інвестиційним портфелем цінних паперів.

Призначення: Інвестування в складні ФІ, що потребують спочатку оцінювання справедливої вартості, це насамперед ПФІ, чия вартість залежить від вартості інших ФІ. Оцінювання вартості ПФІ з метою оптимізації структури виплат інвестиційного портфеля, а також зменшення ризиків [13].

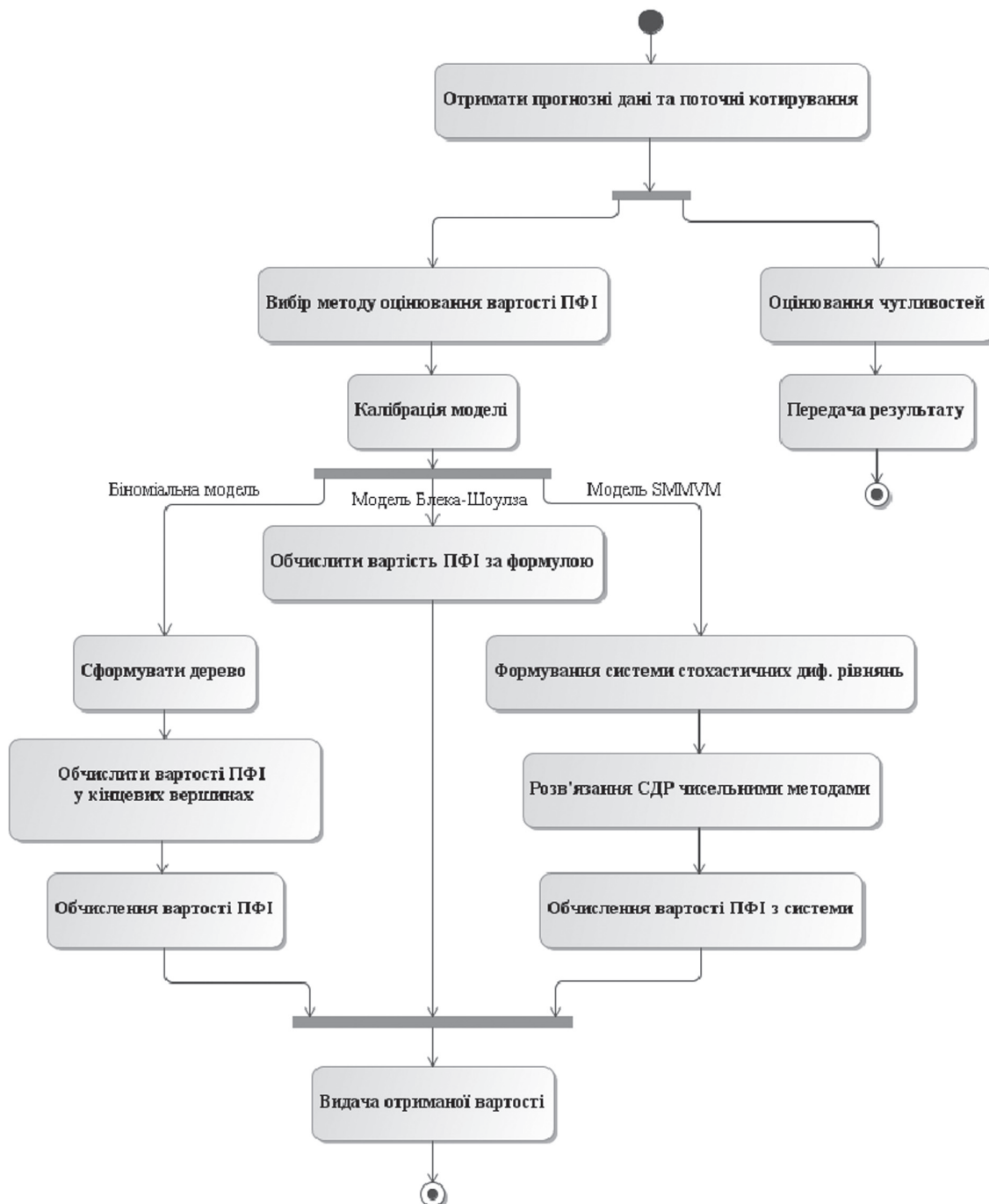


Рис. 3. Модель компонента оцінювання вартості ПФІ. Діаграма діяльності в нотатції UML

Виробник: Лабораторія прикладної математики факультету прикладної математики НТУУ «КПІ».

Функціональність: Функціональність компонента визначається його властивостями, складом та операціями.

Властивості: режим функціонування – реальний перебіг часу. Мінімізація часу реакції на зміни котирувань.

Склад:

– блок методів обчислення вартості ПФІ, в якому здійснюється власне обчислення вартості ПФІ стандартними методами;

– компонент розв’язання стохастичних диференціальних рівнянь, який використовується для знаходження аналітичного або чисельного розв’язку відповідних рівнянь;

– блок чисельних методів обчислення вартості ПФІ, який дозволяє змодельовати вартість ПФІ за допомогою одного з включених до пакету чисельних методів;

– блок калібрування, призначеного для знаходження значень параметрів моделей;

– компонент оцінки чутливості, який дозволяє оцінити залежність вартості фінансового інструмента від ряду параметрів, зокрема, від базового ФІ (Δ – дельта), від часу (θ – тета), від відсоткової ставки (ρ – ро), міру опуклості як другу похідну від вартості базового ФІ (Γ – гама) та залежність вартості опціону від волатильності (v – вега).

Операції – множина засобів, що забезпечує діяльність інвестора з ПФІ: оцінювання справедливої вартості ПФІ, визначення чутливостей, встановлення невідомих параметрів моделей за допомогою калібрації.

Інтерфейси: Компонент оцінювання вартості ПФІ використовує такі інтерфейси:

– IPPD (Interface Portal of Pricing Derivatives)

– IFEPD (Interface Financial Engineering of Pricing Derivatives)

– IPD (Interface of Pricing Derivatives)

– IDPG (Interface Derivatives Pricing Greeks)

– IDPMM (Interface Derivatives Mathematical Methods)

– IMMC (Interface Mathematical Methods Calibration)

– IMMSDE (Interface Mathematical Methods Solving Differential Equations)

– ISDEN (Interface Solving Differential Equations Numerical Methods)

– **Інтеграція:** Інтеграція компонента з усією системою здійснюється через такі порти:

– точка входу до порталу користувача – передбачає отримання даних про поточні котирування ПФІ та прогнозні значення характеристик;

– точка виходу з компонента чутливостей – передбачає видачу обчислених значень чутливостей;

– точка виходу з порталу користувача – повертає інформацію про прогнозовану вартість ПФІ.

Інтероперабельність: Взаємодія компонента з іншими складовими системи здійснюється шляхом видачі зазначеними інтерфейсами інформації про справедливу вартість ПФІ та чутливості. На вхід компонента подаються дані про поточні котирування ПФІ та прогнозні значення волатильності, кореляції та інших характеристик.

4. Вдосконалена модель оцінювання вартості ПФІ

Автори пропонують нову вдосконалену модель оцінки вартості похідних фінансових інструментів: модель випадкового середнього і мультифрактальної волатильності, Stochastic Mean Multifractal Volatility Model (SMMVM). Модель SMMVM дозволяє оцінити вартість ПФІ точніше, ніж традиційними методами, оскільки вона уникає ряду припущень Блека-Шоулза, які найчастіше не виконуються на практиці, та доповнена рядом додаткових умов для врахування специфіки українського ринку [14].

Опис моделі. Запропонована модель SMMVM базується на моделі Чена [15]. Згідно з цією моделлю миттєва дохідність задається системою стохастичних диференціальних рівнянь:

$$\begin{cases} dS_t = (\theta_t - \alpha_t)dt + \sqrt{S_t} \sigma_t dW_t \\ dr_t = (\zeta_t - r_t)dt + \sqrt{r_t} \sigma_t dW_t \\ d\sigma_t = (\beta_t - \sigma_t)dt + \sqrt{\sigma_t} \eta_t dW_t \end{cases}, \quad (1)$$

де S_t – вартість активу; r_t – відсоткова ставка (interest rate); σ_t – волатильність; W_t – вінерівський процес.

Ця система не має аналітичного розв’язку, для її розв’язку застосовуються чисельні методи.

В якості рівняння стохастичної волатильності σ_t пропонується взяти рівняння мультифрактальної моделі з Марковими перемикаваннями режимів, що дозволяє врахувати присутні на ринку ефекти довгострокової пам’яті [16].

Мультифрактальне рівняння волатильності з Марковим перемикаваннями режимів має такий вигляд:

$$\sigma_t^2 = \sigma^2 \prod_{i=1}^n M_t^{(i)}, \quad (2)$$

де σ – сталий масштабуючий множник.

Кожен компонент волатильності $M_t^{(i)}$ оновлюється з ймовірністю γ_i залежно від його рангу в ієрархії множників та залишається незмінним з ймовірністю $(1 - \gamma_i)$

$$\gamma_i = 1 - (1 - \gamma_k)^{b^{i-1}},$$

де параметри $\gamma_k \in [0, 1]$, $b \in (1, +\infty)$.

У [17] пропонується обрати такі значення параметрів: $\gamma_k = 0,5$, $b = 2$. Тоді:

$$\gamma_i = 1 - 0,5^{2^{i-1}},$$

Для компонента волатильності $M_t^{(i)}$ необхідно обрати статичний розподіл, такий що відповідає

вимогам $M_t^{(i)} > 0, E(M_t^{(i)}) = 1$. Автори пропонують обрати логнормальний розподіл, тому що він є неперервний, автоматично задовольняє вимогу $M_t^{(i)} > 0$ і має лише два параметри для калібрації:

$$M_t^{(i)} \sim LN(-\lambda, s^2).$$

З умови збереження стаціонарності $E(M_t^{(i)}) = 1$ випливає, що:

$$e^{-\lambda + 0.5s^2} = 1.$$

Звідки $s = \sqrt{2\lambda}$, і в подальшому маємо оптимізаційну задачу з одним параметром s .

Для прогнозування безризикової ставки дохідності використаємо розширену модель Вашичека [18]:

$$dr_t = (\theta(t) - ar_f)dt + \gamma dW_t, \quad (3)$$

де $\theta(t)$ – функція довгострокового рівня. Всі майбутні траєкторії r будуть прямувати до вигляду функції $\theta(t)$ в довгостроковій перспективі; a – швидкість реверсії, що характеризує швидкість, з якою траєкторії перегрупуються навколо середнього значення $\theta(t)$ в часі; σ – миттєва волатильність, є мірою миттєвої амплітуди випадковості входу в систему. Вища σ означає більшу випадковість.

Зі стандартного рівняння Блека-Шоулза, куди замість швидкості дрейфу μ підставляємо безризикову ставку рефінансування r_t , (1) та (3), формуємо таку систему рівнянь:

$$\begin{aligned} dS_t &= r_t S_t dt + \sigma_t S_t dW_t \\ dr_t &= (\theta(t) - ar_f)dt + \gamma dW_t, \end{aligned} \quad (4)$$

$$\sigma_t = \sqrt{\sigma^2 \prod_{i=1}^n M_t^{(i)}}$$

Розв'язок системи (4) здійснюється шляхом послідовного чисельного розв'язку рівнянь системи знизу догори.

Властивості моделі. У роботі [19] встановлені характеристики фондового ринку України. Врахування якомога більшої кількості цих характеристик необхідне для підвищення точності моделі.

Розглянемо кожну з характеристик та спосіб її врахування в розробленій моделі:

Дохідність:

– Автокореляційна функція дохідності зменшується не експоненціально, а гіперболічно, що призводить до нескінченної суми автокореляції. Реалізація нескінченної суми автокореляцій досягається використанням мультифрактального рівняння для волатильності.

– Дохідність на ринку розподілена не за нормальним законом, а функція щільності ймовірності має товсті хвости, що може призводити до нескінченної дисперсії. Для дохідності використовуємо Леві-розподіл, який має необхідні властивості.

– Наявність довгострокової пам'яті дохідності ПФТС, що підтверджено експонентою Херста, АКФ

та ЧАКФ. Мультифрактальне рівняння волатильності також дозволяє реалізувати цю властивість.

– Порядок дострокової пам'яті змінюється з часом. Використання мультифрактального, а не просто фрактального рівняння волатильність дозволяє досягти змінного порядку пам'яті моделі.

– Виходячи з періодограм встановлено відсутність періодичних процесів в дохідності ПФТС. Періодичні процеси дохідності відсутні в базовій моделі Блека-Шоулза, і модифікація моделі не призводить до появи періодичності.

– Індекс ПФТС не є випадковим блуканням чи мартингалом, що встановлено, спираючись на статистичне тестування співвідношення дисперсій. У запропонованій моделі замість випадкового блукання використовується блукання з поверненням до довгострокового стабільного значення, що забезпечується рівнянням Вашичека для дохідності.

Волатильність:

– Волатильність ПФТС є випадковим процесом, а не константою. Для врахування цього до розробленої моделі включається додаткове рівняння стохастичної волатильності.

– Волатильність ПФТС має довгострокову пам'ять, що підтверджено автокореляційною та частково автокореляційною функцією. Мультифрактальне рівняння волатильності дозволяє реалізувати довгострокову пам'ять.

– Функція щільності ймовірності волатильності має щільність розподілу, відмінну від логнормального. У розглянутій моделі щільність ймовірності волатильності розподілена не за логнормальним законом.

Калібрація моделі. Перед початком розв'язку системи (3) необхідно провести калібрацію другого та третього рівнянь системи. Мультифрактальне рівняння волатильності калібрується узагальненим методом моментів, який описаний у [16].

Калібрація розширеної моделі Вашичека здійснюється згідно з [1, С.658]. Функція $\theta(t)$ вибирається рівною поточній кривій відсоткової ставки. Форму кривої відсоткової ставки зазвичай знаходять методом бутстрепінга цін державних облігацій [1].

Подальше підвищення точності моделі може бути досягнуте завдяки використанню моделі jump-дифузії, що дозволить змоделювати стрибки волатильності. Це зробить модель поведінки вартості ближчою до реального і дозволить врахувати можливі наслідки криз.

Висновки

Застосування прикладного системного аналізу та системної інженерії проектів інформатизації організаційних систем до розробки компонентної моделі дозволяє розробити оптимальну архітектуру компонента оцінювання вартості ПФІ.

Компонента модель оцінювання вартості ПФІ дозволяє формалізувати архітектуру та топологію

сутностей інформатизації бізнес-процесів оптимізації та оцінки ризиків інвестиційного портфеля, що містить ПФІ.

Проведено аналіз існуючих рішень оцінювання вартості похідних фінансових інструментів, оцінено ефективність практичного застосування цих методів. Запропоновано вдосконалений метод оцінювання вартості ПФІ – модель SMMVM – дозволяє підвищити точність оцінювання вартості ПФІ у порівнянні з більшістю основних методів, зокрема, біноміальною та тринomialною моделями, методом Блека-Шоулза та методом Монте-Карло.

Розроблена модель уникає ряду припущень, притаманних іншим методам, які зазвичай не виконуються на практиці, зокрема припущення, що вартість базових ФІ є броунівським рухом з постійною волатильністю. Визначено напрямки подальшого вдосконалення моделі для підвищення точності оцінювання вартості похідних фінансових інструментів.

Практична цінність дослідження полягає в тому, що запропонований метод оцінювання вартості ПФІ дозволяє оцінити їх вартість точніше, ніж у ряді інших методів. Компонент оцінювання вартості ПФІ здатен оцінювати не лише поточну вартість похідних фінансових інструментів, але і прогнозувати майбутню вартість у разі наявності прогнозів для базових фінансових інструментів, що мають бути надані компонентом прогнозування системи.

Список літератури: 1. Hull J. C. Options, Futures, and Other Derivatives [Текст] / Hull J. C. // Prentice Hall.— 2005. — P. 815. 2. Згуровський, М.З. Системний аналіз: проблеми, методологія, застосування [Текст] / М.З. Згуровський, Н.Д. Панкратова. — К.: Наукова думка, 2005. — 743 с. 3. Маслянюк, П.П. Концепція інформатизації корпоративних структур [Текст] / П.П. Маслянюк // Науковій вісті НТУУ «КПІ». — 2001. — № 3 — С. 45-50. 4. Маслянюк, П.П. Основні положення методологій системного проектування інформаційно-комунікаційних систем [Текст] / П.П. Маслянюк // Науковій вісті НТУУ «КПІ». — 2007. — № 6 — С. 54-60. 5. Маслянюк, П.П. Компонентні процеси системного проектування інформаційно-комунікаційних систем [Текст] / П.П. Маслянюк // Науковій вісті НТУУ «КПІ». — 2008. — № 2. — С. 112–121. 6. Маслянюк, П.П. Компонентна модель інформаційно-аналітичної системи та генетичний алгоритм формування оптимального портфеля акцій [Текст] / П.П. Маслянюк, А.В. Рябушенко // Науковій вісті НТУУ «КПІ». — 2009. — № 1. — С. 36-46. 7. Bank for International Settlements. Semiannual OTC derivatives statistics [Електронний ресурс]. — Режим доступу: <http://www.bis.org/statistics/derstats.htm>. 8. Bank for International Settlements. Statistics on exchange traded derivatives [Електронний ресурс]. — Режим доступу: <http://www.bis.org/statistics/extderiv.htm>. 9. Wilmott P. Paul Wilmott on Quantitative Finance [Текст] / Wilmott P. // John Wiley & Sons, Ltd.— 2006. — P. 1380. 10. Cox J.C. Options Pricing: A Simplified Approach [Текст] / Cox J.C., Ross S.A., Rubinstein M. // Journal of Financial Economics — 1979 — №9. — P. 229–263. 11. Black F. The Pricing of Options and Corporate Liabilities [Текст] /

Black F., Scholes M. // Journal of Political Economy. — 1973. — №3. — P. 637–654. 12. Ericsson H.E. Business Modeling with UML: Business Patterns at work [Текст] / Ericsson H. E., Penker M.— Wiley Computer Publishing — 2000 — 350 p. 13. Рябушенко, А.В. Компонент оцінювання вартості похідних фінансових інструментів на фондовому ринку [Текст] / А.В. Рябушенко, М.В. Козленко // Системний аналіз та інформаційні технології: матеріали 12-ї Міжнародної науково-технічної конференції SAIT 2010, 25–29 трав. 2010 р. : зб. тез. доп. — К.: ННК «ІПСА» НТУУ «КПІ», 2010. — С. 148. 14. Маслянюк, П.П. Методи оцінювання вартості похідних фінансових інструментів [Текст] / П.П. Маслянюк, А.В. Рябушенко, М.В. Козленко // Прикладна математика та комп'ютеринг ПМК-2010: наук. конф. магістрантів та аспірантів, 14–16 квіт. 2010 р. : зб. тез. доп. — К.: НТУУ «КПІ», 2010. — С. 60–63. 15. Chen L. Stochastic Mean and Stochastic Volatility — A Three Factor Model of the Term Structure of Interest Rates and its Application to the Pricing of Interest Rate Derivatives [Текст] / Chen L. // Financial markets, institutions, and instruments. — 1996. — № 5. — P.1–88. 16. Lux T. The Markov-switching multifractal model of asset returns: GMM estimation and linear forecasting of volatility [Текст] / Lux T. // Journal of Business and Economic Statistics. — 2008. — № 2. — P.194–210. 17. Calvet L. Forecasting multifractal volatility [Текст] / Calvet L., Fisher A. // Journal of econometrics. — 2001. — № 1. — P.27–58. 18. Vasicek O. An Equilibrium Characterisation of the Term Structure [Текст] / Vasicek O. // Journal of Financial Economics. — 1977. — №5. — P. 177–188. 19. Маслянюк, П.П. Створення компонента стратегічного планування системи управління фінансово-інвестиційною діяльністю [Текст] / П.П. Маслянюк, А.В. Рябушенко // Науковій вісті НТУУ «КПІ». — 2009. — № 4. — С. 53–65.

Надійшла до редколегії 18.05.2010.

УДК 004.942

Компонентная модель оценивания стоимости производных финансовых инструментов на фондовом рынке / П.П. Маслянюк, А.В. Рябушенко, М.В. Козленко // Бионика интеллекта: науч.-техн. журнал. — 2010. — № 3 (74). — С. 19–25.

Целью работы есть разработка компонентной модели оценивания стоимости производных финансовых инструментов (ПФИ) в системе управления финансово-инвестиционной деятельностью.

Предложена модель компонента и усовершенствованная модель оценивания стоимости ПФИ в составе пакета математических методов, которые позволяют точнее оценить стоимость ПФИ.

Ил. 3. Библиогр.: 19 назв.

UDK 004.942

Component model of derivatives pricing on exchange market / P.P. Maslyanko, A.V. Riabushenko, M.V. Kozlenko // Bionics of Intelligence: Sci. Mag. — 2010. — № 3 (74). — P. 19–25.

The objective of the paper is the development of pricing derivatives component model as part of the system of financial investment management.

The component model and the improved model of derivatives pricing are proposed as part of analytical package of mathematical methods that allow getting more accurate estimates.

Fig. 3. Ref.: 19 items.

УДК 519.8



МЕТОДЫ И ИНСТРУМЕНТАЛЬНЫЕ СРЕДСТВА СИСТЕМ ПОДДЕРЖКИ ПРИНЯТИЯ РЕШЕНИЙ ПРИ ОРГАНИЗАЦИОННОМ УПРАВЛЕНИИ СОЦИАЛЬНО-ЭКОНОМИЧЕСКИМИ СИСТЕМАМИ

Э.Г. Петров¹, Е.В. Губаренко²^{1,2} ХНУРЭ, г. Харьков, Украина, ST@kture.kharkov.ua

В статье проанализированы особенности проблемы организационного управления сложными активными системами как интеллектуального процесса принятия решений. Показаны пути ее решения на основе методологии экспертного оценивания. Рассмотрен подход к созданию и развитию методов принятия решений, ориентированных на создание систем поддержки принятия решений при организационном управлении и сформулированы требования к такой системе.

АКТИВНЫЕ СИСТЕМЫ, СИСТЕМА ПОДДЕРЖКИ ПРИНЯТИЯ РЕШЕНИЙ, ЭТАПЫ ПРИНЯТИЯ РЕШЕНИЙ, ЭВРИСТИЧЕСКИЕ МЕТОДЫ, ЧАСТИЧНО-ФОРМАЛИЗОВАННЫЕ МЕТОДЫ, КОМПАРАТОРНАЯ ИДЕНТИФИКАЦИЯ

Введение

С точки зрения системного анализа, любая система S может быть представлена, как декартово произведение (упорядоченное множество) элементов (M), отношений (R) и свойств (P) [1]:

$$S = M \times R \times P. \quad (1)$$

Упорядоченное множество элементов и отношений образует структуру:

$$C = M \times R, \quad (2)$$

которая, в свою очередь, порождает свойства системы

$$P = F(M \times R), \quad (3)$$

где F — оператор, отражающий зависимость свойств от множества элементов и отношений.

Выделяют естественные (нецеленаправленные) и искусственные (целенаправленные системы). Естественные (природные) системы представляют собой устойчивое структурированное объединение элементов, зачастую находящихся в стационарном (равновесном) состоянии. Они сформировались из первичного хаоса в результате естественных процессов, которые базируются на фундаментальных законах, определяющих множество возможных устойчивых состояний (структур) и их свойства. Естественные системы образуют объективную среду обитания человечества.

Изучение естественных систем (1) позволяет идентифицировать фундаментальные закономерности существования устойчивых структур (2) и порождения свойств (3), и на основе этих знаний целенаправленно создавать искусственные системы, обладающие свойствами, обеспечивающими достижения требуемых целей.

В свою очередь искусственные системы могут быть разделены на два класса: пассивные и активные. Отличительная особенность пассивных заключается в том, что все элементы таких систем проявляют следующие свойства: элемент не обла-

дает свободой выбора, не способен к волеизъявлению, не имеет собственных эндогенных целей. Чаще всего к пассивным относят технические системы. В активных системах часть или все элементы являются активными целеустремленными индивидами, обладающими свободой воли, способностью выбирать поведенческий сценарий и проявлять осознанный, научно- и морально-обоснованный выбор. Таким образом, можно выделить автоматические (все элементы пассивны), автоматизированные (часть элементов активна) и социальные, в которых все элементы активные, системы.

Приведенная классификация является достаточно условной, так как в реальных условиях большинство систем смешанные, отличающиеся лишь долей активных и пассивных элементов. Примером смешанной системы могут служить социально-экономические системы (СЭС) различного уровня, элементами которых являются социальные (активные), экономические (автоматизированные) и природно-экологические (пассивные) системы.

Все искусственные системы являются целенаправленными, поэтому для любой из них обязательной является функция управления достижением цели.

В настоящее время создана достаточно совершенная математическая теория оптимального автоматического управления для пассивных систем с экзогенно заданными целями.

Задание экзогенной цели пассивной системы означает, что задано некоторое целевое состояние. В исходный момент времени система находится в некотором начальном состоянии. Для перевода из начального в целевое (конечное, промежуточное) состояние необходимо затратить ресурсы (временные, энергетические, материальные). В этих условиях можно синтезировать некоторое правило (закон, алгоритм) управления, определяющее траекторию изменения состояния. Синтезированное правило должно обеспечивать достижения цели с требуемой

точностью и минимизацией затрат ресурсов. Такая постановка не исключает возможность адаптации закона управления к изменению начальных или целевых (конечных) состояний, но по установленному правилу, определенному заранее.

Теоретической основой автоматического управления является строго доказанное утверждение о существовании и единственности оптимальной траектории достижения цели.

Вместе с этим существует большой класс активных организационных систем, для которых цели управления тоже заданы (экзогенно или эндогенно), но в силу их сложности и несогласованности, влияния большого количества НЕ-факторов и активности элементов, существует не единственное решение, а некоторое множество альтернативных способов достижения цели, наилучший из которых не очевиден. Это порождает проблему выбора решения (способа достижения цели). Таким образом, проблема организационного управления заключается в решении двух задач: принятие эффективного решения и его реализация (управление) [1, 2].

В настоящее время ведутся интенсивные исследования по созданию общей теории принятия решений. В этой области получен ряд фундаментальных результатов, однако в целом проблема далека от исчерпывающего решения. Кроме того, конструктивное решение любой конкретной задачи связано с необходимостью дальнейшего развития, адаптации и конкретизации общих подходов.

Целью данной статьи является анализ и развитие методологии решения актуальных задач организационного управления СЭС, в частности создания автоматизированных систем поддержки принятия управляющих решений, как концептуального этапа управления.

1. Основные условия эффективности принятия организационного решения

Процесс принятия решений является интеллектуальной процедурой, т.е. осуществляется с участием человека. Это означает, что система поддержки принятия решений (СППР) является автоматизированной системой, распределение функций в которой между человеком и ЭВМ зависит от класса объектов управления, их сложности, уровня знаний о них, степени формализации и т.д.

Основной целью создания СППР является повышение эффективности организационного управления. В.М. Глушков [3] сформулировал следующие необходимые условия эффективности организационного управления: своевременность, полнота (комплексность), оптимальность принимаемых решений.

Требования своевременности означают необходимость учета инерционности объекта управления, т.е. наличие лага (запаздывания) между действием

управляющего воздействия и реакцией на него объекта, временной координации взаимодействующих систем. В большинстве случаев это требование выступает как внешнее ограничение на допустимое время принятия и реализации решений.

Условие полноты подчеркивает тот факт, что потенциальная эффективность решений в определенной степени зависит не только от характеристик элементов, отношений и уровня их детализации внутри системы, но и характера взаимодействия со смежными системами и внешней средой. При этом необходимо учитывать, что понятие полноты модели является неисчерпаемым, его следует рассматривать как компромисс между возможным повышением эффективности, с одной стороны, и сложностью модели, уровнем наших знаний, точностью измерений и идентификации параметров и отношений – с другой.

Долгие годы в основе управления СЭС лежала концепция экономического роста, ориентированная на количественный рост экономических показателей: увеличение объемов производства и потребления, максимизацию прибыли и темпов роста ВВП [4]. Эта концепция сыграла свою роль в становлении и развитии индустриального общества. Но в настоящее время возникла и осознана угроза глобальных климатических изменений, истощения и удорожания природных ресурсов, необратимой деградации среды обитания, учащающиеся кризисы локального и глобального характера и связанный с этим рост социальной напряженности. Решение большинства из этих проблем невозможно на национальном уровне, необходима мировая координация, повышение роли и глубины государственного регулирования.

В целом все перечисленные факторы обуславливают необходимость перехода к новой концепции управления, более полно и комплексно учитывающей экономические, социальные и экологические характеристики СЭС и их взаимосвязь. Такая концепция была продекларирована на заседании ООН в 1992 году и получила название концепции устойчивого развития. Однако ее практическая реализация требует как учета этико-моральных норм и бизнес-особенностей, так и развития комплекса формальных математических моделей, инструментальных средств и методологии управления [5]. Переход к любой более полной модели приводит к ее усложнению, росту размерности, необходимости учета нелинейностей, многокритериальности и увеличению неопределенности исходных данных и, как результат, к усложнению процесса принятия решений. Перечисленные задачи являются ключевыми при практической реализации концепции устойчивого развития.

Проблема разработки системы моделей устойчивого развития СЭС является многоаспектной.

Это наглядно видно на примере удовлетворения требования оптимальности принимаемых решений.

С ростом масштабов и значимости принимаемых решений, повышаются требования к их оптимальности. Интуитивные, удовлетворительные, «рациональные» решения оказываются малопримлемыми, так как для большемасштабных, многосвязных систем отклонение от оптимального решения даже на малые доли процента приводит к большим потерям или упущенным возможностям. Это означает, что необходимо определять строгого оптимальные формальные решения x^0 :

$$x^0 = \operatorname{argextr}_{x \in X} K(x), \quad (4)$$

где X – область допустимых решений; $K(x)$ – целевой функционал (критерий эффективности) принимаемого решения.

Пусть каждое решение $x_i \in X, i = \overline{1, n}$ характеризуется кортежем значений $\langle x_{i,j} \rangle, j = \overline{1, m}$, который достаточно полно описывает качественные и количественные значения элементов и отношений системы, т.е. ее состояние. Тогда степень достижения цели (совершенство системы) характеризуется значением скалярной целевой функцией (критерием)

$$K(x_i) = F(x_{ij}),$$

а задача (4) является классической задачей математического программирования [6], типичной для концепции экономического роста.

Для СЭС, в рамках концепции устойчивого развития, характерно наличие некоторого множества несогласованных локальных целей, уровень достижимости которых характеризуется кортежем разнородных по размерности, направлению доминирования, интервалу возможных значений частных критериев

$$\begin{aligned} K(x_i) &= \langle k_l(x_i) \rangle, l = \overline{1, L}, \\ k_l(x_i) &= f_l(x_i) l = \overline{1, L}. \end{aligned} \quad (5)$$

С учетом этого, модель определения оптимального решения (4) превращается в задачу многокритериальной оптимизации (МО)

$$x^0 = \operatorname{argextr}_{x \in X} \langle k_l(x) \rangle, \forall l = \overline{1, n}. \quad (6)$$

Как известно, решение задач МО зависит от структуры допустимого множества решений X [2]. В общем случае множество X представляет собой объединение двух подмножеств

$$X = X^S \cup X^C; X^S \cap X^C = \emptyset,$$

где X^S – подмножество доминируемых (согласованных) решений, т.е. таких, для каждого из которых на множестве X существует решение, лучшее по всем или части показателей при равенстве всех

остальных; X^C – подмножество недоминируемых, т.е. не принадлежащих X^S , решений.

Если $X^C \neq \emptyset$, задача МО является некорректной по Адамару [7], так как не имеет единственного решения. В этом случае можно определить только подмножество $X^C \subset X$, которое часто называют Парето-оптимальным решением. Для практической реализации необходимо единственное решение $x \in X^C$. Выбрать такое решения возможно двумя способами:

- на основе эвристических соображений лица, принимающего решения (ЛПР);
- с помощью некоторой нормативной процедуры.

В первом случае ЛПР (менеджер, эксперт, группа лиц), которому делегированы соответствующие полномочия, на основе знаний, опыта, известных прецедентов, т.е. эвристических соображений, осуществляет выбор единственного решения $x \in X^C$. Во втором случае единственное решение из X^C выбирается на основе некоторого нормативного правила. Такое правило называется регуляризирующим, а его выбор и обоснование – процедурой регуляризации некорректной задачи [8].

Общая идея нормативной регуляризации задачи МО заключается в трансформации некорректной задачи в условно корректную путем ее скаляризации, то есть превращении многокритериальной задачи в задачу однокритериальной, скалярной оптимизации или в последовательность таких задач.

В настоящее время известно большое количество регуляризирующих правил задачи МО, например, принцип главного критерия, схема последовательной (лексикографической) оптимизации, функционально-стоимостной анализ [2]. Однако детальный анализ этих различных регулирующих правил показывает, что все они базируются на общей теории полезности [9]. Согласно этой теории, если заданы два решения $x_1, x_2 \in X$ и на них определено отношение качественного предпочтения, при этом каждое из решений описывается многокритериальным кортежем разнородных характеристик $K(x_1) = \langle k_l(x_1) \rangle, K(x_2) = \langle k_l(x_2) \rangle$, то существует такой обобщенный скалярный показатель $P(x)$ (оценка полезности решения), для которого выполняется соотношение

$$x_1 \succ x_2 \Leftrightarrow P(x_1) > P(x_2).$$

Справедливо и обратное утверждение.

Значение полезности решений описывается функцией вида

$$P(x) = F[A, K(x)], \quad (7)$$

где F – оператор, определяющий вид (структуру) функциональной зависимости (модели); A – параметры модели; $K(x) = \langle k_1(x), k_2(x), \dots, k_L(x) \rangle$ – кортеж локальных критериев конкретного решения $x \in X$.

Теория полезности и, соответственно, функция полезности (7) могут рассматриваться не только с позиции решения задач многокритериальной оптимизации, но и как инструмент определения многофакторных скалярных оценок состояния активных систем и решения на этой основе задач многофакторного сравнения, ранжирования, классификации, кластеризации и других прикладных задач анализа активных систем, в частности СЭС. Для этого необходимо в модели (7) заменить кортеж локальных критериев решений на кортеж, характеризующий состояния СЭС любого уровня – локального или глобального.

При этом задача синтеза локальных моделей (5) является самостоятельной задачей, которую необходимо решать в рамках конкретной проблемы, поэтому она в данной статье не рассматривается.

Основной проблемой конструктивного применения теории полезности является синтез универсальной адаптивной модели многофакторного скалярного оценивания (7). Проблема синтеза любой модели связана, в общем случае, с необходимостью решения задачи ее структурно-параметрической идентификации. Структурная идентификация состоит в определении вида оператора F , а параметрическая – в определении конкретных численных значений параметров A . Кортеж характеристик $K(x)$ является входом модели, характеризующим «качество» решения или «состояние» системы.

Точность и корректность принятого решения, описанным выше методом, зависят от адекватности математической модели многофакторного скалярного оценивания (7). Адекватность любой математической модели определяется точностью решения задач ее структурно-параметрической идентификации.

В настоящее время получила широкое развитие теория идентификации пассивных, технических, технологических, экономических систем, удовлетворяющих условиям наблюдаемости, управляемости и измеримости. Путем пассивных или активных экспериментов (наблюдений) на некотором отрезке времени измеряют синхронизированные по времени значения входных воздействий и реакций системы. Такая последовательность называется временным рядом [10]. Обработка и анализ временной последовательности одним из многочисленных проблемно ориентированных методов, позволяет определить (аппроксимировать) характер зависимости между входом и выходом (структуру) и численные значения параметров модели. Критерием точности и адекватности синтезированной (идентифицированной) модели является близость выхода (реакции) модели и реальной системы на одинаковые входные воздействия.

Модель скалярного многофакторного оценивания является моделью интеллектуального процес-

са выбора, и в силу ограниченности наших знаний в настоящее время невозможно получить временной ряд для любой интеллектуальной деятельности. Можно классическими методами измерять значения входного сигнала, но невозможно прямыми измерениями получить значения реакции системы, то есть значение полезности $P(x)$. Это делает невозможным использование классической теории идентификации и обуславливает необходимость разработки альтернативного подхода.

В качестве альтернативы в настоящее время используется интроспективный подход, основанный на способности индивидуума к самонаблюдению, анализу процессов мышления, их структуризации, оцениванию. Методология практической реализации интроспективного подхода известна как экспертное оценивание.

Методология экспертного оценивания основана на гипотезе, что усреднение субъективных индивидуальных мнений нескольких экспертов позволят получить оценку, приближающуюся к объективной. Для повышения точности таких оценок, обеспечения их устойчивости, воспроизводимости, устранения внешних влияний (авторитарных воздействий, ангажированности) разработано множество правил (технологий) организации, проведения, обработки результатов экспертного оценивания, которые находятся в распоряжении когнетолога (специалиста по извлечению знаний экспертными методами).

Большинство из этих правил имеют проблемную ориентацию и выбираются на основе анализа целей, которые должны быть достигнуты в процессе экспертизы. Можно выделить три группы таких целей:

- непосредственное решение некоторой уникальной проблемы;
- получение некоторой, необходимой для принятия решения с помощью формальных математических процедур информации, которая не может быть получена объективными методами прямого измерения (метод анализа иерархий, ЭЛЕКТРА, оценки общей полезности и другие человеко-машинные методы);
- получение информации для структурно-параметрической идентификации нормативных моделей принятия решений.

2. Этапы принятия решений

Общей для достижения всех перечисленных ранее целей является процедура выбора, то есть принятия решений. Плодотворность использования метода экспертного оценивания как инструментального средства поддержки принятия решений в значительной степени зависит от четкого формулирования конечной цели и постановки промежуточных задач экспертизы. С этой точки зре-

ния процесс принятия решений независимо от его проблемной ориентации можно структурировать на следующие этапы и решаемые задачи:

- формулировка цели, которая должна быть достигнута;
- выделение множества возможных и допустимых решений;
- определение метрики (системы критериев), в которой производится сравнение допустимых решений;
- выбор эффективного решения.

Охарактеризуем каждый из этапов на примере проблемы принятия решений при организационном управлении СЭС.

Формирование цели. Выбор и постановка целей является принципиально важным условием начала процедуры принятия решений. Максимально конкретное описание требований к конечному результату повышает вероятность принятия эффективного и адекватного решения. Цели для СЭС (активных, динамически развивающихся систем) непрерывно изменяются и уточняются как в результате собственного развития, так и взаимодействия с внешней средой. Поэтому цели должны быть согласованы с целями СЭС как более высокого, так и более низкого, уровня; учитывать интересы составляющих элементов и подсистем; формироваться на основе достигнутых ранее результатов (знаний, научно-технического потенциала, инфраструктуры, производственных мощностей); иметь возможность изменяться соответственно требованиям элементов СЭС либо изменяющемуся характеру взаимодействия с внешней средой; предусматривать возможность инноваций и модернизации ресурсного потенциала (т.е. совокупность природных, материальных, трудовых, финансовых и информационных ресурсов) и достижимости глобальной цели.

Формирование целей – крайне ответственный, сложный и трудоёмкий этап, поэтому на ранних стадиях анализа цели формируются на интуитивном уровне, и большую часть требуемой аналитической работы перекладывают на последующие этапы.

Вторым этапом является формирование области возможных вариантов решения (X^B). Базируется данный этап на выделении как можно более полного множества теоретически обоснованных альтернатив, практических прецедентов путем обобщения опыта и знаний экспертов, интуиции и творческой фантазии членов рабочей группы. Рабочая группа – определенное количество экспертов, работающих в непосредственном контакте, либо отдельно друг от друга, решающих общую задачу заранее определенным методом. Информация, получаемая на данном этапе, имеет достаточно «распыленную» форму, она может затрагивать

различные области знаний и различные формы реализации определенной цели. Требования, которые выдвигает данный этап к методам решения, следующие:

- обеспечить эффективное взаимодействие экспертов различных областей знаний, так как СЭС одновременно включает достаточно большое количество областей наук и отраслей производства;
- собрать максимально возможный объем знаний, мнений, предположений и альтернатив, так как именно на этапе формирования области возможных вариантов решения выделяются альтернативы, среди которых, предположительно, находится «верное» решение. И чем шире и разнообразней альтернативы, тем выше вероятность включения действительно эффективных решений;
- систематизировать и упорядочить альтернативы, без критического оценивания каждой, частичное упорядочивание приведет к упрощению работы при дальнейшей обработке альтернатив;
- создать условия креативности и неординальности предлагаемых альтернатив, так как именно оригинальные, нестандартные решения наиболее перспективны.

Результатом выполнения данного этапа является множество возможных вариантов решений, охватывающих как можно большую область и по возможности максимально приближенных к достижению поставленной цели. Следующий шаг этапа состоит в выделении из области возможных вариантов решения (X^B) области допустимых решений (X^D) путем отсечений, не удовлетворяющих достижению поставленной цели и нереализуемых по требуемым временным, энергетическим, материальным, финансовым, трудовым, экологическим и другим ресурсам, альтернатив. Суть данной процедуры заключается в критическом рассмотрении выделенных вариантов решения, оценке соответствия результата решения поставленной цели и возможности реализации данного решения. Требования, которые выдвигаются на данном шаге к методам анализа, следующие:

- каждая альтернатива должна критически рассматриваться экспертом либо группой экспертов, которые специализируются в соответствующих областях знаний и непосредственно обладают знаниями о возможности реализации рассматриваемого решения и необходимых для этого ресурсах;
- принимаются только те альтернативы, которые могут быть реализованы, и результат выполнения принимаемого решения будет соответствовать либо близок к поставленной цели с заданной степенью допущения.

Результатом является область допустимых решений, каждое из которых может быть реализовано (хотя бы в теории), и результат будет соответствовать цели.

Определение критериев оценивания для возможности сравнения альтернатив между собой. В большинстве случаев допустимые альтернативы имеют многокритериальные оценки, которые необходимо учитывать. Результатом выполнения данного этапа должен стать перечень критериев, по которым необходимо оценить каждую альтернативу, приписав ей определенное количественное или качественное значение, обеспечив полноту и всестороннее описание всех рассматриваемых альтернатив средствами выделенных критериев.

Выбор эффективного решения. На этом этапе в зависимости от поставленных целей выбирается решение, которое наиболее эффективно обеспечивает их достижение. Данный этап достаточно хорошо изучен для пассивных систем, но для активных, таких как СЭС, необходимо развитие в рамках теории принятия решения других подходов и инструментальных средств.

3. Классификация методов принятия решений

Основным источником информации о столь сложных системах, как социальные, экономические, экосистемы либо вбирающие в себя признаки сразу нескольких видов системы, являются знания экспертов, которые обобщают фактическую, полученную средствами измерений и наблюдений, эмпирическую, основанную на опыте эксперта, а также интуитивную и субъективную информации. Зачастую даже сугубо экономические системы не могут быть полностью описаны аналитическими моделями, хотя практика анализа хозяйственной деятельности имеет продолжительную и богатую историю применения. Но даже если все элементы системы могут быть измерены, приведены к единой размерности и согласованы друг с другом, именно эксперт выступает заключительным звеном интерпретации полученных результатов. Другими словами, в любых организационных системах: технических, экономических, социальных, экологических, смешанных, — вне зависимости от размерности и вида информации (количественная или качественная) основным источником получения её является эксперт. Изменяются лишь инструментарий измерения и формы представления информации для принятия решений.

С этих позиций выделяют три группы методов принятия решений.

Эвристические методы — методы, основанные на неформализованных правилах непосредственного выбора экспертом варианта решения. Данные методы позволяют:

- в минимальный временной интервал охватить огромные области знаний за счет привлечения различных экспертов;
- найти неординарное решение, не имеющее аналогов и практики применения;

- обеспечить формирование множества возможных вариантов решения, хотя и не все варианты могут быть адекватными и целесообразными;

- такими методами невозможно найти оптимальное решение в меру отсутствия формальной процедуры принятия решения, но решение практически всегда находится достаточно быстро;

- альтернативы очень редко сравнимы между собой, но их разброс позволяет выделять направления, требующие совершенно различных подходов и ресурсов, что может существенно повысить устойчивость системы.

К эвристическим методам относят следующие методы и процедуры.

Метод мозгового штурма. Изобретателем метода мозгового штурма считается Алекс Осборн, сотрудник рекламного агентства BBD&O. Метод мозгового штурма (англ. brainstorming) — оперативный метод решения проблемы на основе стимулирования творческой активности, при котором участникам обсуждения предлагают высказывать как можно большее количество вариантов решения, в том числе самых фантастических. Затем из общего числа высказанных идей отбирают наиболее удачные, которые могут быть использованы на практике. Существует целый ряд разновидностей метода мозгового штурма: метод мозговой атаки, метод коллективной генерации идей, метод совещаний (комиссий, круглого стола), метод суда, метод критической атаки, метод номинальных групп, метод «635», метод синектики [11], метод интеграции решений.

Помимо метода мозгового штурма выделяют следующие методы и процедуры:

- метод деловых игр, основанный на моделировании экспертами функционирования бизнес-процессов и систем управления при выполнении операций, направленных на достижение поставленной цели;

- метод коллективного блокнота — на начальном этапе группе экспертов ставят задачу, затем идет процесс индивидуальной работы эксперта со своим блокнотом в течение определенного времени, на следующем этапе блокноты собираются, полученная информация систематизируется и, после этого, эксперты-участники переходят к формированию решения проблемы в живом совместном обсуждении накопленного и обобщенного материала;

- метод экспертного фокусирования — это форма очного коллективного обсуждения обозначенной проблемы, эксперты-участники тщательно исследуют заданную ситуацию под разным ракурсом, стараются всецело сфокусироваться на ней.

- метод комиссий — этот метод заключается в коллективном обсуждении заданной проблемы, основным отличием от метода фокусирования

является то, что члены экспертной команды стремятся уяснить и сформулировать причину противоречия между разными вариантами предлагаемых решений, найти как можно большее количество объединяющих моментов и прийти к решению, приемлемому для всех членов группы;

– метод «консилиум» – при применении данного метода эксперты копируют форму исследования проблемы у врачей (воспринимая ситуацию как заболевание): классифицируются признаки (симптомы), выявляются причины, приведшие к данной проблеме, проводится анализ, ставится окончательный диагноз и формулируются прогнозы развития проблемы;

– методы типа сценариев – подобные методы предполагают подготовку и согласование представлений о проблеме либо анализируемом объекте, изложенных в письменном виде, называемых сценариями;

– метод индивидуального блокнота – данный метод проведения анализа исключает общение эксперта-участника с исследователями, которые будут проводить опрос;

– метод дискуссии – это один из методов опроса экспертов, представляющий собой относительно свободный диалог между исследователем и экспертом или экспертами на определенную тему, т.е. метод получения информации на основе вербальной коммуникации [12];

– метод групповой дискуссии [12] – члены рабочей группы общаются между собой (ведут спор) и в ходе этого особым образом организованного общения разрешают поставленную проблему;

– методы анкетирования – это проведение опроса в письменной форме, для этого используется набор структурно организованных вопросов (анкета) [12,13];

– методы интервьюирования – это метод получения необходимой информации путем непосредственной целенаправленной беседы в форме «вопрос-ответ»;

– метод Дельфы – это метод группового анкетирования [12,13], используемые процедуры характеризуются тремя основными чертами: анонимностью, регулируемой обратной связью и групповым ответом;

– метод эвристических вопросов – этот метод известен также как метод «ключевых вопросов» и «наводящих вопросов» [13], этот метод целесообразно применять для сбора дополнительной информации в условиях проблемной ситуации или упорядочения уже имеющейся информации в самом процессе решения задачи;

– метод свободных ассоциаций – в процессе зарождения ассоциаций устанавливаются неординарные взаимосвязи между компонентами решаемой проблемы и элементами внешнего мира,

включая компоненты прежнего опыта творческой деятельности лиц, участвующих в коллективном решении проблемы, творческой задачи.

– метод инверсии – представляет собой один из эвристических методов творческой деятельности, ориентированный на поиск идей решения творческой задачи в новых, неожиданных направлениях, чаще всего противоположных традиционным взглядам и убеждениям, которые диктуются формальной логикой и здравым смыслом;

– метод эмпатии (метод личной аналогии) – в решении творческих задач используют различные аналогии: конкретные и абстрактные; ведутся поиски аналогии живой природы с неживой, например в области техники, в этих последних аналогиях могут быть, в свою очередь, установлены аналогии по форме, структуре, функциям, процессам и т.д.;

– метод Борда – процедура голосования для выбора из множества предложенных альтернатив [14,15]. Согласно этому методу результаты голосования выражаются в виде числа баллов, набранных каждой из альтернатив;

– метод Кондорсе – эксперты упорядочивают альтернативы, устанавливая предпочтения в виде цепочки $0_1 \rightarrow 0_2 \rightarrow \dots \rightarrow 0_n$, побеждает та альтернатива, которая побеждает при сравнении один на один с любой из других альтернатив.

Отдельно необходимо выделить методы сравнения и получения количественных оценок.

Ранжирование. Метод представляет собой процедуру упорядочения объектов, выполняемую экспертом. На основе знаний и опыта эксперт располагает объекты в порядке предпочтения, руководствуясь одним или несколькими выбранными показателями сравнения.

Метод парных сравнений [15]. Эксперту предлагают сравнить пары альтернатив, основным недостатком является возможная непоследовательность в ответах и связанное с этим нарушение транзитивности. Эксперты способны принять правильное решение путём попарного сравнения альтернатив без их предварительной структуризации, установив отношения строгого или не строгого порядка.

Множественные сравнения. Они отличаются от парных тем, что экспертам последовательно предъявляются не пары, а n -мерные множества объектов, где n принимает значения $[3, +\infty)$. Эксперт их упорядочивает по важности или разбивает на классы.

Метод гиперупорядочения предполагает рассмотрение не только факторов, но и разностей оценок факторов и их ранжирование. Основным недостатком данного метода является большой объем информации, предоставляемый эксперту для анализа и сравнения.

Метод кортежей предпочтений. Эксперту предъявляется множество частных критериев

$k_1(x), \dots, k_n(x)$ и для каждого $k_i(x), i = \overline{1, n}$, и он должен указать число критериев, более значимых, чем данный, не указывая при этом, какие критерии являются более предпочтительными. Таким образом, каждый критерий получает определённый ранг. Используя отдельно взятый критерий как эталон, эксперт проводит парные сравнения, определяя предпочтение.

Метод классификации. Целью обращения к эксперту является разбиение альтернатив или объектов на классы. Эксперту предлагается соотнести объект к какому-либо классу, сравнив его с эталонными значениями класса и отнеся его к наиболее подходящему. Виды классификаций различны и применяются в соответствии с видом решаемых задач.

Метод непосредственного оценивания. Метод заключается в присваивании объектам числовых значений в шкале интервалов. Эксперту необходимо поставить в соответствие каждому объекту точку на определенном отрезке числовой оси. При этом необходимо, чтобы эквивалентным объектам приписывались одинаковые числа.

Метод шкалы отношений. Измерения в шкале интервалов могут быть достаточно точными при полной информированности экспертов о свойствах объектов. Эти условия на практике встречаются редко, поэтому для измерения применяют относительную балльную оценку. При этом вместо непрерывного отрезка числовой оси рассматривают участки, которым приписываются баллы. Эксперт, приписывая объекту балл, тем самым измеряет его с точностью до определенного отрезка числовой оси.

Метод средней точки – метод непосредственной оценки альтернатив, пошагово эксперт указывает середину выделенного интервала, с каждым шагом уточняя решение.

Метод Чёрчмена-Акофа относится к числу наиболее популярных [16]. В этом методе предполагается последовательная корректировка оценок, указанных экспертами таким образом, чтобы предпочтительный элемент имел оценку выше, чем суммарный вес оценок остальных альтернатив.

Метод парных сравнений Терстоуна [17]. Парно сравниваются объекты, проходя все возможные ситуации. Результаты сравнения заносятся в таблицу. После этого по каждой строке подсчитываем число единичек, т.е. сколько раз сравнение было успешным. После составляются ряды предпочтений для каждого эксперта. Далее рассчитываются общие показатели по всем экспертам: число предпочтений и число предпочтений, приходящихся на одного эксперта. На основании полученной информации можно определить, во сколько и насколько один объект предпочтительнее другого.

Методы средних баллов. В настоящее время распространены экспертные, маркетинговые, ква-

зиметрические, социологические и иные опросы, в которых опрашиваемых просят выставить баллы объектам, изделиям, технологическим процессам, предприятиям, проектам, заявкам на выполнение научно-исследовательских работ, идеям, проблемам, программам, политикам и т.п., а затем рассчитывают средние баллы и рассматривают их как интегральные оценки, выставленные коллективом опрошенных [18].

Метод медиан. Суть метода заключается в упорядочивании списка баллов, выставленных экспертами по неубыванию, другими словами – по возрастанию [18, 19]. Выделяется средняя область, и берутся за основу те значения, которые в нее вошли. При этом медианы вычисляются по обычным правилам статистики – как среднее арифметическое центральных членов вариационного ряда.

Метод фон Неймана – Моргенштерна [20]. Он заключается в получении численных оценок альтернатив с помощью так называемых вероятностных смесей.

Морфологические методы известны так же как методы многомерных матриц или морфологический ящик. Основная идея морфологических методов систематически находить все мыслимые варианты решения проблемы или реализации системы путем комбинирования выделенных элементов или их признаков. В систематизированном виде морфологический подход разработан и применен впервые швейцарским астрономом Ф. Цвики. Выделяют следующие разновидности: метод систематического покрытия поля (МСПП), метод отрицания и конструирования (МОК), метод морфологического ящика (ММЯ).

Частично формализованные методы – методы, обладающие четко структурированной последовательностью принятия решений, но использующие экспертные оценки как метод получения недостающей информации. Методы, имеющую строго регламентированную последовательность, ограничивают экспертов при генерации альтернатив, следовательно, объем возможных решений значительно меньше, нежели при применении эвристических методов. Строгая последовательность действий ускоряет нахождение допустимого множества и оценок альтернатив, а, следовательно, нахождения эффективного решения, его устойчивость и воспроизводимость. При применении частично формализованных методов следует выделить следующие особенности:

- область возможных вариантов решения, используя данные методы, сформировать невозможно, она должна быть уже выделенной и отвечать требованиям;

- исходная информация задается средствами экспертных оценок, но имеет строго установленную форму;

– последовательность действий, расчетов и преобразований задана, что позволяет провести частичную автоматизацию процедур и ускорить процесс принятия решения (на основе именно таких методов построены большинство СППР).

Примерами этой группы являются следующие методы.

Метод анализа иерархий (МАИ) – математический инструмент системного подхода к сложным проблемам принятия решений. МАИ не предписывает ЛПР какого-либо «правильного» решения, а позволяет ему в интерактивном режиме найти такой вариант (альтернативу), который наилучшим образом согласуется с его пониманием сути проблемы и требованиями к ее решению. При подходе МАИ усилия ЛПР направляются на сравнение только заданных альтернатив [14,21].

Каждый элемент иерархии может представлять различные аспекты решаемой задачи, причем во внимание могут быть приняты как материальные, так и нематериальные факторы, измеряемые количественные параметры и качественные характеристики, объективные данные и субъективные экспертные оценки. Иными словами, анализ ситуации выбора решения в МАИ напоминает процедуры и методы аргументации, которые используются на интуитивном уровне. Следующим этапом анализа является определение приоритетов, представляющих относительную важность или предпочтительность элементов построенной иерархической структуры, с помощью процедуры парных сравнений. Безразмерные приоритеты позволяют обоснованно сравнивать разнородные факторы, что является отличительной особенностью МАИ. На заключительном этапе анализа выполняется синтез приоритетов на иерархии, в результате которой вычисляются приоритеты альтернативных решений относительно главной цели. Лучшей считается альтернатива с максимальным значением приоритета.

Метод порогов несравнимости (МПН). В конце 60-х годов группа французских ученых во главе с профессором В. Руа предложила новый подход к проблеме принятия решений при многих критериях [14,22]. Метод порогов несравнимости известен как outranking approach (в мировой литературе), Разработка Индексов Попарного Сравнения Альтернатив (РИПСА) – в работах Ларичева [14]. Связь между любой парой альтернатив определяется последовательностью бинарных отношений. Если бинарное отношение является отношением доминирования одной альтернативы над другой, при котором одна альтернатива имеет по всем критериям не худшие, а хотя бы по одному из критериев лучшие оценки, то ядро является множеством Парето. Выделенное множество Парето может использоваться как самостоятельный предваритель-

ный этап в других методах принятия решений. Для методов порогов несравнимости следует выделить два основных этапа:

– этап разработки, на которой строятся один или несколько индексов попарного сравнения альтернатив;

– этап исследования, на котором построенные индексы используются для ранжирования (или классификации) заданного множества альтернатив.

В настоящее время имеется много методов принятия решений, принадлежащих к методам порогов несравнимости. Из них наиболее известна группа методов ELECTRE (Elimination Et Choix Traduisant la Realité – исключение и выбор, отражающие реальность). Существуют ELECTRE I, ELECTRE II, ELECTRE III, ELECTRE IV, ELECTRE IS и ELECTRE A.

Основные этапы метода ELECTRE:

а) На основании заданных оценок двух альтернатив подсчитываются значения двух индексов: согласия и несогласия. Эти индексы определяют согласие и несогласие с гипотезой, что альтернатива А превосходит альтернативу В.

б) Задаются уровни согласия и несогласия, с которыми сравниваются подсчитанные индексы для каждой пары альтернатив. Если индекс согласия выше заданного уровня, а индекс несогласия – ниже, то одна из альтернатив превосходит другую. В противном случае альтернативы несравнимы.

с) Из множества альтернатив удаляются доминируемые. Оставшиеся образуют первое ядро. Альтернативы, входящие в ядро, могут быть либо эквивалентными либо несравнимыми.

д) Вводятся более «слабые» значения уровней согласия и несогласия (меньший по значению уровень согласия и больший уровень несогласия), при которых выделяются ядра с меньшим количеством альтернатив.

е) В последнее ядро входят наилучшие альтернативы. Последовательность ядер определяет упорядоченность альтернатив по качеству

Методы структурно-параметрической идентификации формальных нормативных моделей многофакторного оценивания и многокритериального выбора. Теоретической основой является многокритериальная теория полезности (МКТП) (Multi-Attribute Utility Theory) [9, 20], утверждающая существование скалярной многофакторной оценки (функции полезности) вида (5, 6). Экспертные оценки применяются для решения следующих задач:

– выделения множества допустимых альтернатив (X^D);

– формирования списка локальных критериев (кортежа $K(x)$);

– определения структуры функции полезности (вида оператора F) (6);

– решения задач параметрической идентификации, т. е. экспертного определения значений параметров модели оценивания (A);

Стремление повысить точность решения задачи структурно-параметрической идентификации привело к разработке метода компараторной идентификации [1,4]. В этом случае экспертная процедура используется для в качестве компаратора, устанавливающего отношения строгого или не строгого порядка на множестве допустимых альтернативных решений $X = \{x_i, i = \overline{1, n}\}$,

$$x_1 \succ x_2 \succ \dots \succ x_n,$$

$$x_1 \sim x_2 \succ \dots \succ x_n.$$

По этой исходной информации на основе (5) строится система вида:

$$(E_1) > (E_2);$$

...

$$(E_{n-1}) > (E_n).$$

Откуда следует система ограничений:

$$(E_1) - (E_2) \geq 0;$$

...

$$(E_{n-1}) - (E_n) \geq 0,$$

(8)

определяющая область возможных моделей (структуры и параметров) многофакторного оценивания. На основе (8) формальными методами определяются возможные варианты модели многофакторного скалярного оценивания.

Выводы

Отказ от концепции экономического роста и переход к концепции устойчивого развития приводит, с одной стороны, к резкому усложнению процессов принятия решений за счет увеличения размерности системы, необходимости учета многокритериальности, роста уровня неопределенности исходных данных, а с другой – к повышению требований своевременности и эффективности (оптимальности) принимаемых решений.

Кардинальным направлением решения указанной проблемы является переход от креативных (методы, использующие уникальные, абсолютно новые идеи и подходы для решения одной конкретной задачи) к нормативным, формализованным методам принятия решений, что открывает широкую перспективу автоматизации этих процессов и создания функциональных специализированных СППР.

Однако процесс принятия решений является интеллектуальной процедурой, а теория принятия решений, как научное направление, не обладает необходимыми инструментальными и формальными средствами для непосредственного получения необходимой информации и синтеза моделей, по аналогии с идентификацией моделей натуральных

систем с наблюдаемыми процессами. Для интеллектуальных процессов такую информацию можно получить только опосредовано, используя способность индивида к самонаблюдению, анализу процессов мышления, их структуризации, оцениванию. Методология получения знаний на основе интроспективного подхода известна как экспертное оценивание.

Метод экспертного оценивания играет значительную роль как инструмент решения общей проблемы принятия решений. На первых этапах своего развития он был ориентирован на непосредственное, эвристическое, неформализованное решение проблем принятия решений, а усилия теоретиков были направлены на структурирование процесса, регламентацию процедур получения информации, обоснование методов математической обработки результатов экспертизы с целью обеспечения их сходимости, устойчивости, воспроизводимости. Однако по мере усложнения объектов анализа, роста их размерности, многосвязности, нелинейности, многокритериальности, неполноты исходной информации пришло понимание ограниченности возможности методов, вытекающее из особенностей человеческой психики и характера мышления.

В этих условиях в настоящее время интенсивно развивается два альтернативных направления:

– создание нормативных человеко-машинных процедур принятия решений (методы анализа иерархий (МАИ), порогов несравнимости (ELECTRE) и другие);

– использование экспертных оценок для получения исходных гипотез и необходимой информации для решения задач синтеза (структурно-параметрической идентификации) формальных моделей принятия решений в рамках теории полезности (методы компараторной идентификации).

Полученные результаты открывают перспективы практического создания СППР широкого назначения.

Список литературы: 1. *Овезгельдыев, А.О.* Синтез и идентификация моделей многофакторного оценивания и оптимизации [Текст] / Э.Г. Петров, К.Э. Петров – Киев: «Наукова думка», 2002. – 164с. 2. *Петров, Е.Г.* Методи та засоби прийняття рішень у соціально-економічних системах [Текст] / Е.Г. Петров, М.В. Новожилова, І.В. Гребенник. – К.: Техніка, 2004. – 256 с. 3. *Глушко, В.М.* Введение в АСУ [Текст] / В.М. Глушко. – Изд. 2-е, исправленное и дополненное. – К.: «Техніка», 1974, – 320 с. 4. *Губаренко, Е.В.* Методология формирования количественных оценок уровня развития региона [Текст] / Е.В. Губаренко, Н.В. Подмогильный // Вестник ХНТУ, 2010. – № 2 (38). – С. 76-80. 5. *Згуровский, М.З.* Роль инженерной науки и практики в устойчивом развитии общества [Текст] / М.З. Згуровский, Г.А. Статюха // Системні дослідження та інформаційні технології. – 2007. – № 1. – С. 19-38. 6. *Мину, М.* Математическое программирование. Теория

и алгоритмы: Пер. с фр. Предисловие А.И. Штерна. — М.: Наука. Гл. ред. физ.-мат.лит., 1990.— 488 с. 7. *Адамар, Ж.* Задача Коши для линейных уравнений с частными производными гиперболического типа [Текст]/Ж Адамар. — М.: Наука, 1978. — 351 с. 8. *Тихонов, А.Н.* Методы решения некорректных задач [Текст] / А.Н. Тихонов, В.Я. Арсенин. — М.: Наука, 1986. — 288 с. 9. *Фишберн, П.* Теория полезности для принятия решений [Текст] / П. Фишберн. — М.: Наука, 1978. — 352 с. 10. *Бриллинджер, Д.Р.* Временные ряды. Обработка данных и теория [Текст] / Д.Р. Бриллинджер. М: Мир, 1980. — 536 с. 11. *Gordon W.J.J.* Sinectics: The Development of Creative Capacity — New York, 1961. — 180p. 12. *Никандров, В.В.* Вербально-коммуникативные методы в психологии [Текст] / В.В. Никандров. СПб.: Речь, 2002. — 72 с. 13. Метод анкетирования, применения для сбора первичных данных по редким видам на примере снежного барса [Текст] / Е.А. Быкова, А.В. Есипов, Б. Аромов, Е.А. Крейцберг, Э.В. Вашетко // Охраняемые территории Центральной Азии. Ташкент, 2004. — С. 208-214. 14. *Ларичев, О.И.* Теория и методы принятия решений, а также Хроника событий в Волшебных странах [Текст] / О.И. Ларичев: Учебник. Изд. второе, перераб. и доп. — М: Логос, 2002. — 392 с: ил. 15. *Девид, Г.* Методы парных сравнений [Текст] / Г. Девид. — М.: Статистика, 1978. — 144 с. 16. *Churchman C.W.* Methods of Inquiry / C.W. Churchman, R.L. Ackoff. — St. Louis: Educational Publishers, 1950. — 134 p. 17. *Thurstone L.L.* The measurement of values / L.L. Thurstone. — Chicago: Univ. Chicago Press, 1959. — 268 p. 18. *Орлов, А.И.* Менеджмент [Текст]: Учебник. / А.И. Орлов. — М.: Издательство «Изумруд», 2003. — 298 с. 19. *Орлов, А.И.* Экспертные оценки [Текст] / А.И. Орлов // Заводская лаборатория. — 1996. Т.62. № 1. С.54-60. 20. *Нейман, Дж.* Теория игр и экономическое поведение [Текст] / Дж. Нейман, О Моргенштерн. — М.: Наука, 1970. — 124 с. 21. *Саати, Т.*, Аналитическое планирование. Ор-

ганизация систем [Текст] / Т. Саати, К. Керне. — М.: Радио и связь, 1991. — 224 с. 22. *Руа, Б.* Классификация и выбор при наличии нескольких критериев (метод ЕЛЕКТРА) [Текст] / Б.Руа // Вопросы анализа и процедуры принятия решений. — М.: Мир, 1976. — С. 80-108.

Поступила в редколлегию 21.05.2010.

УДК 519.8

Методи і інструментальні засоби систем підтримки ухвалення рішень при організаційному управлінні соціально-економічними системами / Е.Г. Петров, Є.В. Губаренко // Біоніка інтелекту: наук.-техн. журнал. — 2010. — № 3 (74). — С. 26–36.

Стаття присвячена процесу ухвалення рішень. Були виділені етапи ухвалення рішень та сформовані вимоги до методів вирішення задач, виникаючих на кожному з етапів. Надається класифікація та огляд методів експертного оцінювання: евристичні, частково формалізовані та компараторні. Наводяться переваги та недоліки кожного класу методів при застосуванні їх у розробці, створенні чи функціонуванні систем підтримки рішень.

Бібліогр.: 22 найм.

UDK 519.8

Methods and tools of the systems of support of making decision at an organizational management the socio-economic systems / E.G. Petrov, E.V. Gubarenko // Bionics of Intelligence: Sci. Mag. — 2010. — № 3 (74). — P. 26–36.

The article is sanctified to the process of making decision. There were the distinguished stages of making decision and formed requirements to the methods of decision of tasks arising up on each of the stages. Classification and review of methods of expert evaluation are given: heuristic, partly formalized and komparator. Advantages and lacks of every class of methods are pointed at application of them in development, creation or functioning of the systems of support of decisions.

Ref.: 22 items.

УДК 519.688:004.896



КОМПЛЕКСИРОВАНИЕ СРЕДСТВ ПОДДЕРЖКИ ПРИНЯТИЯ РЕШЕНИЙ ПРИ СИНТЕЗЕ ТЕРРИТОРИАЛЬНО РАСПРЕДЕЛЕННЫХ ОБЪЕКТОВ

В.В. Бескоровайный¹, О.В. Зыкова²

¹ ХНУРЭ, г. Харьков, Украина, beskorovainyi@kture.kharkov.ua

² ХНУРЭ, г. Харьков, Украина, zykovaolga@online.ua

Предложен состав средств системы поддержки многокритериального оценивания и выбора решений при проектировании и управлении крупномасштабными объектами. Система поддержки принятия решений включает математические модели и методы классической количественной теории полезности, анализа иерархий и определения отношения превосходства по качеству *ELECTRE*. Система с таким набором средств за счет более тщательной экспертизы позволяет определять более эффективные проектные и управленческие решения.

ТЕРРИТОРИАЛЬНО РАСПРЕДЕЛЕННЫЕ ОБЪЕКТЫ, СТРУКТУРНЫЙ СИНТЕЗ, ОПТИМИЗАЦИЯ, МНОГОКРИТЕРИАЛЬНОЕ ОЦЕНИВАНИЕ, ПОДДЕРЖКА ПРИНЯТИЯ РЕШЕНИЙ, ФУНКЦИЯ ОБЩЕЙ ПОЛЕЗНОСТИ, ИДЕНТИФИКАЦИЯ

Введение

Эффективность антропогенных систем во многом определяется способом их структурной организации. Практически любая система может быть реализована на различных наборах элементов, различных схемах и типах связей между элементами. В процессах проектирования, планирования развития или реорганизации систем неизбежно возникает проблема выбора наиболее рационального способа их структурной организации с учетом множества разнородных показателей качества. Такой выбор, как правило, производится с использованием систем поддержки принятия решений (СППР, *DSS* – *Decision Support System*), построенных на основе теории полезности и принятия решений [1]. При этом в СППР все шире используются достижения бионики интеллекта, а состав их средств во многом определяется предметной областью и особенностями решаемых задач.

1. Анализ предметной области и постановка задачи

Для формализованного представления СППР традиционно используется математический аппарат теории множеств, теории категорий, теории графов и теории операторов.

Операторная модель СППР может быть представлена в виде [2]: $Au = z$, где $u \in R_u$, $z \in R_z$ – элементы некоторых метрических пространств; A – оператор модели, переводящий элементы ресурсов $u \in R_u$ в элементы результатов $z \in R_z$.

Архитектура СППР *ADSS* может быть представлена тремя составляющими: объектами O – функциональными и обеспечивающими подсистемами, внешними и внутренними возмущающими воздействиями; связями S , описывающими взаимодействия между объектами; поведением D – изменяющимися состояниями подсистем, зафиксированных в локальных хранилищах данных:

$$ADSS = \langle O, S, D \rangle.$$

Характерной особенностью территориально распределенных объектов (ТРО) является то, что с ростом их масштабов их стоимостные и функциональные характеристики становятся все более зависимыми от топологии (территориальной или пространственной организации). Это требует совместно с традиционными задачами их структурного синтеза решать комплексы задач топологической оптимизации и, таким образом, приводит к проблеме их структурно-функционально-параметрического и топологического синтеза [3].

Проблема является многогранной, включает комплексы задач выбора структуры, топологии, технологии функционирования, параметров элементов и связей, всесторонней многокритериальной оценки и выбора вариантов на различных этапах ее жизненного цикла. При этом проекты создания или реконструкции ТРО требуют привлечения значительных материальных, трудовых, финансовых ресурсов и занимают продолжительное время.

Целью создания ТРО является удовлетворение потребностей множества территориально рассредоточенных пользователей, а степень ее достижения оценивается множеством его важнейших свойств, характеризующих эффекты от его функционирования и затраты, связанные с его созданием и эксплуатацией. Упомянутые свойства при выборе решений используются в качестве частных критериев эффективности ТРО $K = \{k_i\}_{i=1}^n$, где n – количество частных критериев.

В основе современных подходов к выбору решений лежит парадигма максимизации полезности [4]. Считается, что лицо, принимающее решение (ЛПР), при выборе вариантов из множества допустимых $x \in X$ приписывает им некоторую полезность $P(x)$, значения которой и определяют его

выбор: $\forall x, y \in X: x \sim y \leftrightarrow P(x) = P(y); x \succ y \leftrightarrow P(x) > P(y); x \geq y \leftrightarrow P(x) \geq P(y)$. При этом одной из важнейших задач является определение количественной или порядковой метрики.

При выборе комплекса средств многокритериального оценивания и выбора проектных решений учитывают следующие особенности проблемы синтеза ТРО [5].

1. Тесная взаимосвязь и неполная информационная определенность задач $\{Task_l\}_{l=1}^t$ (где t – количество задач проблемы) выбора принципов построения ТРО, структурного, топологического, параметрического и технологического синтеза, а также анализа и выбора проектных решений по наборам исходных данных $InDat_l$ и ограничений Res_l , обуславливает итерационный характер методов и процедур их решения. Таким способом обеспечивается разрешимость комплекса задач $\{Task_l\}_{l=1}^t$ по входам.

2. Высокая сложность методов решения (решающих процедур), обусловленная комбинаторным характером большинства задач $\{Task_l\}_{l=1}^t$, и широкий диапазон условий их решения требуют при решении каждой из них использования множества методов $\{MetDec_l^k\}_{k=1}^{l_k}$ (где l_k – количество методов для решения l -й задачи), имеющих различную сложность и точность решения. Это обеспечивает разрешимость задач системного проектирования по ресурсам.

3. Для более полного использования опыта проектировщиков и учета трудно формализуемых факторов процесс решения целесообразно строить на основе интерактивных (человеко-машинных) процедур. Процесс поиска проектного решения при этом будет состоять из взаимодополняющих процедур автоматического и интеллектуального синтеза с участием системных аналитиков и операторов, применением CASE-средств и экспертных систем.

4. На всех этапах проектирования целесообразно использовать приемы, снижающие трудоемкость решения задач системного проектирования $\{Task_l\}_{l=1}^t$. С этой целью могут быть использованы различного рода эвристики, учитывающие специфику задач, решения, полученные с помощью “быстрых” процедур, формальные или экспертные оценки.

Особенности ТРО выдвигают противоречивые требования к средствам многокритериального оценивания и выбора проектных решений. Одна часть средств СППР должна работать в составе автоматических процедур проектирования (следовательно, быть достаточно «быстрыми»), другая часть, определяющая глобальные проектные решения, используется для принятия решений в условиях неопределенности и риска, наличия трудноформализуемых факторов (следовательно, требует привлечения аналитиков и экспертов).

В первом случае необходима количественная оценка проектных решений, что требует использования кардиналистических (количественных) методов теории полезности, во втором – достаточным является упорядочение решений, что может быть осуществлено и помощью ординалистических (порядковых) методов. В настоящее время для решения этой задачи используются три основных подхода: классической количественной многокритериальной теории полезности; аналитической иерархии; определения отношения превосходства по качеству [6].

Целью данной работы является анализ существующих средств многофакторного оценивания и выбора решений, а также их комплексирование в СППР для решения задач оптимизации крупномасштабных объектов.

2. Модели и методы классической количественной многокритериальной теории полезности

Этот подход развивается в рамках многокритериальной теории полезности *MAUT (Multi-Attribute Utility Theory)* [6]. В рамках этого подхода количественная оценка предпочтений вариантов выражается с помощью функций общей полезности (ФОП) $P(x)$, $x \in X$, которая строится на основе функций полезности частных критериев (ФПЧК) $\xi_i(k_i(x))$, $i = \overline{1, n}$ (где n – количество частных критериев) [4].

Современные модели многокритериального оценивания и выбора строятся на основе аддитивных, мультипликативных или смешанных ФОП [4, 7–9]:

$$P(x) = \sum_{i=1}^n \lambda_i \xi_i(x); \quad (1)$$

$$P(x) = \prod_{i=1}^n [\xi_i(x)]^{\lambda_i}; \quad (2)$$

$$P(x) = \sum_{i=1}^n \lambda_i \cdot \bar{k}_i(x) + \sum_{i=1}^n \sum_{j=i}^n \lambda_{ij} \cdot \bar{k}_i(x) \cdot \bar{k}_j(x) + \dots; \quad (3)$$

$$P(x) = [\beta \cdot \sum_{i=1}^n \lambda_i \cdot \bar{k}_i(x)] + \{(1 - \beta) \cdot \prod_{i=1}^n [\bar{k}_i(x)]^{\lambda_i}\}, \quad (4)$$

где $P(x)$ – полезность варианта x ; n – количество частных критериев; λ_i – коэффициент важности критерия k_i , выбираемый с учетом условий $\sum_{i=1}^n \lambda_i = 1$, $\lambda_i \geq 0$; $i = \overline{1, n}$, $\xi_i(x) = \xi_i(k_i(x))$ – ФПЧК критерия k_i ; λ_{ij} – весовые коэффициенты произведения нормированных критериев $\bar{k}_i \cdot \bar{k}_j$, $i, j = \overline{1, n}$; β – адаптационный параметр, определяющий вид схемы, $0 \leq \beta \leq 1$. При $\beta = 1$ (4) реализует аддитивную схему (1), при $\beta = 0$ – мультипликативную схему вида (2).

Недостатком аддитивных моделей вида (1) считается то, что они не отражают объективно роль частных критериев и допускают практически неограниченную компенсацию одних свойств решения другими, а мультипликативных моделей вида

(2) – то, что они компенсируют недостаточную величину одного критерия избыточной величиной другого [8].

Комбинированные модели (3)–(4) являются развитием моделей (1)–(2) и объединяют их достоинства. Недостатком ФОРП вида (3) считается ее сложность, а (4) – трудность определения значения адаптационного параметра β .

В рамках этого подхода определение метрики для ранжирования альтернатив в виде функции общей полезности (ФОРП) $P(x)$ представляет, по сути, решение задачи идентификации. В общем случае в процессе идентификации требуется решение вопросов, связанных с выбором критериев подобия, входных сигналов, структуры и параметров модели, оценки ее точности и адекватности. Наибольший интерес как в теоретическом, так и в практическом плане представляют вопросы выбора структуры и параметров (весовых коэффициентов и параметров ФПЧК, адаптационных параметров) моделей оценивания $P(x)$. Традиционно такие задачи решаются методами экспертного оценивания (приписывания, ранжирования, парных сравнений, множественных сравнений), что требует больших затрат времени на проведение экспертизы.

Особенностью процессов идентификации моделей многофакторного оценивания и выбора решений является то, что в качестве выходных сигналов в таких задачах выступают оценки и решения человека, носящие приближенный или качественный характер. Это делает плодотворным использование для решения таких задач метода компараторной идентификации [4].

Суть общей задачи компараторной идентификации модели многофакторного оценивания состоит в следующем. ЛПР на основе оценок по множеству частных критериев $\{k_i(x)\}_{i=1}^n$ устанавливает качественную полезность на множестве альтернатив $x \in X$. Она может быть выражена множеством бинарных отношений эквивалентности $R(X) = \{(x, y) : x, y \in X, x \sim y\}$, нестрогого предпочтения $R_{NS}(X) = \{(x, y) : x, y \in X, x \geq y\}$, строгого предпочтения $R_S(X) = \{(x, y) : x, y \in X, x \succ y\}$ и представлена порядком одного из видов:

$$R^o(X) = x^o \succ x_i \succ x_j \succ \dots \succ x_m, \quad (5)$$

$$R^o(X) = x^o \succ x_i \sim x_j \succ \dots \succ x_m, \quad (6)$$

$$R^o(X) = x^o \geq x_i \geq x_j \geq \dots \geq x_m. \quad (7)$$

Требуется для установленного ЛПР порядка $R^o(X)$ (5), (6) или (7) выбрать вид модели многофакторного оценивания $P_g(x)$ из множества допустимых (1)–(4) и подобрать наилучшие значения ее параметров $q \in Q$.

В качестве критериев идентификации в зависимости от условий задачи используются: минимум суммарной (средней, максимальной, суммарной

квадратичной) абсолютной, относительной погрешности оценки общей полезности $P(x)$, максимум силы предпочтений, средней точки, максимизации функции правильности выбора или погрешности восстановления порядка альтернатив [9].

На основе метода компараторной идентификации к настоящему времени разработаны: для ФОРП вида (1) математические модели и методы точечного оценивания параметров [10] и методы идентификации интервальных групповых предпочтений [11]; для ФОРП вида (2) метод, позволяющий свести задачи параметрической идентификации к задачам линейного или нелинейного математического программирования [12]; для множества ФОРП (1)–(4) методы решения общей задачи идентификации, включающей не только выбор вида модели, но и определение весовых коэффициентов, а также параметров функций полезности частных критериев [13].

3. Метод анализа иерархии

Метод анализа иерархии МАИ (*AHP – Analytic Hierarchy Process*) используется для решения слабо структурированных и неструктурированных проблем, т.е. в тех случаях, когда эксперты (или ЛПР) не могут дать абсолютные оценки альтернатив по критериям, а пользуются более слабыми сравнительными измерениями [14].

В общем случае в процессе реализации МАИ выделяют четыре основных этапа [6]:

1. Структуризация задачи в виде иерархической структуры с несколькими уровнями: цели – критерии – альтернативы.
2. Парные сравнения элементов каждого уровня.
3. Вычисление коэффициентов важности для элементов каждого уровня.
4. Расчет количественных оценок качества каждой из альтернатив и определение наилучшей среди них.

Суть МАИ применительно к решению задач многокритериального синтеза ТРО состоит в декомпозиции исходной проблемы с выделением цели их создания (удовлетворение потребностей множества территориально распределенных пользователей), частных критериев эффективности (оперативность реализации функций, надежность, живучесть, затраты на создание и эксплуатацию) и собственно вариантов построения (структура, параметры, топология, технология функционирования).

Далее производятся парные сравнения компонентов решаемой проблемы с использованием шкалы словесных определений уровня важности (равная важность, умеренное превосходство, сильное превосходство, значительное превосходство, очень большое превосходство) с количественной оценкой в баллах, например, от 1 до 9. Результаты записываются в виде матриц парных сравнений

$$A = \begin{bmatrix} a_{11} & a_{12} & \dots & a_{1m} \\ a_{21} & a_{22} & \dots & a_{2m} \\ \vdots & \vdots & \ddots & \vdots \\ a_{m1} & a_{m2} & \dots & a_{mm} \end{bmatrix}. \quad (8)$$

На главных диагоналях ($a_{11}, a_{22}, a_{33}, \dots, a_{mm}$) матриц (8) окажутся сравнительные характеристики по одному и тому же элементу, т.е. $a_{ii} = 1, \forall i = \overline{1, m}$. Такие матрицы имеют свойство обратной симметрии

$$a_{ij} = \frac{1}{a_{ji}}, \quad i, j = \overline{1, m}. \quad (9)$$

В результате выполнения $m \cdot (m - 1) / 2$ попарных сравнений элементов проблемы одного уровня иерархии на основе экспертных оценок формируется часть элементов матриц, расположенных, выше $A = [a_{ij}], i = \overline{1, m}, j = \overline{i, m}$ (или ниже $A = [a_{ij}], i = \overline{1, m}, j = \overline{1, i}$) главных диагоналей. Элементы, расположенные ниже $A = [a_{ij}], i = \overline{1, m}, j = \overline{1, i}$ (выше $A = [a_{ij}], i = \overline{1, m}, j = \overline{i, m}$) главных диагоналей, вычисляются с использованием свойства обратной симметрии (9).

Для полученных матриц парных сравнений (8) определяются их собственные векторы

$$\alpha_i = m \sqrt[m]{\sum_{j=1}^m a_{ij}}. \quad (10)$$

По полученным значениям собственных векторов $\alpha_i, i = \overline{1, m}$ (10) путем нормирования их координат определяют коэффициенты важности элементов соответствующего иерархического уровня $\beta_i = \alpha_i / \sum_{j=1}^m \alpha_j, i = \overline{1, m}$.

Для повышения достоверности оценок при заполнении каждой из матриц (путем сравнения со случайно заполненной матрицей) осуществляется проверка согласованности суждений ЛПР.

Оценка показателей важности вариантов построения ТРО осуществляется на основе аддитивного соотношения

$$P_j = \sum_{i=1}^n w_i v_{ji}, \quad (11)$$

где P_j – показатель качества j -го варианта построения ТРО; w_i – вес i -го частного критерия; v_{ji} – важность j -го варианта построения ТРО по i -му критерию.

К числу недостатков МАИ относят [6]: введение нового, недоминирующего варианта в общем случае может привести к изменению предпочтений между двумя ранее заданными вариантами; недостаточно обоснованный переход к числам при проведении измерений, а также оторванность метода объединения оценок от предпочтений ЛПР.

Дальнейшее развитие этот подход получил в виде методов мультипликативной аналитической иерархии и метода *МАСВЕТН*.

4. Методы качественного ранжирования многокритериальных альтернатив

В рамках этого подхода разработан ряд методов исключения и выбора, объединенных в семейство *ELECTRE (ELimination Et Choix Traduisant la Realite)* [4, 6]. В отличие от методов *MAUT* и *AHP* в методах этого семейства оценка каждой альтернативы является не количественной, а относительной (по сравнению с другой). Методы *ELECTRE* направлены на решение задач с уже заданными многокритериальными альтернативами.

Связь между любой парой альтернатив $x, y \in X$ определяется последовательностью бинарных отношений $xR_1y, xR_2y, \dots, xR_ny$. Для каждой пары $x, y \in X$ можно установить предпочтительность, равноценность или несравнимость. Бинарное отношение превосходства задается уровнем индексов согласия и несогласия.

На основе выбранного бинарного отношения осуществляется попарное сравнение всех альтернатив $x, y \in X$. Варианты, которые оказались лучшими при всех сравнениях $x \succ y, x, y \in X$, выделяются в новое множество, называемое ядром. После выделения ядра его элементы объявляются несравнимыми. За первым бинарным отношением xR_1y следует второе xR_2y , далее – третье xR_3y и т.д. Каждое последующее отношение является более слабым. Ядра, соответствующие последующим отношениям, содержат в общем случае меньшее количество несравнимых элементов. Процесс получения ядер заканчивается, когда достигается требуемое количество элементов в ядре. Эти элементы вместе с последним бинарным отношением являются решением задачи.

Бинарные отношения между альтернативами $x, y \in X$ строятся следующим образом. Каждому из n критериев, имеющих числовые шкалы, ставится в соответствие число $p_i, i = \overline{1, n}$, характеризующее важность критерия $k_i, i = \overline{1, n}$. В частности, можно рассматривать p_i как «количество голосов» членов жюри, отдавших предпочтение критерию $k_i, i = \overline{1, n}$. Выдвигается гипотеза о превосходстве альтернативы x над y , т.е. $x \succ y$. Множество $K = \{k_i\}$, состоящее из n критериев, разбивается на три подмножества:

- $K^+(x, y)$ – подмножество критериев, по которым x предпочтительнее y , т.е. $x \succ y$;
- $K^=(x, y)$ – подмножество критериев, по которым x равноценно y , т.е. $x \sim y$;
- $K^-(x, y)$ – подмножество критериев, по которым y предпочтительнее x , т.е. $x \prec y$.

Далее формулируется индекс согласия с гипотезой о превосходстве x над y . В частности, этот индекс может быть определен как отношение суммы весов критериев подмножеств K^+ и $K^=$ к общей сумме весов всех критериев:

$$c_{xy} = \frac{\sum_{i \in (K^+ \cup K^-)} p_i}{\sum_{i=1}^n p_i}. \quad (12)$$

Индекс несогласия d_{xy} с гипотезой о превосходстве x над y определяется на основе самого «противоречивого» критерия – критерия, по которому y в наибольшей степени превосходит x . Для учета наибольшей разницы длин шкал критериев разность оценок y и x относят к длине наибольшей шкалы

$$d_{xy} = \max_{i \in K^-} \{ |k_i(x) - k_i(y)| / L_i \}, \quad (13)$$

где $k_i(x)$, $k_i(y)$ – оценки альтернатив x и y по i -му критерию; L_i – длина шкалы i -го критерия.

Очевидно, что $0 \leq c_{xy} \leq 1$, $0 \leq d_{xy} \leq 1$.

Бинарное отношение превосходства задается уровнем индексов согласия и несогласия. Если $c_{xy} \geq c_1$ и $d_{xy} \leq d_1$ (где c_1 и d_1 – заданные уровни), то альтернатива x объявляется превосходящей альтернативу y . Если же при этих уровнях сравнить альтернативы не удалось, то они объявляются несравнимыми. Уровни c_1 и d_1 позволяют выделить ядро, в которое входят доминирующие и несравнимые элементы.

Выводы

На основе анализа особенностей ТРО как объектов проектирования и управления предложен состав средств системы поддержки многокритериального оценивания и выбора проектных и управленческих решений, включающей модели и методы классической количественной теории полезности, анализа иерархии и определения отношения превосходства по качеству. СППР с таким набором средств позволит определять наилучшие проектные (управленческие) решения как в автоматическом, так и в интерактивном режимах функционирования.

Предложенные средства реализованы программно, апробированы и показали свою работоспособность на контрольных примерах. Относительная сложность предложенной СППР компенсируется большей экономией средств, используемых на создание и эксплуатацию крупномасштабных объектов, за счет более тщательной экспертизы решений.

Список литературы: 1. Катуплев, А.Н. Математические методы в системах поддержки принятия решений [Текст] / А.Н. Катуплев, Н.А. Северцев. – М.: Высшая школа, 2005. – 311 с. 2. Авраменко, В.П. Система поддержки принятия решений по ликвидации нештатных ситуаций [Текст] / В.П. Авраменко // Радиоэлектроника. Информатика. Управление. – 2009. – №2. – С. 107–111. 3. Петров, Э.Г. Территориально распределенные системы обслуживания [Текст] / Э.Г. Петров, В.П. Пискалова, В.В. Бескоровайный. – Киев: Техника, 1992. – 208 с. 4. Овезгельдыев, А.О. Синтез и идентификация моделей многофакторного оценивания и оптимизации [Текст] / А.О. Овезгельдыев,

Э.Г. Петров, К.Э. Петров. – К.: Наук. думка, 2002. – 164 с. 5. Бескоровайный, В.В. Формирование и выбор решений задачи системного проектирования территориально распределенных систем обработки информации [Текст] / В.В. Бескоровайный // Системы обработки информации. – 2002. – Вып. 6 (22). – С. 243–247. 6. Ларичев, О.И. Теория и методы принятия решений, а также Хроника событий в Волшебных Странах [Текст] / О.И. Ларичев. – М.: Логос, 2000. – 296 с. 7. Брахман, Т.Р. Многокритериальность и выбор альтернативы в технике [Текст] / Т.Р.Брахман. – М.: Радио и связь, 1984. – 287 с. 8. Анфилатов, В.С. Системный анализ в управлении [Текст] / В.С. Анфилатов, А.А. Емельянов, А.А. Кукушкин. – М.: Финансы и статистика, 2003. – 368 с. 9. Бескоровайный, В.В. Структурно-параметрична ідентифікація моделей багатofакторного оцінювання [Текст] / В.В.Бескоровайный, И.В.Трофименко // Системи озброєння і військова техніка. – 2006. – № 3 (7). – С. 56–59. 10. Петров, Э.Г. Методика оценки адекватности моделей точечной идентификации индивидуальных предпочтений ЛПР [Текст] / Э.Г. Петров, Н.С. Шило // Радиоэлектроника и информатика. – 2003. – №2. – С. 97–103. 11. Петров, Э.Г. Модель выбора многокритериального решения при интервальном задании весовых коэффициентов [Текст] / Э.Г.Петров, Л.В.Батий // Вестник Херсонского государственного технического университета. – 2002. – № 1 (14). – С. 28–31. 12. Бескоровайный, В.В. Параметрическая идентификация мультипликативных моделей для многофакторного выбора решений [Текст] / В.В.Бескоровайный, И.В.Трофименко // Збірник наукових праць Харківського університету повітряних сил. – Харків: ХУ ПС, 2005. – Вып. 5 (5). – С. 74–78. 13. Бескоровайный, В. В. Метод решения задачи компараторной идентификации моделей многофакторного оценивания [Текст] / В. В. Бескоровайный, Э. Г. Петров, И. В. Трофименко // Бионика интеллекта. – 2006. № 2 (65). – С. 3–7. 14. Саати, Т. Принятие решений. Метод анализа иерархий [Текст]: пер. с англ. / Т. Саати. – М.: Радио и связь, 1993. – 278 с.

Поступила в редколлегию 24.05.2010.

УДК 519.688: 004.896

Комплексування засобів підтримки прийняття рішень при синтезі територіально розподілених об'єктів / В.В. Бескоровайний, О.В. Зикова // Біоніка інтелекту: наук.-техн. журнал. – 2010. – № 3 (74). – С. 37–41.

Запропоновано склад засобів системи підтримки багатокритеріального оцінювання й вибору рішень при проектуванні й керуванні великомасштабними об'єктами. Система підтримки прийняття рішень включає математичні моделі й методи класичної кількісної теорії корисності, аналізу ієрархії й визначення відносин переваги за якістю *ELECTRE*.

Бібліогр.: 14 найм.

UDK 519.688: 004.896

Integration of decision-support tools in the synthesis of geographically distributed objects / V.V.Beskorovainyj, O.V.Zykova // Bionics of Intelligence: Sci. Mag. – 2010. – № 3 (74). – P. 37–41.

The composition of the decision support system for multi-criteria estimation and choice of solutions used for designing and managing large-scale objects is proposed. The decision support system includes mathematical models and methods of classical multi-attribute utility theory, hierarchy analysis and definitions of relationship superiority by quality *ELECTRE*.

Ref.: 14 items.

УДК 519.876.2



МЕТОД ИССЛЕДОВАНИЯ ПРОЦЕССОВ ПРИНЯТИЯ РЕШЕНИЙ В ГОСРЕГУЛИРОВАНИИ ТЕЛЕКОММУНИКАЦИОННОЙ ОТРАСЛИ

А.В. Калмыков¹¹ Национальный аэрокосмический университет им. Н.Е. Жуковского «ХАИ», г. Харьков, Украина

Рассматриваются процессы принятия решений в сложных организационных структурах. Используется методика формирования системных моделей организационных систем на основе декомпозиции на уровни иерархии и старты (аспекты) деятельности. Предложено представление аспектов деятельности в виде вложенной иерархии решающих и исполнительных компонентов. Показана многоуровневая структура информационных связей между управляющей системой и управляемым объектом.

СИСТЕМНЫЙ ПОДХОД, ИЕРАРХИЧЕСКАЯ ОРГАНИЗАЦИОННАЯ СИСТЕМА, ДЕКОМПОЗИЦИЯ, ЦЕЛЕВАЯ МОДЕЛЬ, ФУНКЦИОНАЛЬНАЯ МОДЕЛЬ, РЕШАЮЩИЙ ЭЛЕМЕНТ

Введение

Современное развитие телекоммуникационной отрасли определяется многими внутренними и внешними факторами, среди которых важнейшее место занимает государственное регулирование. Данные процессы включают такие направления:

- разрешение на вид деятельности (лицензирование);
- распределение ресурсов (номерных, радиочастотных);
- определение перечня и потребительской стоимости социально-значимых услуг;
- допуск к использованию новых видов технологий и оборудования;
- формирование принципов экономико-технологического взаимодействия предприятий-операторов телекоммуникаций и осуществление контроля их соблюдения;
- определение производственно-технологических требований (лицензионные условия) к предприятиям-операторам телекоммуникаций и контроль их выполнения.

Следовательно, процессы государственного регулирования отрасли телекоммуникаций во многом определяются факторами внешнего окружения, прежде всего, социальными, экономическими и технологическими, а эффективность деятельности и выполнения предприятиями задач и функций отрасли в значительной степени зависит от «качества» регулирования. Проблемы реальной управленческой деятельности в телекоммуникационной отрасли отражены в документе «Концепція розвитку телекомунікацій в Україні» [1], где отмечено следующее:

- «...недостатній регуляторний вплив держави на ринок телекомунікацій;
- недостатне фінансове та матеріально-технічне забезпечення розроблення наукового підходу до визначення принципів державної політики щодо регуляторного впливу на ринок телекомунікацій;
- низький рівень координації органами виконавчої влади діяльності з розвитку мереж загального користування;

– недосконалість нормативної бази з питань провадження інвестиційної діяльності;...»

Часть из указанных недостатков обусловлена объективными ресурсными ограничениями, однако, другие вызваны проблемами внутренней организации аппарата управления отрасли и взаимодействия с внешним окружением. Такая ситуация приводит к неэффективному управлению ресурсами, например, к запоздалой приватизации основного государственного актива в отрасли телекоммуникаций – ОАО Укртелеком, отсутствию четкой программы развития беспроводных технологий.

Очевидно, что в данном контексте целесообразно рассмотреть процессы деятельности аппарата управления отрасли с учетом внешнего и внутреннего информационного обмена и схем принятия управленческих решений.

1. Постановка задачи исследования

Таким образом, можно сформулировать объективную научно-прикладную задачу исследования реальных процессов управления отраслью. Перспективной целью таких исследований является прогнозирование результатов различных внешних и внутренних воздействий на аппарат управления отраслью, определение критических точек при принятии управленческих решений.

Государственная система управления отраслью является сложной многоуровневой структурой, деятельность которой регламентируется законами Украины («Про телекомуніції», «Про радіочастотний ресурс України»), множеством внутренних положений, инструкций. Исследование таких больших объектов предполагает использование методов анализа больших технических систем [2], которые предполагают представление сложной системы в виде многоуровневой иерархической структуры. Подобные методы анализа применяются и для анализа сложных систем управления [3] и, следовательно, могут быть использованы и для исследования больших организационных

структур, которые, по сути, являются смешанными человеко-машинными системами управления. При этом следует отметить, что большинство известных методов исследования организационных систем основано на использовании структурно-функционального подхода и принципов бизнес-моделирования, которые предполагают анализ экономической и операционной эффективности управленческих структур [4], [8], [9]. Вместе с тем, для телекоммуникационной отрасли в силу ряда политических, социальных факторов интересным представляется изучение процессов принятия управленческих решений с тем, чтобы определить возможные условия, причины и механизмы возникновения неэффективных решений.

В работе [3] рассматривается прикладная задача системного проектирования производственно-техно-логических систем, в ходе решения которой последовательно изучаются различные аспекты деятельности исследуемой системы. Для этого определяется иерархия аспектов деятельности или страт системы: целевая, функциональная структурно-организационная, инфологическая, ресурсная, алгоритмическая и формируются соответствующие системные модели. Предлагаемая методика проектирования сложных систем предполагает исследование объекта в первую очередь на основе целевой, затем функциональной системных моделей. Последующие шаги состоят в формировании на основе полученных результатов системных моделей оптимальных структур системы, её информационного и алгоритмического обеспечения.

Данный подход с некоторыми допущениями может быть использован и для анализа процессов принятия решений в существующих организационных системах, при этом отправной точкой будет не целевой аспект, а организационная структура. Таким образом, с учетом сформулированной в данной статье задачи, далее рассматриваются алгоритмический и информационный аспекты деятельности на основе системной модели существующей структуры управления отраслью.

2. Системные модели структуры и функций аппарата управления

Отрасль телекоммуникаций можно представить как метасистему, состоящую из множества производственных систем, выполняющих схожие функции и имеющих подобные цели, разнящиеся масштабом деятельности и способами решения задач, и одной регулирующей (управляющей) системы. Такое представление отрасли показано на рис. 1.

Введём такие обозначения [5]:

- представление отрасли как совокупности систем – *MetSys*;
- система управления и регулирования отрасли – *СУР MetSys*;

- предприятия – участники отрасли – Sys_i ;
- СУР MetSys* имеет сложную структуру, которая состоит из:

- центрального органа, регулирующего деятельность в отрасли на концептуальном и стратегическом уровне;
- нескольких исполнительных институтов, осуществляющих, в основном, функции контроля и регулирования деятельности предприятий отрасли на оперативном и производственном уровнях.

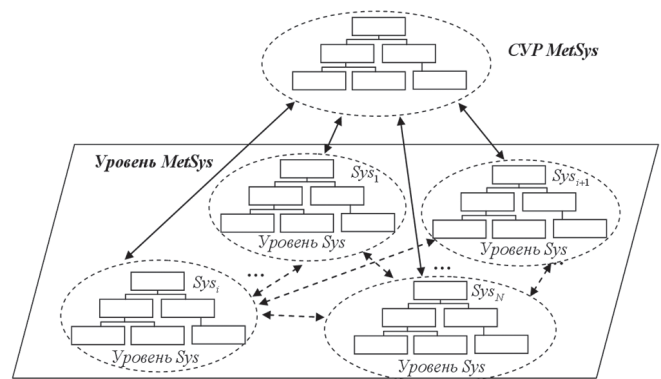


Рис. 1. Представление отрасли на уровне *MetSys* → *Sys*

Исходя из предложенной в [5] методики проектирования сложных систем, для *СУР* телекоммуникационной отрасли определим такие уровни организационной структуры:

- *СУР Metsys* – уровень метасистемы (аппарат управления);
- *СУР Sys* – уровень системы (обособленное подразделение, институт, инспекция и т.п.);
- *СУР UnSys* – уровень подсистемы (департамент или управление);
- *СУР GrEl* – уровень группы элементов (отдел, рабочая группа);
- *СУР El* – уровень элементов (непосредственный исполнитель).

В соответствии с современными фактическими данными системная структурно-организационная модель *СУР* телекоммуникационной отрасли будет выглядеть как многоуровневая иерархия. На рис. 2 показан фрагмент такой системной модели.

Полученная структурная модель является базисом для формирования функциональной модели фактического состояния организационной системы «как есть». Каждый элемент организационной структуры имеет определённую функциональную нагрузку, при этом функции элементов более высокого уровня состоят из нескольких (или большего количества) функций низкого уровня. Следовательно, принципы декомпозиции функциональной страты будут такими же, как и для структурной модели. При этом системные модели целей и функций могут быть нетождественными. На рис. 3. показан фрагмент функциональной системной модели *СУР* отрасли.

Выполнение функций связано с выполнением набора действий или алгоритмов и обменом информацией между элементами организационной структуры. Таким образом, для изучения принципов функционирования и принятия решений в системе управления отраслью целесообразно рассмотреть системные алгоритмическую и информационную модели СУР отрасли.

3. Системные алгоритмическая и информационная модели аппарата управления

Системная алгоритмическая модель реальной организационной системы управления телекоммуникационной отрасли формируется в базе функциональной системной модели, исходя из допущения: каждой функции соответствует алгоритм ее выполнения. При этом алгоритм высокого уровня включает в себя несколько алгоритмов более низкого уровня. Фрагмент алгоритмической системной модели СУР телекоммуникационной отрасли показан на рис. 4. Стрелками обозначены внутриуровневые связи, которые показывают логическую последовательность выполнения алгоритмов элементами структуры.

Принятие решений или выполнение соответствующих алгоритмов осуществляется на основании входной информации, поступающей по междууровневым и внутриуровневым связям от элементов СУР и управляемого процесса.

Информационная модель СУР описывает такие потоки информации между элементами структуры, а также между СУР и управляемым процессом. Каждый j -й элемент h -ого уровня организационной структуры взаимодействует:

- с вышестоящим по иерархии i -ым элементом: γ_{ji}^{h-1} – входящий координирующий информационный сигнал, ϖ_j^h – исходящий информационный сигнал обратной связи;
- с подчиненными по иерархии n -ми элементами: γ_{jn}^h – исходящий координирующий информационный сигнал, ϖ_n^{h+1} – входящий информационный сигнал обратной связи;
- с управляемым процессом или объектом (для нижних уровней): m_j – исходящий управляющий информационный сигнал, z_j – входящий информационный сигнал обратной связи;
- с элементами своего уровня иерархии в соответствии со структурой внутриуровневых алгорит-

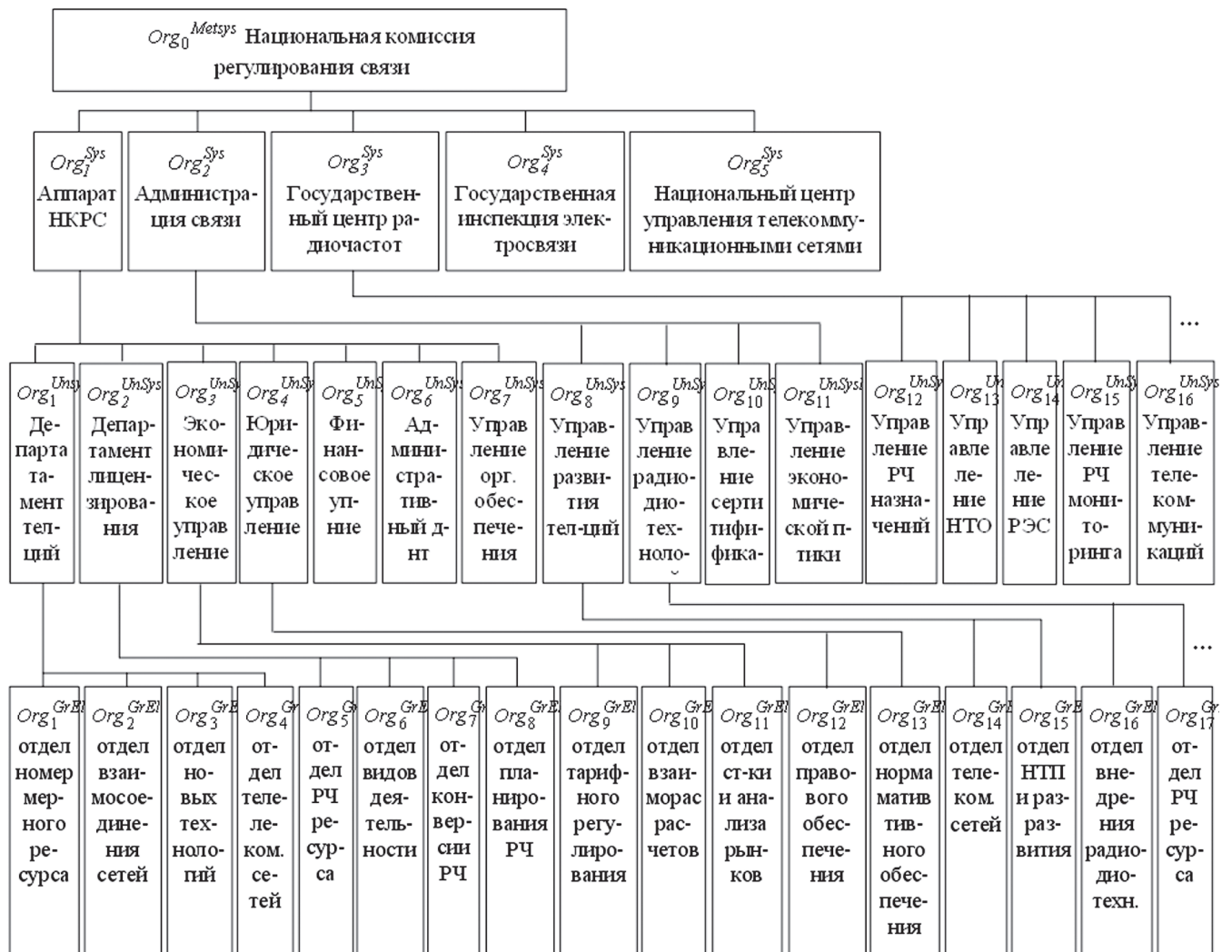


Рис. 2. Системная структурно-организационная модель СУР телекоммуникационной отрасли

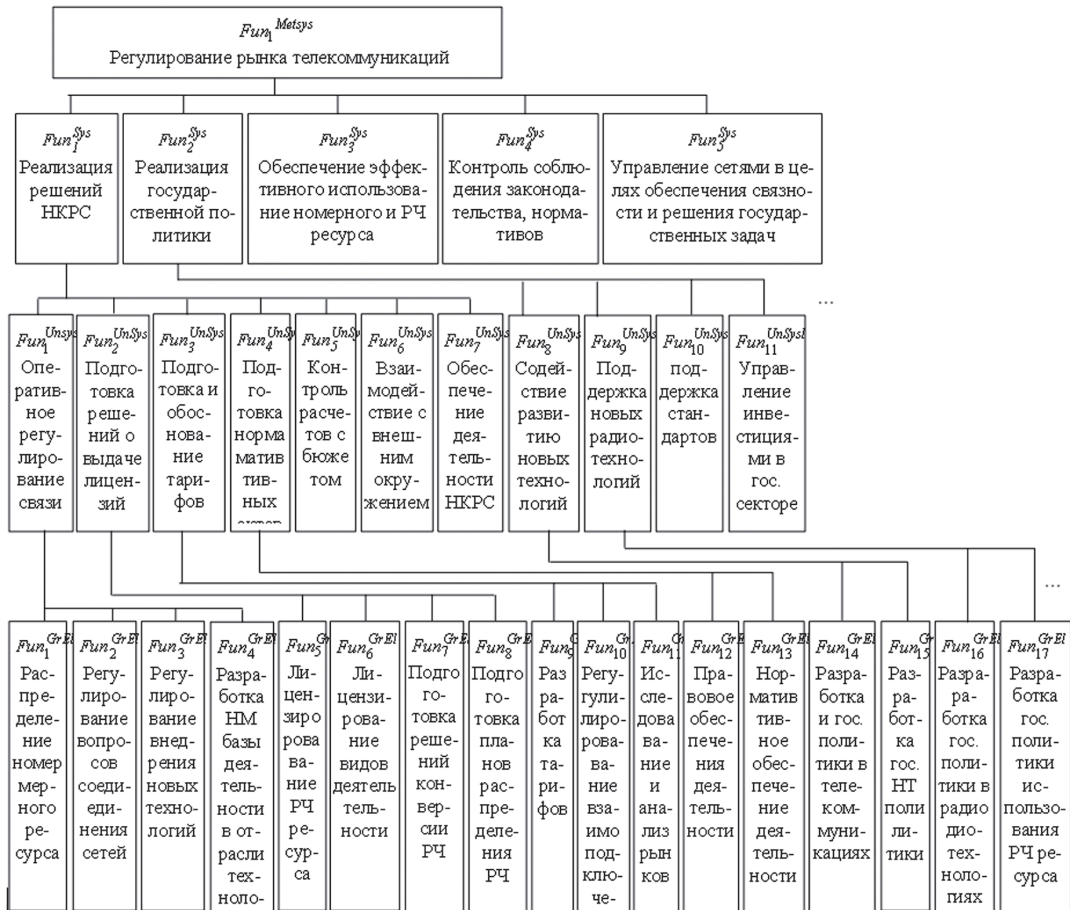


Рис. 3. Системная функциональная модель СУР телекоммуникационной отрасли

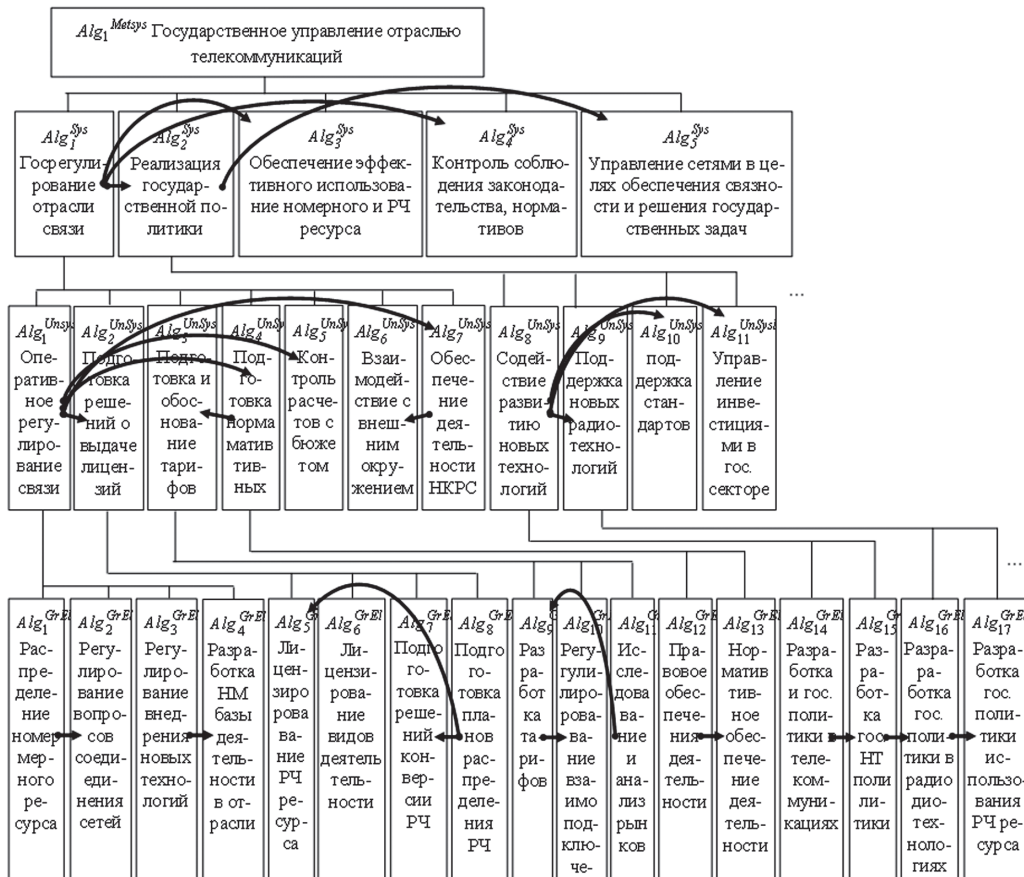


Рис. 4. Фрагмент системной алгоритмической модели системы регулирования отрасли телекоммуникаций

мических связей: out_{ji}^h – исходящий сигнал к i -ому элементу, in_{ij}^h – входящий сигнал от i -ого элемента, причём $out_{ji}^h = in_{ij}^h$.

На рис. 5 показана системная информационная модель фрагмента многоуровневой СУ, где Inf_i^h – процесс-преобразователь информации, соответствующий i -ому элементу h -ого уровня организационной иерархии.

Соответственно, управляемый процесс P взаимодействует:

- с элементами нижнего уровня (s -ми) СУ: m_s – входящий управляющий информационный сигнал, m_s – исходящий информационный сигнал обратной связи с СУ;

- с внешним окружением: ω – входящее внешнее возмущение, y – выход управляемого процесса P .

Для СУР отраслью перечисленные выше информационные сигналы имеют практические смысл и содержание (см. табл. 1). Формирование выходных информационных сигналов осуществляется в соответствии с алгоритмами функционирования элементов СУР отрасли. В формализованном виде такие алгоритмы представлены в виде положений, должностных, профессиональных инструкций, правил. Информационное обеспечение каждого элемента Org_i организационной структуры СУР описывается выражениями (1-3):

$$\bigcup_j \gamma_{0j}^{MetSys} = Inf_0^{MetSys} \left(\bigcup_j \omega_j^{Sys} \right), \quad (1)$$

Таблица 1
Содержание информационных сигналов между элементами организационной системы и управляемым процессом. Часть 1

γ – координирующий информационный сигнал между элементами организационной структуры	Резолюции, приказы, указания, утверждения, согласования
ω – входящий информационный сигнал обратной связи между элементами организационной структуры	Отчеты, акты, проекты приказов, решений, положений
m – управляющий информационный сигнал к управляемому процессу	Решения о предоставлении лицензии, номерного и РЧ ресурса, предписания, рекомендации, решения о наложении штрафов, отзыве лицензии, номерного и РЧ ресурса
z – информационный сигнал обратной связи от управляемого процесса	Отчеты, письма, акты, запросы
inp, out – информационные сигналы между элементами одного уровня в процессе выполнения задач	Письма (запросы, ответы)

$$\left(\bigcup_j \gamma_{ij}^h, \bigcup_k out_{ik}^h, \omega_i^h \right) = Inf_i^h \left(y_{xi}^{h-1}, \bigcup_j \omega_j^{h+1}, \bigcup_k inp_{ki}^h \right), \quad (2)$$

$$\left(m_s, \omega_s^g, \bigcup_k out_{sk}^g \right) = Inf_s^g \left(y_{xs}^g, z_s, \bigcup_k inp_{ks}^g \right), \quad (3)$$

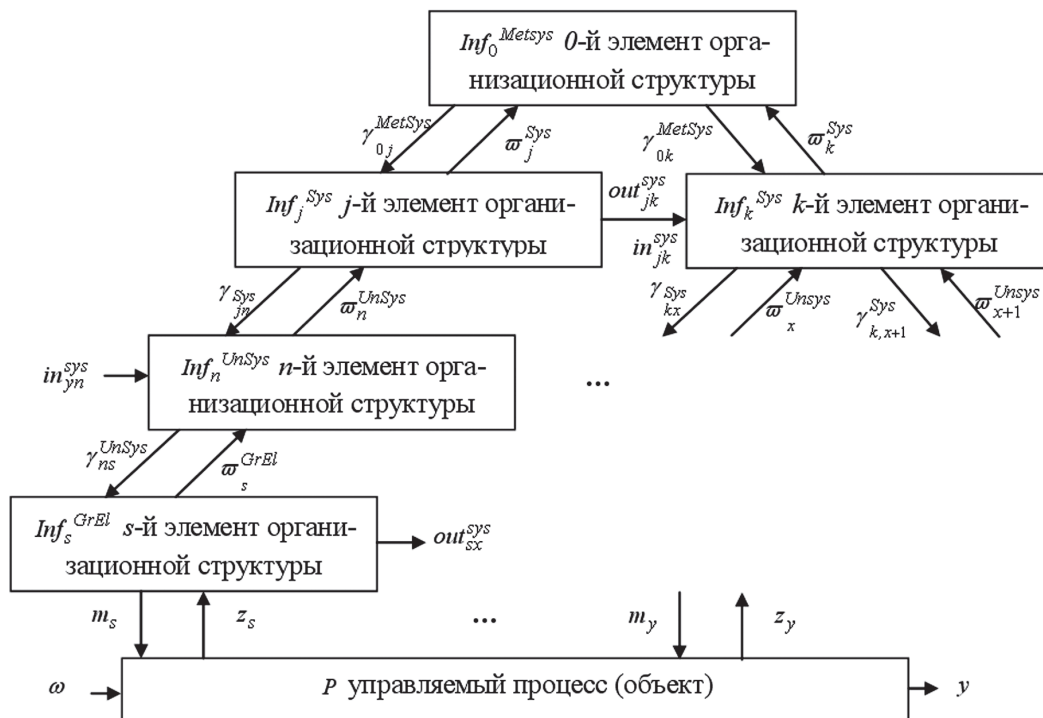


Рис. 5. Системная информационная модель

где g – нижний уровень иерархии, $h \neq Metsys, h \neq g$;
 $\bigcup_j \varpi_j^{h+1}$ – множество сигналов обратной связи от элементов организационной структуры, подчинённых Org_i ; $\bigcup_k \text{Inp}_{ki}^h, \bigcup_k \text{Out}_{ki}^h$ – множество внутриуровневых информационных сигналов (входящих и исходящих) от/до элемента Org_i .

Традиционный подход к исследованию многоуровневых организационных систем предполагает, что непосредственно с управляемыми процессами взаимодействуют только элементы нижнего уровня организационной структуры [6]. Однако на практике, например в отрасли телекоммуникаций, управляемые процессы и фрагмента организационной структуры СУР отрасли фактически взаимодействуют на всех уровнях организационной иерархии.

Рассмотрим данные подробно, для чего внесём изменения в системную модель, изображённую на рис. 5, введя дополнительные информационные сигналы обратной связи z_i^k от управляемого объекта (в данном случае совокупности предприятий отрасли) к i -ому элементу k -ого уровня иерархии аппарата управления и регулирования (см. рис. 6). Назовём данные сигналы «обратной связью второго рода» СУР отрасли.

Информационное обеспечение каждого элемента Org_i организационной структуры с учётом сигналов «обратной связи второго рода» тогда описывается выражениями (4-6):

$$\bigcup_j \gamma_{0j}^{MetSys} = Inf_0^{MetSys} \left(\bigcup_j \varpi_j^{Sys}, z_0^{MetSys} \right), \quad (4)$$

$$\left(\bigcup_j \gamma_{ij}^h, \bigcup_k \text{Out}_{ik}^h, \varpi_i^h \right) = Inf_i^h \left(y_{xi}^{h-1}, \bigcup_j \varpi_j^{h+1}, \bigcup_k \text{Inp}_{ki}^h, z_i^h \right), \quad (5)$$

$$\left(m_s, \varpi_s^g, \bigcup_k \text{Out}_{sk}^g \right) = Inf_s^g \left(y_{xs}^g, z_s, z_s^g, \bigcup_k \text{Inp}_{ks}^g \right), \quad (6)$$

где g – нижний уровень иерархии $h \neq Metsys, h \neq g$;

$\bigcup_j \varpi_j^{h+1}$ – множество сигналов обратной связи от элементов организационной структуры, подчинённых Org_i ; $\bigcup_k \text{Inp}_{ki}^h, \bigcup_k \text{Out}_{ki}^h$ – множество внутриуровневых информационных сигналов (входящих и исходящих) от/до элемента Org_i .

Возможное практическое содержание таких сигналов показано в табл. 2.

Таблица 2

Содержание информационных сигналов между элементами организационной системы и управляемым процессом. Часть 2

z_i^k – информационный сигнал «обратной связи второго рода» от управляемого процесса	– отраслевые конференции, «круглые столы»; – смешанные рабочие группы и научно-технические и общественные советы с участием представителей аппарата управления и предприятий отрасли; – неофициальные лоббистские контакты
--	--

Рассматриваемые информационные сигналы в зависимости от элемента структуры, к которому они направлены, несут разную смысловую нагрузку. Выбор сигнала «обратной связи второго рода» определяется управляемым процессом (предприятием отрасли) в зависимости от целей, на которые

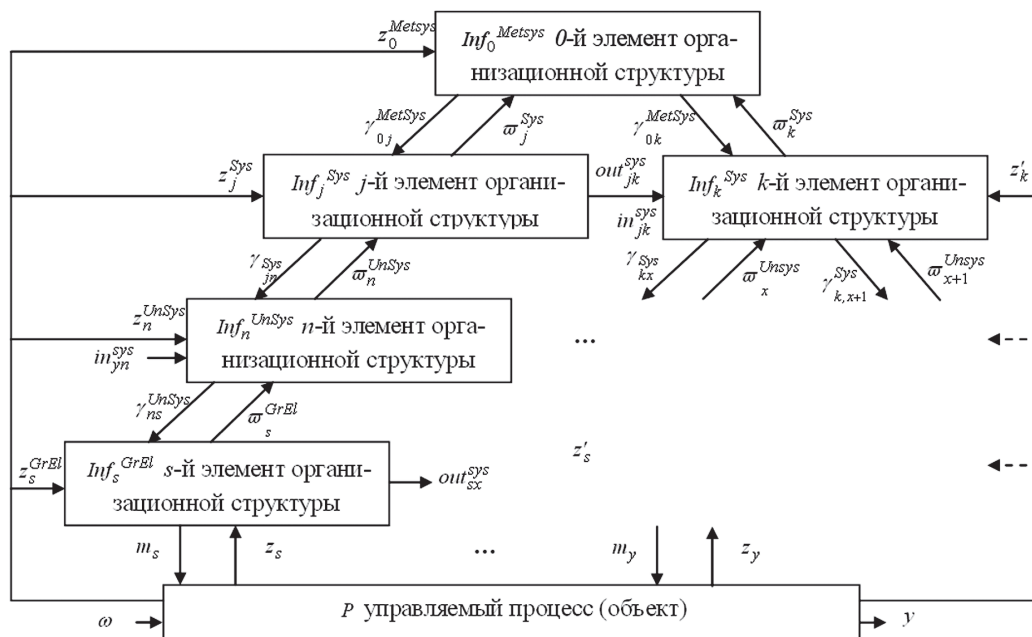


Рис. 6. Фрагмент системной информационной модели СУР отрасли с учетом обратной связи «второго рода»

Каждому компоненту системной алгоритмической модели организационной структуры ставим в соответствие элемент информационной модели, сохраняя межуровневые и внутриуровневые связи. Так, каждому решающему компоненту $Comp_i^h$ соответствует блок обработки информации $InfC_i^h$, отвечающий за обработку информационных сигналов, каждому исполнительному компоненту Exe_i^{h+1} соответствует блок обработки информации $InfE_i^{h+1}$ (см. рис. 8).

Системная информационная модель может быть записана как совокупность элементов, каждый из которых состоит из преобразователя информации Inf_i^h , входящих и исходящих сигналов, приведенных в табл. 2, 3. В терминах РССМ каждый i -ый элемент h -ого уровня иерархии будет описан выражениями (9-11):

$$R Inf_0^{MetSys} = \left[\begin{array}{c} (z_0^{MetSys} \vee e) \cdot Inf_0^{MetSys} \cdot \\ \left[\left[\gamma_1^{MetSys} \wedge \dots \wedge \gamma_n^{MetSys} \right] \wedge \varpi_0^{MetSys} \right] \end{array} \right], \quad (9)$$

$$R Inf = \left[\begin{array}{c} (z_0^{MetSys} \vee e) \cdot InfC_0^{MetSys} \cdot \left[\left[\gamma_1^{MetSys} \wedge \gamma_2^{MetSys} \wedge \gamma_3^{MetSys} \wedge \gamma_4^{MetSys} \wedge \gamma_5^{MetSys} \right] \right] \cdot \\ \left[\left[InfE_1^{Sys} \cdot \left[\left[InfE_2^{Sys} \cdot InfE_5^{Sys} \right] \wedge InfE_3^{Sys} \wedge InfE_4^{Sys} \right] \right] \right] \end{array} \right] =$$

$$= \left[\begin{array}{c} (z_0^{MetSys} \vee e) \cdot InfC_0^{MetSys} \cdot \left[\left[\gamma_1^{MetSys} \wedge \gamma_2^{MetSys} \wedge \gamma_3^{MetSys} \wedge \gamma_4^{MetSys} \wedge \gamma_5^{MetSys} \right] \right] \cdot \\ \left[\begin{array}{c} \left[\begin{array}{c} (z_1^{Sys} \vee e) \cdot InfC_1^{Sys} \cdot \left[\left[\gamma_1^{Sys} \wedge \dots \wedge \gamma_7^{Sys} \right] \right] \wedge \\ \left[\left[out_2^{Sys} \wedge \dots \wedge out_5^{Sys} \right] \wedge \varpi_1^{Sys} \right] \end{array} \right] \cdot \left[\begin{array}{c} Exe_1^{UnSys} \cdot \left[\begin{array}{c} Exe_2^{UnSys} \wedge \left[\begin{array}{c} Exe_4^{UnSys} \cdot Exe_3^{UnSys} \end{array} \right] \wedge \\ Exe_5^{UnSys} \wedge \left[\begin{array}{c} Exe_7^{UnSys} \cdot Exe_6^{UnSys} \end{array} \right] \end{array} \right] \end{array} \right] \\ \left[\begin{array}{c} (z_2^{Sys} \vee e) \cdot InfC_2^{Sys} \cdot \left[\left[\gamma_8^{Sys} \wedge \dots \wedge \gamma_{11}^{Sys} \right] \right] \wedge \left[\left[out_1^{Sys} \wedge out_3^{Sys} \wedge out_4^{Sys} \wedge out_5^{Sys} \right] \wedge \varpi_2^{Sys} \right] \cdot \\ \left[\begin{array}{c} Exe_8^{UnSys} \cdot \left[\begin{array}{c} Exe_9^{UnSys} \wedge Exe_{10}^{UnSys} \wedge Exe_{11}^{UnSys} \end{array} \right] \cdot Exe_5^{Sys} \end{array} \right] \end{array} \right] \wedge Exe_3^{Sys} \wedge Exe_4^{Sys} \end{array} \right] \end{array} \right] \quad (12)$$

$$R Inf_i^h = \left[\begin{array}{c} \left(z_i^h \right) \cdot Inf_i^h \cdot \\ \left[\begin{array}{c} \left[\gamma_1^h \wedge \dots \wedge \gamma_n^h \right] \wedge \\ \left[\begin{array}{c} out_1^h \wedge \dots \wedge out_{i-1}^h \wedge \\ \dots \wedge out_{i+1}^h \wedge \dots \wedge out_l^h \end{array} \right] \wedge \\ \varpi_i^h \end{array} \right] \end{array} \right], \quad (10)$$

$$R Inf_i^g = \left[\begin{array}{c} \left(z_i^g \right) \cdot z_i \cdot Inf_i^h \cdot \\ \left[\begin{array}{c} m_i \wedge \\ \left[\begin{array}{c} out_1^g \wedge \dots \wedge out_{i-1}^g \wedge \\ \dots \wedge out_{i+1}^g \wedge \dots \wedge out_l^g \end{array} \right] \wedge \\ \varpi_i^g \end{array} \right] \end{array} \right], \quad (11)$$

где c_i^h – условие, в соответствии с которым активируется обратная связь «второго рода» к i -ому элементу h -ого уровня; n – количество подчинённых элементов; l – количество элементов h -ого уровня организационной структуры; g – нижний уровень иерархии, $h \neq Metsys, h \neq g$.

Тогда информационная модель фрагмента СУР телекоммуникационной отрасли на уровнях $MetSys \rightarrow Sys$ запишется в терминах РССМ как (12).

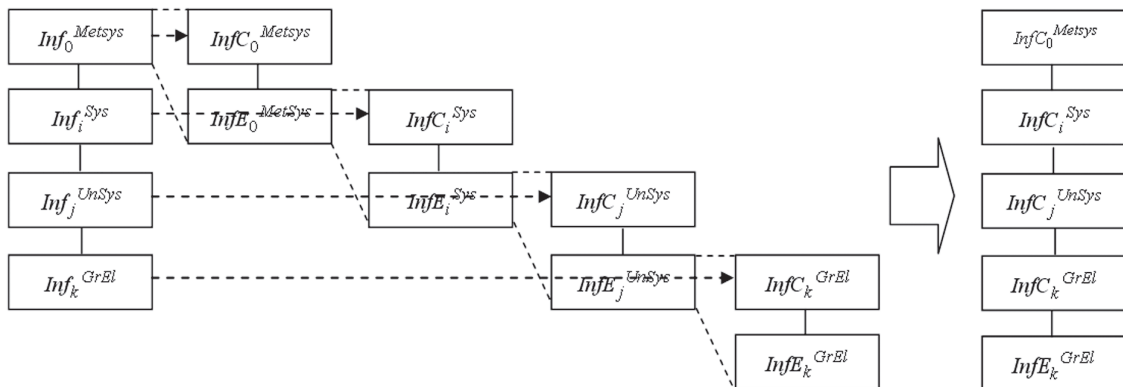


Рис. 8. Представление системной информационной модели в виде иерархии информационных блоков решающих и исполнительных компонентов

5. Применение системных моделей при исследовании процессов принятия решений

Полученные системные алгоритмическая и информационная модели являются основой для системного моделирования процессов функционирования и принятия решений в сложной организационной структуре СУР телекоммуникационной отрасли. Использование формализованного представления позволяет описать структуру, связи и зависимости между элементами исследуемой системы в виде, пригодном для использования в математических и компьютерных моделях. При этом отметим, что решающие компоненты элементов алгоритмической модели, блоки обработки информации соответствуют точкам принятия решений в фактической организационной структуре. Моделирование таких точек принятия решений (фактически локальных процессов принятия решений) может выполняться при помощи вероятностных моделей, основанных на методах нечеткой логики. Для этого входящие и исходящие информационные сигналы описываются как наборы их возможных состояний с вероятностью и задержкой времени их возникновения. Данные, полученные на входе элемента системной модели, обрабатываются решающими компонентами на основании заложенных правил. В свою очередь, правила формируются на основании положений законодательства. На выходе решающего компонента также получаем наборы возможных состояний исходящих сигналов с вероятностью и задержкой времени возникновения. Полученные таким образом системные модели позволяют применить методы динамического и имитационного моделирования процессов принятия решений.

На примере СУР телекоммуникационной отрасли показана *методика формирования системных моделей реальных систем принятия решений*, которая состоит в следующей последовательности действий:

1. Формирование структурной модели аппарата управления на основе доступной информации (штатное расписание, инструкции и т.п.).
2. Определение функций структурных единиц и формирование функциональной модели.
3. Формирование алгоритмической модели на основании функциональной с учетом логики выполнения задач на каждом уровне организационной иерархии. Представление элемента алгоритмической модели как совокупности решающей и исполнительных компонентов.
4. Формирование информационной модели на основании алгоритмической с межуровневых и внутриуровневых информационных связей. Представление элемента информационной модели как совокупности блоков преобразования информации решающего и исполнительных компонентов.

5. Формирование моделей локального принятия решений для решающих компонентов, например, на основе нечеткой логики.

Необходимо отметить, что важной особенностью рассматриваемых объектов является наличие многоуровневых информационных сигналов обратной связи «второго рода». Такое отличие от традиционных организационных систем обусловлено спецификой регулирования отрасли телекоммуникаций (и ряда других) в современном состоянии:

- либеральные принципы управления отраслью, отсутствие жестких связей управления предприятиями отрасли со стороны государства;
- быстрые темпы развития технологий, необходимость внедрения которых быстрее осознается в предпринимательской среде, т.е. внутри управляемого процесса (субъекта);
- ограниченность ресурсов, конкуренция за них между участниками отрасли.

Информационные сигналы обратной связи «второго рода» также представляются в виде набора возможных состояний в зависимости от задачи, решаемой управляемым процессом. Для формирования подобных наборов состояний предполагается использовать экспертные оценки.

Выводы

Представленный метод позволяет сформировать системные модели, используемые для анализа процессов принятия решений в сложных организационных системах. Данные модели позволяют исследовать эффективность организационной системы управления с учетом действующих в отрасли норм и правил, исследовать взаимодействие системы управления и управляемых процессов. На основе моделей имеется возможность оценить вероятности и длительности получения того или иного управляющего сигнала (управленческого решения) как результата последовательного принятия решений элементами организационной структуры.

Корректно построенные и детально разработанные системные модели допускают исследование эффективности аппарата управления отраслью при изменении действующего законодательства, норм и правил, применяемых в отрасли.

Дальнейшие исследования процессов принятия решений в сложных управляющих системах целесообразно развивать в таких направлениях:

- применение методов нечеткой логики для моделирования принятия решений в решающих компонентах;
- моделирование на основе игровых методов процессов конкуренции за ресурсы между управляемыми процессами и их влияние на принятие решений.

Список литературы. 1. Концепція розвитку телекомунікацій в Україні [Електронний ресурс]/ КМУ. — режим

доступа: <http://www.rada.gov.ua>. 2. *Спицнадель, В.Н.* Основы системного анализа [Текст]: учеб. пособие / В.Н. Спицнадель. — СПб.: «Издательский дом «Бизнес-пресса», 2000. — 326 с. 3. *Илюшко, В.М.* Модели и методы информационной технологии проектирования метасистем [Рук.]: дис. д-ра техн. наук: 05.13.06/ В.М. Илюшко. — Харьков, 1998. — 451 с. 4. Менеджмент в телекоммуникациях [Текст] / под ред. Н.П.Резниковой, Е.В.Деминой. — М.: Эко-Трендз, 2005. — 392 с. 5. *Илюшко, В.М.* Системное моделирование в управлении проектами [Текст]: монография / В.М. Илюшко, М.А. Латкин. — Харьков.: Национальный аэрокосмический университет им. Н.Е.Жуковского «ХАИ», 2010. — 220 с. 6. *Месарович, М.* Теория иерархических многоуровневых систем [Текст]: пер. с англ. / М.Месарович, Д.Мако, И.Такахара. — М.: «Мир», 1978. — 312 с. 7. *Саати, Т.* Аналитическое планирование. Организация систем [Текст]: пер. с англ / Т.Саати, К.Кернс. — М.: «Радио и связь», 1991. — 223 с. 8. *Мильнер, Б.З.* Системный подход к организации управления [Текст] / Б.З Мильнер, Л.И.Евченко, В.С Рапопорт. — М.: Экономика, 1983. — 224 с. 9. *Истомин, Е.П.* Теория организации: системный подход [Текст] / Е.П.Истомин, А.Г.Соколов. — СПб.: ООО «Андреевский издательский дом», 2009. — 314 с.

Поступила в редколлегию 28.05.2010.

УДК 519.876.2

Метод дослідження процесів прийняття рішень у держрегулюванні телекомунікаційної галузі / А.В. Калмыков // Біоніка інтелекту: наук.-техн. журнал. — 2010. — № 3 (74). — С. 42–51.

На прикладі системи керування галуззю телекомунікацій розглядаються питання аналізу процесів прийняття рішень у великих організаційних структурах. Для коректної декомпозиції такої складної соціальної й тех-

нічної системи на складові елементи використовується системний підхід та методи структурного аналізу. На базі даного підходу для дослідження процесів прийняття рішень у складних керуючих системах показано формування системних алгоритмічних й інформаційних моделей. Показана декомпозиція елементів системних моделей на вирішальний і виконавчий компоненти, обґрунтована багаторівнева структура зворотних зв'язків між апаратом керування й керованим процесом. Запропонована концепція моделювання процесів прийняття рішень в складних організаційних структурах. Показані практична значущість та перспективні напрямки досліджень процесів прийняття рішень у багаторівневих ієрархічних структурах.

Табл. 2. Іл. 8. Бібліогр.: 9 найм.

UDC 519.876.2

Research method of decision making processes at state adjustment of telecommunication industry / A.V. Kalmykov // Bionics of Intelligence: Sci. Mag. — 2010. — № 3 (74). — P. 42–51.

The paper considers questions of the analysis of decision making processes in a large organizational structure on an example of a management system of the telecommunication industry. For correct decomposition of such difficult social and technical system on constituting elements the structural analysis method is applied. According to it algorithmic and information models for research of decision making in difficult managing systems are applied. The approach implies decomposition of system models elements on solving and executive components. For it the multi-level structure of feedback between management system and managed process is proved. Approach for modeling of decision making processes in complex organization systems is proposed on this basis. Also the practical importance and perspective directions of researches of decision making processes in multi-level hierarchical structures are suggested in the article.

Tab. 2. Fig. 8. Ref.: 9 items.

УДК 656:004.75



ІНФОРМАТИЗАЦІЯ ТРАНСПОРТНОЇ ІНФРАСТРУКТУРИ, МАШИН ТА СИСТЕМ

О.П. Алексієв¹, В.О. Алексієв², В.О. Хабаров³, Г.Г. Четвериков⁴

¹ ХНАДУ, м. Харків, Україна, o.p.alex@mail.ru

² ХНАДУ, м. Харків, Україна, aleksiyev@gmail.com

³ НТЦ «Дор'якість», м. Харків, Україна, mechatronics@list.ru

⁴ ХНУРЕ, м. Харків, Україна, chetvergg@kture.kharkov.ua

Розглянуто інформаційний розвиток транспортних систем у єдиному інформаційному просторі транспортних організацій.

ТРАНСПОРТНА ІНФРАСТРУКТУРА, ІНФОРМАЦІЙНИЙ РОЗВИТОК, МЕХАТРОНІКА, ТЕЛЕМАТИКА, СИНЕРГЕТИКА, КОМП'ЮТЕРНІ СИСТЕМИ

Вступ

Зараз склалася ситуація, коли практика створення сучасних автомобільних приладів та пристроїв, агрегатів та систем випередила теорію інформаційного аналізу та синтезу транспортних систем. Існуючі окремі рішення з інформаційного забезпечення автомобільного транспорту потребують узагальнення, стандартизації та уніфікації, визначення нових спеціальних вимог до створення комп'ютерних обчислювальних систем та мереж на транспорті.

У зв'язку з постійним інформаційним розвитком суспільства та його промислової складової нові транспортні системи і машини досягли високого інформаційного рівня досконалості. Відповідно з'явилося нове протиріччя між стрімким розвитком засобів та методів інформатизації складних об'єктів і систем та гетерогенним характером існуючих підсистем та ланок транспортного комплексу. Вирішення цього протиріччя дозволить на всіх рівнях транспортної інфраструктури поліпшити обслуговування мешканців міст і регіонів, удосконалити перевізні процеси, уникнути існуючих негативних впливів: збоїв в організації руху, незадовільного стану шляхів сполучень, нераціонального використання коштів, що виділяються на ремонт, експлуатацію та облаштованість транспортних маршрутів.

1. Аналіз досліджень та публікацій

Можна виділити три послідовні етапи інформатизації транспортних систем: виконання комп'ютерних обчислень для вирішення окремих складових проблем та задач транспортного типу [1]; створення галузевих автоматизованих систем управління рухом, транспортними підприємства-

ми й організаціями, впровадження інформаційних технологій на транспорті [2–4]; обладнання транспортних засобів мехатронними вузлами, агрегатами та системами, розвиток транспортної телематики, інтелектуалізація транспортних машин, систем та шляхів сполучення [5, 6].

У дослідженні [7] саме інтелектуалізація, застосування на транспорті мехатроніки, синергетики та телематики розглянуто як наукова засада інформатизації транспортних систем. Що торкається сучасного протиріччя між досягнутим рівнем інформатизації суспільства, технічним удосконаленням транспортних машин та реальним інформаційним станом транспортної інфраструктури міст та регіонів, у роботах [8, 9] визначено його розв'язання шляхом створення єдиного інформаційного простору транспортних організацій. Розглянемо таке рішення як інформаційний розвиток транспортних систем.

2. Постановка задачі

Транспортна інфраструктура залежить та підпорядковується структурі телекомунікаційних зв'язків комп'ютерних систем, що забезпечують інформаційні послуги. Її основний компонент – телематика. Наслідком цього є необхідність застосування синергетичного підходу до вирішення проблем створення складних об'єктів та систем, які мають гетерогенний характер (рис. 1). Його відмінність від звичайного системного підходу полягає у необхідності урахування та використання механізму самоорганізації інформаційного простору транспортних систем.

На рис. 2 наведено схему, що пояснює взаємозв'язок синергетичного об'єднання телематичних комплексів та мехатронних вузлів із інтелектуальними складовими транспортної інфраструктури.

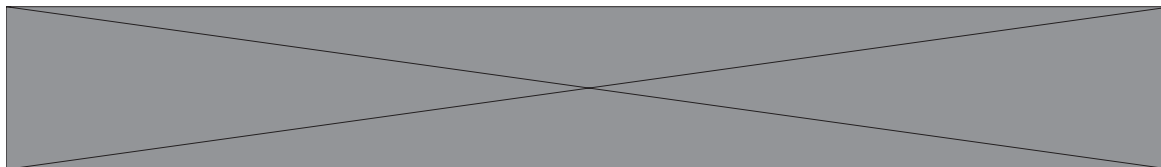


Рис. 1. Синергетичний підхід



Рис. 2. Співвідношення мехатроніки, телематики та синергетики

Такий зв'язок ґрунтується на стандартизації програмно-апаратних рішень та уніфікації інтерфейсів і протоколів передачі даних на засадах синергетики, телематики та мехатроніки. Мехатронні системи стали основною частиною транспортного середовища. Їх створення та розвиток відповідають процесам вдосконалення та оптимізації транспортної системи (ТС). Розглянемо інформаційний розвиток ТС як перетворення Q існуючого стану транспортної інфраструктури (сукупності n змінних $k_1(t), k_2(t), \dots, k_n(t)$) до нової сукупності n змінних $q_1(t), q_2(t), \dots, q_n(t)$. Прообрази $k_i(t)$ оператора Q – характеристики пересування у просторі та часі транспортних засобів, а $q_i(t)$ – нова сукупність оцінок цього пересування. Кожна змінна нової сукупності є точкою, що визначає найбільш раціональну концепцію розвитку ТС. Тоді Q є таким оператором, що $q_i(t) = Q[k_i(t), t]$, де $k_i(t)$ – динамічна функція, яка відповідає i -му об'єкту у досліджуваній системі незалежно від його специфіки. Інтерпретація $q_i(t)$ – ефективність роботи i -частини системи. Подання цього параметра не як пропускної здатності, а як ефективності узагальнює властивості розвитку транспортної системи.

Припустимо, L – множина точок транспортної комунікації (шляху сполучення (ШС) у його лінійному уявленні), що складається з підмножин L^α, L^p, L^y відповідно зареєстрованих значень поточних параметрів швидкості руху, наповненості рухомої одиниці (РО) та ШС. Таким чином, $L = L^\alpha \cup L^p \cup L^y$. Відповідно стан рухомого об'єкта, процес його пересування в просторі L транспортного засобу і характеристики дорожнього середовища можна уявити як суму трьох операторів $S = S' \cup S'' \cup S'''$, де S, S', S'', S''' – оператори перетворення спостережуваних динамічних процесів $k_i(t)$ (СДП) пересування рухомих об'єктів (РО) у певний момент часу t (поточні характеристики РО, ШС).

Таким чином, створюється своєрідний ланцюжок від інформаційних приладів окремої транспортної машини, систем внутрішньої та зовнішньої транспортної телематики до потужних комп'ютерних

систем на базі локальних обчислювальних мереж транспортних організацій, органів самоврядування та державних інформаційних систем на усіх рівнях світової Глобальної комп'ютерної розподіленої мережі. Розглянемо задачу сполучення та об'єднання комп'ютерних ресурсів для забезпечення інформаційного розвитку транспортних систем.

3. Транспортні процеси та системи

В основу узагальнення оцінки транспортних машин, систем, ШС та транспортних процесів (ТП) покладено механізм взаємодії транспортного засобу (ТЗ) та ШС, який слід вважати керованим процесом. Кожен ТЗ характеризується своїм поточним напрямком пересування й маршрутом руху. Визначення режиму пересування такого ТЗ відбувається на основі аналізу ситуацій на відповідному ШС, прогнозу умов руху та супроводжується формуванням керуючих впливів.

На перше місце у керуванні транспортними процесами виходить сукупність дій осіб, що приймають рішення (ОПР) про організацію руху. Пересування окремого АТЗ теж підпорядковується відповідним діям ОПР (рис. 3).

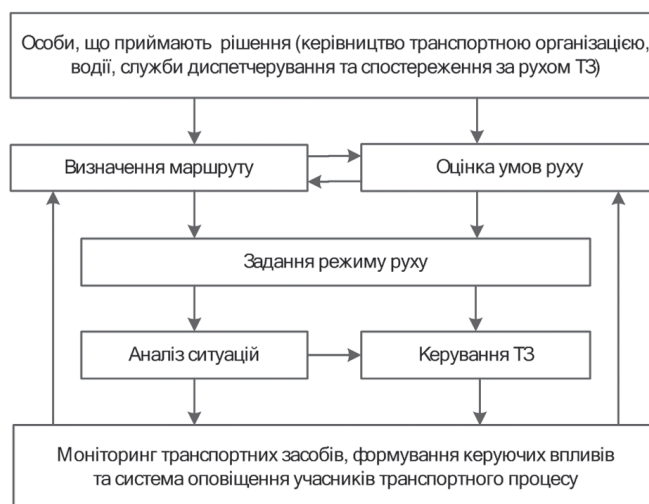


Рис. 3. Транспортний процес

Аналіз транспортних процесів передбачає їх уявлення як послідовності дій учасників руху ТС, яку формально визначимо як множину $T = Z \cup K$, де Z – засоби пересування, K – шляхи сполучення. Логічним продовженням такого узагальнення є використання інформаційно-логічної моделі (ІЛМ) транспортно-експлуатаційних параметрів.

Рух окремого транспортного засобу визначається за результатами оцінки ШС та моніторингу стану ТЗ (рис. 4). Йому відповідає множина транспортно-експлуатаційних параметрів, які так чи інакше впливають один на одного.

Можна показати, що аналогічний висновок випливає із аналізу ІЛМ взаємозв'язку цих параметрів $\{d_i\}$. Відповідна матриця D відбиває змістовну складову моніторингу транспортного процесу.

Якщо розглядати граф, вузлам якого відповідають параметри $\{d_j\}$, використані для оцінки складної автоматизованої системи, а дугам – зв’язки між ними, то оцінку аналізованої множини об’єктів можна виконати за їх структурною складністю D_i :

$$D_i = \sum_j b_{ij}, \quad (1)$$

де $b_{ij} \in \{1,0\}$, 1 – параметри зв’язані; 0 – параметри незв’язані.

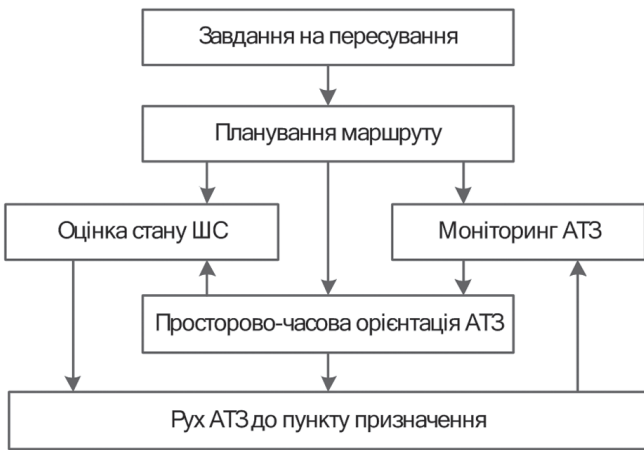


Рис. 4. Рух окремого транспортного засобу

За результатами розрахунку складності задач можна стверджувати, що для прийняття рішень необхідно з первинних параметрів оцінювати швидкість руху, геометрію дороги, її рівність і зчпні якості.

Тоді комплексна узагальнена оцінка K_i стану частини транспортної магістралі або варіанта i маршруту руху АТЗ є середньозваженим коефіцієнтом:

$$K_i = \frac{1}{n} \cdot \sum_j \sum_\gamma \sum_l b_{ij\gamma}^k \cdot \frac{l_{ij\gamma}}{L_{ij}}, \quad (2)$$

де $b_{ij\gamma}^k$ – обмірюване (розраховане і нормоване та зведене до єдиної системи відліку) значення параметра k по γ ділянці дороги j , що входить в i -й маршрут; L_{ij} – довжина дороги j на i -му маршруті; n – кількість параметрів, за якими ведеться оцінка альтернативних маршрутів руху. Індекси i, j, γ відповідно визначають: варіант маршруту, включену до маршруту дорогу та номер елемента.

Знаючи значення K_i оцінки множини альтернативних маршрутів, можна вибрати на основі теорії математичного програмування найбільш раціональний напрямок руху. Інтелектуалізація транспортних засобів і систем, створення спеціальної апаратури, що має забезпечувати рішення функціональних задач організації руху, нерозривно пов’язані зі схематичними рішеннями в цій сфері та цифровою обробкою даних. Відповідний простий математичний опис такої обробки даних – лінійна система з постійними параметрами, яка є співвідношенням вхідної і вихідної послідовності $x(n)$ та $y(n)$ відповідно:

$$y(n) = Y[x(n), t], \quad (3)$$

де Y – оператор перетворення прообразу оператора $x(n)$ в образ оператора $y(n)$, визначені на множині n відкликів в області T поточного часу t .

Слід зазначити, що в даному випадку $Y: T \rightarrow T$, де T – числова вісь, у відповідність якій ставиться час t . У загальному випадку T – простір, на якому існують свої множини часових відкликів T_i зі своїм розрізненням у часі ξ_{T_i} , причому $\bigcup T_i = T$, але $\bigcap T_i = \emptyset$ (пуста множина). У реальних системах одновимірне перетворення (3) трансформується у двовимірні перетворення, а у більш загальному випадку – у багатовимірні перетворення послідовностей сигналів, що реєструються. Це є співвідношенням типу згортка:

$$y(n) = \sum_{m=-\infty}^{\infty} h(m) \cdot x(n-m), \quad (4)$$

де $h(m)$ – відклик на одиничний стрибок або імпульсну характеристику відповідної системи реєстрації даних.

За аналогією з одновимірною лінійною системою для двовимірної системи з вихідною послідовністю $y(n_1, n_2)$ і вхідною – $x(n_1, n_2)$ справедливо:

$$y(n_1, n_2) = \sum_{m_1=-\infty}^{\infty} \sum_{m_2=-\infty}^{\infty} x(m_1, m_2) \cdot h(n_1 - m_1, n_2 - m_2), \quad (5)$$

де $h(n_1, n_2)$ – імпульсна характеристика такої лінійної системи. Тоді:

$$y(n_1, n_2, \dots, n_i, \dots, n_k) = \sum_{m_1=-\infty}^{\infty} \sum_{m_2=-\infty}^{\infty} \dots \sum_{m_i=-\infty}^{\infty} \dots \sum_{m_k=-\infty}^{\infty} x(m_1, m_2, \dots, m_i, \dots, m_k) \times (6)$$

$$\times h(n_1 - m_1, n_2 - m_2, \dots, n_i - m_i, \dots, n_k - m_k).$$

Відповідно можна навести операторний опис цієї оцінки:

$$y(n_i) = Y_i[x(m_i), t], \quad (7)$$

де Y_i – оператор k -вимірної лінійної системи цифрової реєстрації й обробки даних про стан транспортних засобів.

Для того щоб фізично реалізувати перетворення Y_i , необхідним є виконання умов фізичної реалізації і стійкості створюваної системи моніторингу транспортного комплексу: $y(n_0)$ залежить тільки від $x(n_0)$, якщо $n \leq n_0$. Для відповідності системи вимогам стійкості необхідно передбачити, що

$$\sum_{m=-\infty}^{\infty} |h(n)| < \infty.$$

На часовому просторі T послідовності t_i можуть належати зонам з різним часовим розрізненням. Так, транспортний засіб зі своїм часом існує в діапазоні T_1 , водій і пасажир – T_2 , а транспортна комунікація – T_3 . Твердження $T_1 \cup T_2 \cup T_3 = T$ не завжди є справедливим. Можна визначити такий діапазон T_4 з T , для якого:

$$\forall T_4 | \exists T_4 = T_1 \cap T_2 \cap T_3. \quad (8)$$

Слід брати до уваги, що прообрази СДП ТП є функціями часу. Оцінка $q(t)$ транспортної системи в цілому (результат загального спостереження) визначається операторним співвідношенням:

$$q(t) = Q[h(l), t], \quad (9)$$

де $h(l)$ – динамічна функція аргументу l , що визначається в часі t у часовому просторі T_3 .

Оператор P є оператором із пам'яттю, тому що в зазначеній постановці важливими є не тільки прості визначення або оцінка $h(l)$ у точці $l = l_i$, що відповідає точці $s = s_i$ та моменту часу $t \in T$, але й аналіз ансамблю значень $\{x(t), y(t), z(t)\}$ на множині звітів x_j зі своєю «вагою» M_i або (у загальному випадку функції) коефіцієнта переваги. Процедура сканування ШС для $l = l_i$ буде мати відповідний вигляд:

$$X_j = \frac{\sum_{j=i-n}^{j=i+n} M_i X_j}{\sum_{j=i-n}^{j=i+n} M_i}. \quad (10)$$

Для визначення метрики простору існування множини функцій, що реєструються, СДП $x(t)$ слід розглядати у вигляді сукупності точок x_i на числовій осі. Розглянемо, як система моніторингу (підсистема – «спостерігач») розрізняє ці точки, визначаючи процес зміни станів. Для будь-якої пари x_i та x_j ($i \neq j$) можна вказати такі величини ξ , для яких значення x_i стають невизначеними або втрачають свій фізичний зміст. Ця величина визначає поріг розрізнення станів, причому для конкретного елемента можна одержати безліч значень порогів розрізнення. Величина найменшого порогу ξ визначає нижню межу надійного розходження двох значень СДП $x(t)$ або їхнього існування. Ця величина – поріг розрізнення: $\xi = \inf \sup \rho(x_i - x_j)$, для всіх $i \neq j$.

Слід відрізнити фізичне і споживче розрізнення. Точність перетворення даних характеризується «споживчим» порогом розрізнення ξ_y , а $x(t)$ – «фізичним» порогом розрізнення ξ_x і характеризує чутливість системи. При цьому умова спостереження об'єкта $\rho(x_i, x_j) \geq \xi_y$, тому що у протилежному випадку неможливо буде відмітити зміну стану процесу, який спостерігається.

Тоді графічно можна представити це перетворення у вигляді узагальненої схеми, що наведена на рис. 5.

Слід визначити, що завжди є така часова послідовність t_i , протягом якої вказівки «керуючого центру» $I(t)$ відсутні і керуючий вплив $U(t)$ формується автономно (на борту транспортного засобу) за оцінкою $Q(t)$:

$$\exists t_i \forall t_i \in T | Q(t) = Q_i[I(t), U(t)], I(t_i) = 0. \quad (11)$$

Поряд з цим можна висловити твердження про те, що контроль вузлів, агрегатів та модулів транс-

портного засобу має виконуватися рівномірно лише в окремому випадку. Вимірювання та оцінку параметрів керованої ланки мехатронної системи слід проводити у той час, коли вплив цієї ланки на роботу системи є найбільш відчутним.

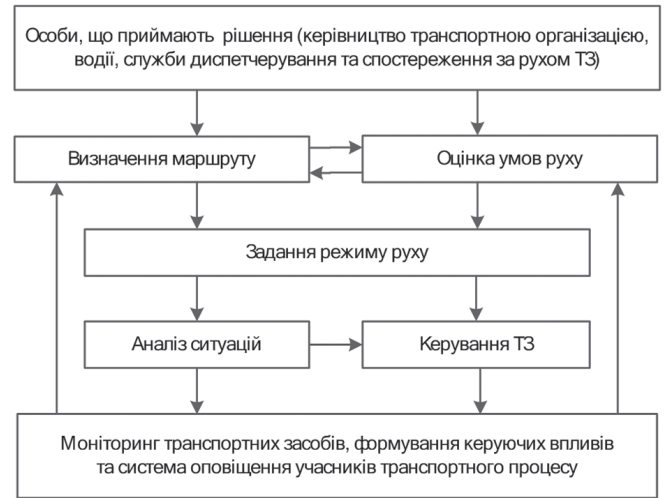


Рис. 5. Узагальнена схема оцінки стану транспортної системи

З визначення інформаційного розвитку ТС як мехатронізації транспортних машин та систем, їх інтелектуалізації поряд із застосуванням технології *X-by-Wire* можна стверджувати, що транспортна машина або РО мають стати своєрідною мережею – мережевим транспортним засобом (*Network Vehicle*). Відповідно за аналогією ТС включно з ШС являють собою технічний аналог нейронної мережі. Тому доцільно для розробки алгоритмів підготовки керуючих рішень на різних рівнях транспортної системи застосувати математичний апарат штучних нейронних мереж (ШНМ). З інженерної точки зору ШНМ – паралельно розподілені системи обробки інформації, утворені простими обчислювальними вузлами та мають властивість до накопичення експериментальних знань. Їх відмінною особливістю є наявність процедури навчання T . Таку процедуру можна представити як послідовність кроків: розрахунок результату; порівняння з точним значенням; вимірювання ваги; оцінка результатів перетворювань відповідно до досягнення потрібного рівня точності, а в негативному випадку повторення послідовності спочатку (розрахунок результату).

Ефективність застосування ШНМ для рішення задач у транспортних системах пояснюється тим, що для таких, порівняно з існуючими у природі, систем характерна врівноваженість, здатність до відбудови, можливість самоорганізації, росту, розвитку, узгодженості всіх складових частин та наявність механізмів компенсації зовнішніх впливів.

4. Впорядкованість та інформаційний розвиток

Питання про оптимальну впорядкованість та організацію особливо гостро стоїть при досліджен-

нях глобальних проблем – енергетичних, екологічних та багатьох інших, що потребують залучення великих ресурсів. Тут немає можливості шукати відповідь методом проб та помилок, а «нав'язувати» системі необхідну поведінку є дуже складним завданням. Набагато доцільніше діяти, спираючись на знання внутрішніх властивостей системи, законів її розвитку. В такій ситуації значення законів самоорганізації, формування упорядкованості у фізичних, біологічних та штучних нейронних системах важко переоцінити. Синергетика як метод стає інструментом пошукової діяльності. Вона виходить далеко за межі вузької спеціалізації при застосуванні загальних теоретичних моделей складної поведінки. Самоорганізація – одне з ключових понять синергетики. В аспекті утворення – це означає самоутворення. Відповідно найкраще керування – це самокерування.

Синергетичне об'єднання різних транспортних засобів, систем та комп'ютерних ресурсів для підвищення рівня інформаційного забезпечення транспортних систем ґрунтується на раціональному сполученні програмних, апаратних і математичних засобів для створення мехатронних систем на транспорті.

Висловимо твердження, що зміна функціональних та кількісних характеристик складових частин, вузлів або агрегатів транспортних машин не надає значного впливу на удосконалення їхніх конструктивних особливостей. Тому для створення концептуальних автомобілів, транспортних роботів важливими є моделі, які відрізняються не кількісними, а якісними характеристиками.

Відомий закон Мура та принцип Макімото слід розглядати як частковий випадок більш загального принципу синергетичного об'єднання складних об'єктів та систем.

Слід визначити цей принцип як невпинний розвиток. Логічним продовженням цього принципу є принцип інтелектуалізації – додання транспортній системі властивостей розумної поведінки.

Основною властивістю мехатронних транспортних систем, що підтверджує принцип невпинного розвитку, є співвідношення між спеціалізацією, уніфікацією та універсалізацією систем, вузлів та агрегатів транспортних засобів. Прототипом своєрідного інтелектуального засобу є універсаль-

на мобільна автотранспортна лабораторія (УМТЛ), що створена для доведення вірогідності зазначених вище тверджень та принципів (рис. 6).

Загалом, розвиток будь-якої транспортної системи базується на отриманні даних для формування динамічного банку відомостей про її попередній та поточний стан. Формування таких відомостей має забезпечувати інтелектуальна система моніторингу, що є аналогом порівняно простих систем опитування мешканців міст та регіонів, що охоплені транспортною інфраструктурою. Це є задачами, які вирішує транспортний ситуаційний центр (ТСЦ) за допомогою УМТЛ (рис. 7).

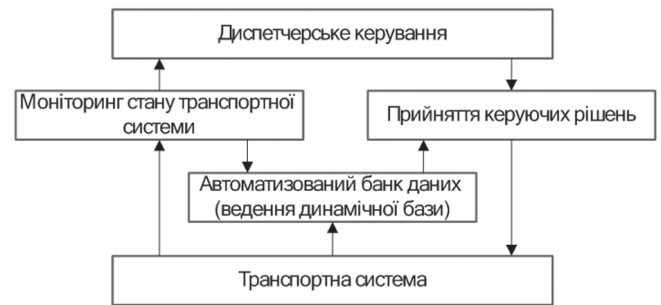


Рис. 7. Задачі, що вирішуються ТСЦ

Інформаційному розвитку транспортної інфраструктури відповідають чотири рівні синергетичного об'єднання транспортних машин, систем та їх програмно-апаратних засобів керування ТП:

- мехатронні модулі автотранспортних машин або інформаційно-керуючі автомобільні вузли та агрегати;
- телематичні системи, що дозволяють комплексно вирішувати проблеми просторово-часової орієнтації автотранспортних засобів;
- інтелектуальні транспортні технології, які забезпечують керування транспортними потоками та підтримують відповідний рівень сучасної транспортної інфраструктури міст чи регіонів;
- логістичні транспортні комплекси, в яких елементи, що керуються, є інтелектуальними транспортними засобами.

Окремим рішенням інформатизації роботи усіх транспортних організацій є комплексний розвиток та управління рухом усіх видів ГПТ міста (автобус, трамвай, тролейбус). Відбиття інформації про цей розвиток є змістом своєрідного порталу – мереже-

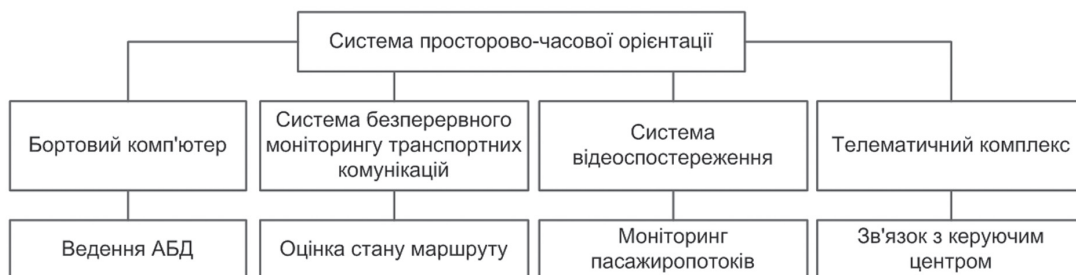


Рис. 6. Структурна схема обладнання УМТЛ

вого простору, у якому “існують” усі учасники дорожнього руху. У якості фізичної реалізації такого простору виступає розподілена система управління транспортним комплексом (рис. 8).

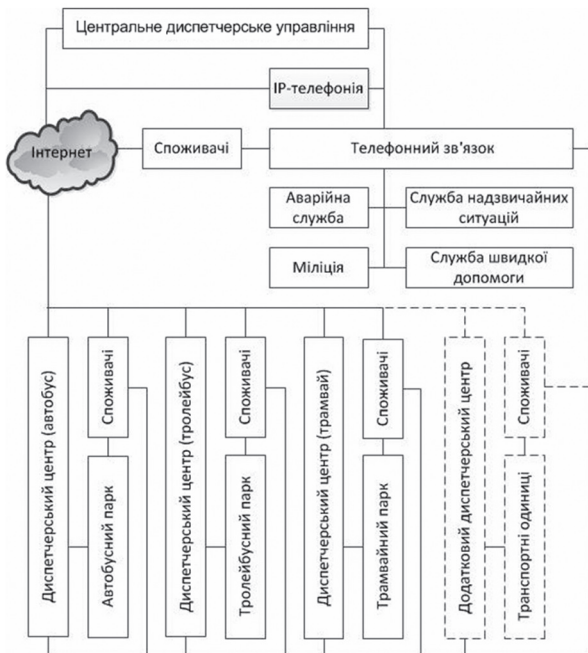


Рис. 8. Єдиний інформаційний простір.
Технічна реалізація

На першому етапі реалізації відповідного проекту джерелом інформації про стан транспортної системи може бути одна (або декілька УМТЛ), що виступає у ролі мобільного диспетчерського пункту, системи моніторингу маршрутної мережі та інструментального засобу апробації проектних рішень з створення внутрішньої та зовнішньої телематики РО.

Висновки

На основі оцінки протиріччя між стрімким розвитком засобів та методів інформатизації складних об'єктів та систем, їх уніфікації та гетерогенним характером існуючих підсистем та ланок транспортного комплексу, його інформаційного забезпечення передбачено втілення розрізаних комп'ютерних ресурсів транспортних організацій до єдиного інформаційного простору. Завдяки можливості використання існуючих гетерогенних комп'ютерних систем та Internet для автоматизації управління транспортом значно підвищується інформативність, технологічність, скорочується термін впровадження та як наслідок зменшення витрат на комп'ютеризацію підсистем та ланок транспортного комплексу (до двох разів порівняно з існуючими окремими рішеннями з автоматизації керування рухом).

Є істинними формальні висловлювання про необхідність синергетичного об'єднання комп'ютерних ресурсів всіх учасників дорожнього руху від окремої транспортної машини до корпоратив-

ного рівня транспортних систем. В основу такого об'єднання покладено застосування мехатроніки, телематики і синергетичного підходу до створення мережних транспортних засобів та систем.

Список літератури: 1. Алексієв, О.П. Нова інформаційна технологія управління дорожньою галуззю [Текст] / О.П. Алексієв та ін. – Київ: НМК ВО, 1992. – 212 с. 2. Шалабаєв, Е.В. Теоретические и практические проблемы развития мехатроники [Текст] / Е.В. Шалабаєв // Современные технологии. – СПб., 2001. – С. 46–67. 3. Алексієв, В.О. Мехатроніка транспортних засобів та систем [Текст] / В.О. Алексієв, В.П. Волков, В.І. Калмиків. – Харків: ХНАДУ, 2004. – 176 с. 4. Інтелектуальні технології організації руху пасажирського транспорту міста [Текст] / А.М. Туренко, В.О. Богомолов, О.П. Алексієв, В.О. Алексієв // Автомобільні дороги та дорожні будівництва. – Київ: УТУ, 2004. – Вып. 4. – С. 305–311. 5. Информационные технологии на автомобильном транспорте [Текст] / В.М. Власов, В.Б. Николаев, А.В. Постолиит, В.М. Приходько – М.: МАДИ (ГТУ), 2006. – 283 с. 6. Пржибил, П. Телематика на транспорте [Текст] / П. Пржибил, М. Свитек – М.: МАДИ (ГТУ), 2003 – 540 с. 7. Алексієв, В.О. Управління розвитком транспортних систем : монографія / В.О. Алексієв. – Харків: ХНАДУ, 2008. – 268 с. 8. Богомолов, В.О. Проблема створення єдиного інформаційного простору транспортних організацій [Текст] / В.О. Богомолов, В.О. Алексієв // Автомобільний транспорт : Сборник науч. трудов. – 2009. – Вып. 25. – С. 222–225. 9. Богомолов, В.О. Концептуальне обґрунтування та синергетичний підхід до розвитку транспортних систем [Текст] / В.О. Богомолов, В.О. Алексієв // Інформаційно-керуючі системи на залізничному транспорті : Наук.-техн. журнал. – 2009. – № 5 (78). – С.59–63.

Надійшла до редколегії 01.06.2010.

УДК 656:004.75

Информатизация транспортных систем / О.П. Алексієв, В.О. Алексієв, В.О. Хабаров, Г.Г. Четвериков // Бионика интеллекта: науч.-техн. журнал. – 2010. – № 3 (74). – С. 52–57.

Предложено решение проблемы информатизации транспортных систем на основе мехатроники, телематики и основных принципов синергетического объединения всех составляющих транспортной инфраструктуры городов и регионов. Его реализация обеспечит необходимой информацией всех участников дорожного движения, перевозочные процессы и на всех уровнях транспортной инфраструктуры улучшит транспортное обслуживание населения городов и регионов.

Ил. 8. Библиогр.: 9 назв.

UDC 656:004.75

Informatization of transportation systems / O.P. Alekseyev, V.O. Alekseyev, V.O. Habarov, G.G. Chetverikov // Bionics of Intelligence: Sci. Mag. – 2010. – № 3 (74). – P. 52–57.

Proposed solution to the problem of informatization transport systems based on mechatronics, telematics and the basic principles of the synergistic integration of all components of the transport infrastructure of cities and regions. Its implementation is fully provide the necessary information to all road users, the transportation process and at all levels of transport infrastructure will improve transport services for the population of cities and regions.

Fig. 8. Ref.: 9 items.

УДК 004.932.2:004.93'1



СТРУКТУРНО-ИЕРАРХИЧЕСКИЕ МЕТОДЫ РАСПОЗНАВАНИЯ ИЗОБРАЖЕНИЙ ПРИ ВЛИЯНИИ ПРОСТРАНСТВЕННЫХ ПОМЕХ

В.А. Гороховатский

ХНУРЭ, г. Харьков, Украина, gorohovatsky-v@rambler.ru

Рассмотрена проблема распознавания изображений объектов при действии помех пространственного типа. Приведены результаты разработки структурно-иерархических методов, которые обеспечивают высокую вероятность правильного распознавания по сравнению с традиционными подходами. Обсуждаются пути оптимизации и оценки помехозащищенности, приводятся результаты экспериментов по распознаванию изображений реальных объектов.

РАСПОЗНАВАНИЕ ИЗОБРАЖЕНИЙ, СТРУКТУРНО-ИЕРАРХИЧЕСКИЕ МЕТОДЫ, ПРОСТРАНСТВЕННЫЕ ПОМЕХИ, СТРУКТУРНОЕ ОПИСАНИЕ ВИДЕО-ОБЪЕКТОВ

Введение

Существующие методы анализа объектов на изображениях реальных визуальных сцен не позволяют в достаточной степени эффективно осуществить распознавание в условиях искажения отдельных частей объектов, т.к. модель их построения принципиально не учитывает влияния внешних воздействий, приводящих к формированию частичных и ложных описаний [1]. На рис. 1 демонстрируется действие фона и локальных помех как разновидностей пространственных помех. Искажение уже 30% описания видео-объекта существенно снижает вероятность его идентификации до уровня 0,6–0,7 [2]. Необходима разработка более универсальных структурных подходов, в основу которых должен быть положен иерархический анализ описаний пространственных признаков объектов и возможность принятия решения по представлению, включающему ложные структуры [3].

Для успешной реализации обсуждаемого подхода необходимо решение ряда самостоятельных задач, от которых в конечном итоге зависит эффективность распознавания. Для получения структурного представления нужно разработать надежные методы выделения структурных элементов (СЭ) в виде характерных признаков (ХП, англ. Keypoints), которые отражают локально-пространственные свойства объектов в виде структурного описания. Вторая задача связана с достоверным установлением сходства СЭ объекта и эталонов. Третья состоит в выборе способа функционального интегрирования элементарных сходств для формирования иерархического подобия как основы глобального решения о классе объекта. Функциональное интегрирование состоит в агрегировании и логической обработке множества локальных сходств СЭ и построении на этой основе функционала структурного сходства описаний [4]. Обсуждаемые методы опираются на структурное стратифицированное представление, характеризуются многоэтапной обработкой с включением иерархического анализа и названы структурно-иерархическими методами

(СИМ).

Цель работы – развитие теоретических основ построения и экспериментальное исследование СИМ как средства для эффективного решения проблемы распознавания видео-объектов при влиянии геометрических преобразований и искажений, вызванных воздействием пространственных помех.

Основной научный результат состоит в теоретическом обосновании принципа построения СИМ для обеспечения высокой достоверности распознавания объектов в реальной обстановке внешней среды.

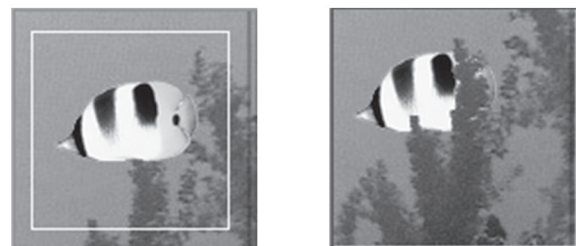


Рис. 1. Действие фона (а) и фона совместно с локальными помехами (б)

1. Принцип построения СИМ

Каждая точка (x, y) изображения $B(x, y)$ – это объект $B_0[T(x, y)]$, полученный из эталона $B_0(x, y)$ путем геометрических преобразований T , либо фон (локальная помеха), на которые дополнительно накладывается аддитивный шум. Функцию изображения $B(x, y)$ в целях структурного анализа представим в виде конечного множества фрагментов $\{b_k(x, y)\}$ или множества $\Lambda = \{\lambda^k\}$ ХП. При знаку λ^k соответствует элемент $c^k = (x^k, y^k)$ множества координат $C = \{c^k\}$, который определяет пространственные свойства описания. Визуальный объект представляется в виде кортежа $Z = \langle \Lambda, C \rangle$. Получение множества ХП как описания объекта направлено на решение проблемы распознавания и является ее важным этапом. От эффективности формирования ХП напрямую зависит результат распознавания.

Путем перехода к множеству ХП или множеству информативных фрагментов удается достичь существенного снижения объема информации, в то время как характеристики распознавания (достоверность, помехозащищенность) при этом практически не снижаются. В пространстве ХП данные рассматриваем как мультимножества [5], т.к. значения ХП бывают близкими между собой даже внутри одного описания. В большинстве применений ХП представляет собой числовой вектор $\lambda^k = (\lambda_1, \dots, \lambda_n)^k$, $\lambda_i \in R^1, i = \overline{1, n}$, $\lambda^k \in R^n$. В некоторых предметных областях типы ХП классифицируют путем присвоения меток [1, 5].

Далее нужно оценить эквивалентность объектов по их структурным описаниям Z_i, Z_m , что сводится к установлению соответствий ХП и получением на их основе значения глобального соответствия [3]. Наличие геометрических преобразований значительно усложняет анализ соответствий ХП. Проблема инвариантности актуальна как на локальном уровне формирования ХП, так и на этапе сопоставления описаний. Существующие подходы решают проблему распознавания путем «согласования» описаний в некотором признаковом пространстве, что соответствует значению критерия – функционала близости $\vartheta(Z_i, Z_m)$.

Наиболее распространена идея получения меры $\vartheta(Z_i, Z_m)$ в виде оценки части общих элементов. Получение точного значения меры с учетом значений и структурных взаимоотношений ХП представляет собой сложную в вычислительном плане проблему. В качестве компромисса целесообразно рассмотреть меры, основанные на принципах голосования СЭ, в которое могут быть включены как ХП, так и их отношения, формирующие многоуровневые признаки. Голосование основано на оптимальных статистических решениях относительно ХП и позволяет отделить множество ХП объекта от ложных элементов. Подход представляется универсальным в силу учета многообразия возможных ситуаций, связанных с видом изображений, эталонов, сегментов, фоновых образований, помех.

Множество отношений $R_Z \subseteq Z \times Z \times \dots \times Z$ для ХП опишем в виде предиката $R_Z : Z \times Z \times \dots \times Z \rightarrow \{0, 1\}$, где 1 означает истинность R_Z для набора аргументов. В силу простоты описания и анализа распространены бинарные отношения $R_Z \subseteq Z \times Z$. Имеем три типа отношений ХП: R_Λ – на множестве Λ , R_C – на множестве C , а также комбинированные отношения $R_Z = R_{\Lambda C}$ вида $([\lambda^i, c^i], [\lambda^k, c^k])$, заданные на множестве пар признаков – координаты. Отношения R_Λ реализуют групповые свойства ХП, а R_C и R_Z отражают пространственные связи ХП [3, 6]. Использование отношений развивает систему распознавания путем дополнения ее ресурсов многоуровневыми призна-

ками. Комбинации признаков в виде отношений и построение решений на использовании групп ХП заложены в природе структурного представления как способа грануляции данных и получения общих знаний о распознаваемом объекте [6, 7].

Концепция применения СИМ при распознавании основана на оптимизации близости между описаниями объекта и эталонов. Иерархическое сходство описаний вычисляется через сходство ХП. Модель применяемой в СИМ меры локально-пространственного подобия представима как функция $\vartheta(Z^1, Z^2) = \Upsilon[\{\rho(\lambda_i^1, \lambda_k^2)\}]$ от симметрической матрицы $\{\rho(\lambda_i^1, \lambda_k^2)\}$ значений метрики ρ для пар $\lambda_i^1 \in \Lambda^1, \lambda_k^2 \in \Lambda^2$ [3]. Инвариантность построения Λ^1, Λ^2 абстрагирует от пространственной информации и упрощает решение. На локальном уровне предпочтение отдано применению метрики $\rho(\lambda_i^1, \lambda_k^2)$, за счет чего оценка соответствия ХП определяется однозначно. Для устранения пространственных помех введем функцию $\Psi_1[\rho(\lambda_i^1, \lambda_k^2)]$ логического анализа значений $\rho(\cdot)$. Одним из вариантов есть предикат вида

$$\Psi_1[\rho(\lambda_i^1, \lambda_k^2), \delta] = \begin{cases} 0, & \rho(\lambda_i^1, \lambda_k^2) \leq \delta, \\ 1, & \text{otherwise,} \end{cases} \quad (1)$$

параметрически определяемый порогом δ . Управление изменением δ в (1) в допустимом диапазоне $[\delta_{\min}, \delta_{\max}]$ задает окрестность в пространстве ХП, в пределах которой λ_i^1, λ_k^2 считаются эквивалентными. На практике порог δ выбирают экспериментально для анализируемой базы данных. Предикат в виде отрицания $\Psi_1[.]$ в теории интеллекта назван предикатом «узнавания».

Вторым важным моментом есть способ Ψ_2 обобщения (свертывания) в общую глобальную меру локальных подобий, прошедших обработку Ψ_1 . Здесь могут быть применены корреляционные, ассоциативные, дисперсионные меры подобия, а также метрики. Одним из вариантов есть метрика Танимото

$$\Psi_2(\Lambda^1, \Lambda^2) = \rho(\Lambda^1, \Lambda^2) = \mu(\Lambda^1 \Delta \Lambda^2) / \mu(\Lambda^1 \cup \Lambda^2), \quad (2)$$

где $\mu(\cdot)$ – мощность множества; Δ, \cup – операции симметрической разности и объединения множеств; эквивалентность элементов понимается в смысле предиката (1). Конкретно СИМ в соответствии с (1)-(2) задается кортежем $\langle \Xi, \Psi_1, \Psi_2 \rangle$, включающим способ Ξ формирования ХП и процедуры Ψ_1, Ψ_2 анализа множеств ХП на локальном и глобальном уровнях обработки.

Распознавание объекта Z на основе СИМ рассмотрим как отображение $\mathfrak{R} : Z \rightarrow \{Z(j)\}$, где $\{Z(j)\}$ – множество эталонных описаний, и реализуем его путем решения задачи оптимизации

$$j^* = \arg \max_{j \in J} \vartheta(Z, Z(j)), \quad j^* \in J. \quad (3)$$

В процессе решения (3) вычисляется мощность $\mu(Z_j)$ подмножества $Z_j \subseteq Z(j)$ описания объекта Z , для которых установлена эквивалентность элементам $Z(j)$ и определяется Z_j наибольшей мощности с номером j^* . Выражения $\mu(Z_{j^*})/\mu(Z)$ или $\mu(Z_{j^*})/\mu(Z(j))$, отражающие значение доли отданных за класс $Z(j^*)$ голосов, трактуем как оценку апостериорной вероятности отнесения Z к классу j . Вычисление $\mu(Z_j)$ в теоретико-множественном представлении соответствует голосованию и сводится к построению на множестве Λ (или Z) представления $\Lambda = \cup \Lambda_j$ в виде разбиения ($\Lambda_i \cap \Lambda_j = \emptyset$) или покрытия ($\Lambda_i \cap \Lambda_j \neq \emptyset$) из элементов, которые получают метку $i \neq j$ класса j путем отображения $\Lambda \rightarrow J$ [3]. Распознавание с применением СИМ включает применение основных способов изучения структуры объектов в системном анализе, связанных как с «препарированием», так и с «проектированием» объекта на совокупность близких объектов. Схема применения СИМ при распознавании представлена на рис. 2.

Важным фактором при принятии решения (3) есть мощность $\mu(Z_{j^*})$ результирующей коалиции голосов. При выборе $\mu(Z_{j^*}) < 0,5$ требуется дополнительное обоснование [3], т.к. при этом предполагается возможность решения на основе относительно небольшого числа ХП, что не вписывается в рамки традиционных статистических решений, где принято $\mu(Z_{j^*}) \geq 0,5$. Заметим, что отождествление описаний по части одинаковых ХП или фрагментов в СИМ приводит к отождествлению целых классов изображений, где эти элементы совпадают, а другие в то же время могут различаться. Это требует контроля решений.

Этапы СИМ: 1) формирование описания в виде множества ХП или фрагментов; 2) сопоставление ХП для вычисления локальных характеристик подобия; 3) установление меры локально-пространственного подобия между описаниями объекта и эталона; 4) оптимизация на множестве описаний эталонов. Каждый из этапов имеет внутреннюю структуру.

2. Оптимизация и помехозащищенность

СИМ есть обобщение традиционных интегральных подходов на случай влияния пространственных помех, где уровень локальности при анализе задается параметрически. Принцип оптимальности выбора параметров СИМ представляется как оптимизация в соответствии с аналитически заданными статистическими критериями с учетом меры подобия ХП, разновидности Ξ, Ψ_1, Ψ_2 , а также распределений помех [2].

Сложность и многообразие представления аналитического вида критерия на примере среднего риска определяется распределениями эталона, объекта, аддитивной и пространственной помех, а также функцией сходства ХП. Одно из аналитических выражений для случая равномерного распределения эталона, функции сходства в виде суммы модуля разности и фрагментов без общих точек имеет вид

$$R(\delta_1, \delta_2) = \sum_{k=\delta_2+1}^s C_s^k p_1^k q_1^{s-k} + \sum_{k=1}^{\delta_2} C_s^k p_0^k q_0^{s-k}, \quad (4)$$

где δ_1, δ_2 – пороги (параметры оптимизации); p_0, p_1 – вероятности отклонения значения сходства в пределах порога δ_1 , которые вычисляются путем интегрирования n -мерной плотности (n – размерность ХП). Из-за сложного нелинейного вида критерия (4) оптимальные величины δ_1, δ_2 получены численно.

Еще более сложный аналитический вид имеют математические выражения для критерия среднего риска или вероятности ошибки с учетом модели пространственной помехи [2].

Анализ показывает, что при существенных отличиях характеристик помехи и эталона СИМ имеет ощутимые преимущества: численно минимум вероятности ошибки в эксперименте составил 0,16 против 0,35 для традиционного метода. В то же время при незначительных отличиях, задаваемых отклонением по яркости, преимущества практически нет: минимумы примерно одинаковы. Это объясняется более существенными интегральными свойствами традиционного метода, потому

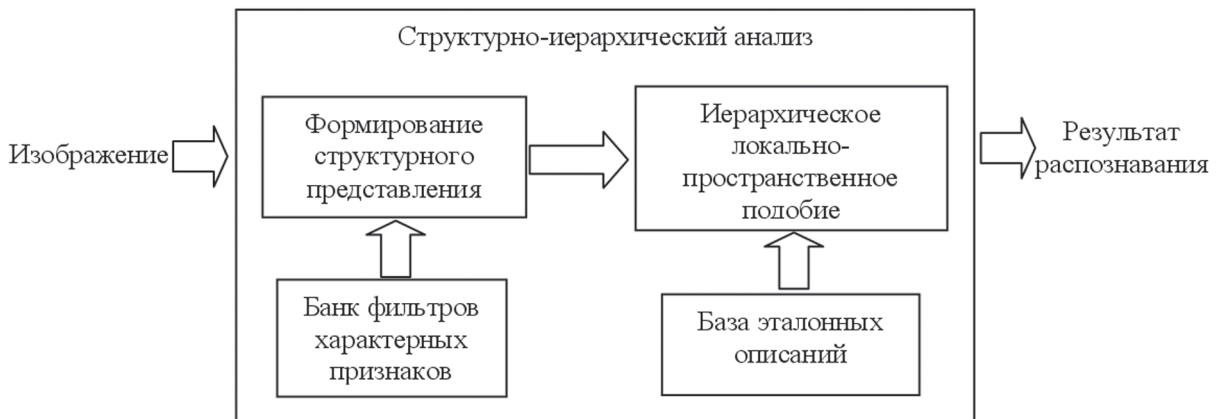


Рис. 2. Применение СИМ при распознавании

что при незначительных отличиях действие пространственных помех можно считать эквивалентным аддитивным флуктуациям. Этот факт подчеркивает тесную связь более универсального СИМ и традиционного метода, который можно считать его частным случаем. Преимущества применения СИМ усиливаются с увеличением отличий между значениями помехи и распознаваемого объекта.

3. Экспериментальные результаты

Проведено компьютерное моделирование с целью сравнительной оценки предложенных методов распознавания (СИМ) для тематической базы данных из 30 полутоновых изображений аквариумных рыб [2]. ХП получены методом SIFT, для сопоставления дескрипторов ХП использовано евклидово расстояние, а порог δ выбран как 1% от максимума метрики. По результатам экспериментов наилучшие показатели отличия в пространстве признаков и адекватной реакции на локальные искажения имеют модификации СИМ на базе метрик Хемминга, Танимото, а также мер с применением подсчета бинарных соответствий ХП.

Допустимый уровень локальных искажений для СИМ составил 75-80% от объема описания при вероятности распознавания больше 0,98. Это значительно эффективнее традиционных методов (до 30%). Время обработки с использованием СИМ меньше, чем для известных методов. Анализ бинарных отношений в мерах, основанных на многозначных соответствиях, обеспечивает лучшую чувствительность к помехам исключения и большую достоверность решений, определяемую относительным локальным оптимумом значений мер.

Выводы

Применение структурно-иерархических методов обеспечивает высокую вероятность распознавания и дает возможность осуществить оценку параметров объектов в условиях действия искажений пространственного типа. Проведенные эксперименты для систематизированных баз видеоданных в целях анализа вероятности распознавания, достоверности, помехозащищенности и быстродействия дают возможность получить обобщенные оценки для произвольных классов изображений. Решение практических задач идентификации протяженных объектов, космического мониторинга внешней среды, при контроле качества оборудования подтверждает прикладную ценность разработанных методов. Полученные результаты в плане эффективности распознавания открывают перспективу научно обоснованного решения ряда новых прикладных задач, где анализ осуществляется в присутствии ложных объектов.

Перспективы дальнейших исследований связаны с развитием адаптационных подходов в целях оптимального выбора параметров СИМ в зависимости от уровня пространственных помех.

Список литературы: 1. Шапиро, Л. Компьютерное зрение [Текст] / Л. Шапиро, Дж. Стокман.; [пер. с англ. А.А. Богуславского, под ред. С.М. Соколова]. — М.: БИНОМ. Лаборатория знаний, 2006. — 752 с. 2. Путятин, Е.П. Методы та алгоритми комп'ютерного зору [Текст]: навч. посіб. / Е.П. Путятин, В.О. Гороховатський, О.О. Матат. — Х.: ТОВ «Компанія СМІТ», 2006. — 236 с. 3. Гороховатский, В.А. Структурное распознавание изображений на основе моделей голосования признаков характерных точек [Текст] / В.А. Гороховатский, Е.П. Путятин // Реєстрація, зберігання і обробка даних. — 2008. — Т.10. — №4. — С.75-85. 4. Gorokhovatsky V.A. Image Likelihood Measures of the Basis of the Set of Conformities / V.A. Gorokhovatsky, Ye. P. Putyatın // Telecommunications and Radio Engineering. — 2009, 68 (9). — P. 763–778. 5. Шлезингер, М. Десять лекций по статистическому и структурному распознаванию [Текст] / М. Шлезингер, В. Главач. — К.: Наукова думка, 2004. — 535с. 6. Гороховатский, В.А. Применение отношений на множестве характерных признаков изображений при распознавании на основе голосования [Текст] / В.А. Гороховатский // Бионика интеллекта. — 2008. — №1(68). — С. 87–93. 7. Kinoshenko D. A Partition Metric for Clustering Features Analysis / D. Kinoshenko, V. Mashtalir, V. Shlyakhov // International Journal «Information Theories and Applications». — 2007. — V. 14. — №3. — P. 230–236.

Поступила в редакцию 3.06.2010.

УДК 004.932.2:004.93'1

Структурно-ієрархічні методи розпізнавання зображень при впливі просторових завад / В.О. Гороховатський // Біоніка інтелекту: наук.-техн. журнал. — 2010. — № 3 (74). — С. 58–61.

У роботі запропоновано та досліджено структурно-ієрархічні методи розпізнавання зображень об'єктів при дії завад просторового типу. Наведено принципи оптимізації та оцінки завадостійкості розроблених методів, приводяться результати комп'ютерного моделювання.

Л. 2. Бібліогр.: 7 найм.

UDC 004.932.2:004.93'1

Structural-hierarchical methods of image recognition under the conditions of the influence of spatial distortions / V.O. Gorokhovatskyi // Bionics of Intelligence: Sci. Mag. — 2010. — № 3 (74). — P. 58–61.

The structurally-hierarchical methods of image objects recognition under the conditions of spatial distortions influence are suggested and investigated. Principles of optimization and noise-stability estimation of the suggested methods and the results of computer modeling are shown.

Fig. 2. Ref.: 7 items.

УДК 004.93



ОЦЕНКА КАЧЕСТВА ПОДЧЕРКИВАЮЩИХ ПРЕОБРАЗОВАНИЙ ДЛЯ ЗАДАЧИ СЕГМЕНТАЦИИ ИЗОБРАЖЕНИЙ

М. В. Полякова¹, В. Н. Крылов², Н. А. Гуляева³

¹ ОНПУ, г. Одесса, Украина, marina_polyakova@rambler.ru

² ОНПУ, г. Одесса, Украина, victor_krylov@inbox.ru

³ ОНПУ, г. Одесса, Украина, gulyaevser@rambler.ru

Статья посвящена формированию обобщенного показателя качества подчеркивающего преобразования для процедуры сегментации изображений, обеспечивающего повышение эффективности обработки изображений в системах компьютерного распознавания зрительных образов. Для этого проанализированы существующие показатели качества подчеркивающего преобразования. Те из них, которые учитывают имеющуюся априорную информацию об изображении и цели обработки, выделены в обобщенный векторный показатель качества подчеркивающего преобразования.

СЕГМЕНТАЦИЯ ИЗОБРАЖЕНИЙ, ПОДЧЕРКИВАЮЩЕЕ ПРЕОБРАЗОВАНИЕ, ВЕЙВЛЕТ-ПРЕОБРАЗОВАНИЕ, ПОКАЗАТЕЛЬ КАЧЕСТВА СЕГМЕНТАЦИИ

Введение

Одной из базовых процедур в системе компьютерного распознавания зрительных образов (СКРЗО) является сегментация изображений на однородные по какому-либо признаку или набору признаков области (сегменты). Целью этой процедуры является уменьшение объема обрабатываемой при распознавании информации за счет выделения границ однородных областей. Составной частью процедуры сегментации изображений является подчеркивающее преобразование (ПП), цель которого — концентрация энергии сигнала вблизи границ сегментов. Именно ПП определяет главным образом помехоустойчивость процедуры сегментации и погрешность выделения границ однородных областей. Кроме того, при распознавании иерархических объектов, т. е. имеющих структуру «объект-подобъект», а также при анализе сцен ПП позволяет выделять объекты требуемых геометрических размеров.

На современном этапе в методах контурной и текстурной сегментации изображений при обнаружении и локализации объектов распознавания применяются различные линейные и нелинейные ПП, в частности, вейвлет-преобразование (ВП). Качество ПП оценивается, как правило, по результатам сегментации. Поэтому представляется целесообразным разработать метод оценивания качества самого ПП, позволяющий обоснованно проводить его выбор в зависимости от свойств изображений и цели обработки.

При оценке качества ПП целесообразно учитывать априорную информацию об изображении. В большинстве случаев известен уровень помех на изображении (или его можно оценить), а также максимальное значение погрешности, которая допустима при определении координат точек границ объекта на изображении. Последняя вычисляется исходя из заданного качества распознавания объектов.

По аналогии с определением показателя качества СКРЗО в целом и отдельных процедур СКРЗО, в частности, процедуры сегментации [1], введем определение показателя качества этапа сегментации.

Определение. Показателем качества этапа сегментации называется расстояние в смысле некоторой метрики между результатом обработки изображения при моделировании этапа сегментации и идеальным результатом.

Оператор системы, осуществляющей ПП, определен в [2] как состоящий из линейной фильтрации и нелинейного преобразования. Матрица значений признака сегментации преобразуется линейными фильтрами. В соответствии с целями обработки с помощью априорной информации об изображении линейные фильтры для ПП выбираются таким образом, чтобы ширина полосы пропускания фильтров уменьшалась с увеличением уровня преобразования. Линейная фильтрация используется для подчеркивания границ однородных областей изображения на разных уровнях преобразования. Если нужно обеспечить инвариантность к знаку результата ПП, после линейной фильтрации на каждом уровне выполняется нелинейное преобразование, например, вычисляется абсолютное значение. Результатом ПП является упорядоченная последовательность матриц изображения с подчеркнутыми границами областей разных размеров. К этой последовательности матриц изображения далее может быть применено композиционное преобразование. Оно объединяет результаты обработки изображения на нескольких уровнях в единую матрицу, например, с помощью применения в каждой точке изображения операторов среднеквадратического значения функции.

1. Постановка задачи

В литературе рассматриваются отдельные показатели качества ПП, что затрудняет оценку различных операторов. Для адекватного учета частных

показателей качества целесообразно ввести обобщенный векторный показатель качества ПП. Последний представляет собой многомерный вектор, зависящий от изображения значений признака сегментации $g(x, y)$, результата ПП $r(x, y)$ этого изображения и вектора параметров ПП α :

$$W = Q(\alpha, g(x, y), r(x, y)),$$

где Q — функционал, задающий значения компонентов векторного показателя качества ПП.

При разработке векторного показателя включенные в него показатели разбиваются на группы, отражающие разные аспекты качества ПП. Однако при таком подходе остается ряд трудностей:

1. Определение частных показателей, входящих в обобщенный векторный показатель. Трудности связаны с тем, что частные показатели содержат элементы субъективизма, а также не учитывают отдельные конкретные цели обработки.

2. Учет взаимосвязи между частными показателями для разных уровней обработки в методах сегментации изображений. Многоэтапность методов сегментации приводит к иерархичности ПП и, соответственно, иерархичности оценок его качества. Значения частных показателей верхнего уровня преобразования основываются на значениях показателей качества ПП на нижних уровнях.

3. Совместное использование в обобщенном векторном показателе качества ПП частных показателей, имеющих разнородный характер.

Обобщенный векторный показатель качества позволяет производить сравнение различных ПП, осуществлять выбор их параметров и добиваться наилучших значений показателей качества сегментации с учетом исходных данных метода сегментации.

Такого показателя качества, отражающего рассматриваемые исходные данные метода сегментации в полном объеме, на современном этапе не существует. В этом плане актуальным является выделение отдельных частных показателей качества ПП с последующим исследованием их взаимосвязи и введением векторного показателя качества, компоненты которого формируются из частных показателей качества ПП.

Цель данной работы заключается в формировании обобщенного показателя качества ПП для процедуры сегментации изображений, который обеспечивает повышение эффективности обработки изображений в СКРЗО.

Для достижения поставленной цели решены следующие задачи:

- анализ существующих показателей качества ПП;
- введение показателей качества ПП, учитывающих имеющуюся априорную информацию об изображении и цели обработки, а также выделение

этих показателей в обобщенный векторный показатель качества ПП.

2. Оценка погрешности определения координат граничных точек изображения

При сегментации изображений в СКРЗО необходимо учитывать количество и ценность информации на разных уровнях обработки. Меры количества и ценности информации рассматриваются в трех аспектах: структурном, статистическом и семантическом [3, 4]. Последний аспект определяется по тому эффекту, который информация, полученная в результате ПП, оказывает на функционирование СКРЗО в целом. Этот аспект не рассматривается в данной работе ввиду сложности оценки влияния результата ПП на функционирование СКРЗО в целом.

Используем при определении обобщенного векторного показателя качества ПП структурный и статистический аспекты измерения количества информации. Структурный аспект предполагает измерение геометрического параметра информационного сообщения (длина, площадь, объем и др.) путем подсчета информационных элементов. Для изображений информационный элемент представляет собой пиксель. Тогда геометрический параметр информационного сообщения рассчитывается в пикселях. В рамках структурного аспекта оценки качества ПП вычисляется погрешность определения координат граничных точек изображения.

Последнюю характеристику оценим с помощью эффективной протяженности импульса, полученного в результате ПП изображения перепада интенсивности. Эту эффективную протяженность импульса определим из условия двукратного уменьшения амплитуды сигнала изображения в пространстве ПП. В частности, если перепад интенсивности ориентирован на изображении вертикально, то эффективную протяженность импульса

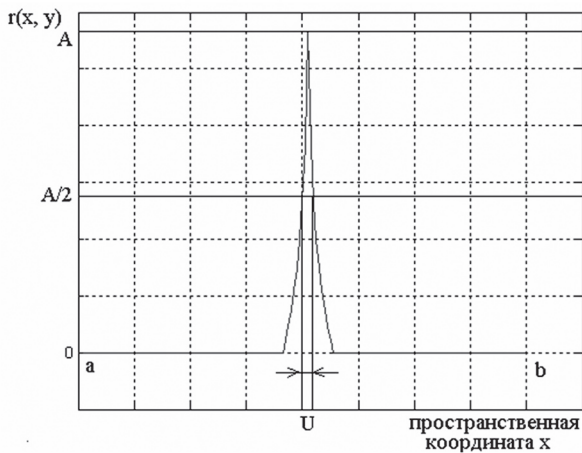
L определяем из условия $\frac{r(x, y)}{\max_{x \in [a, b]} r(x, y)} = 0,5$, где

$r(x, y)$ — результат ПП перепада интенсивности (рис. 1, а).

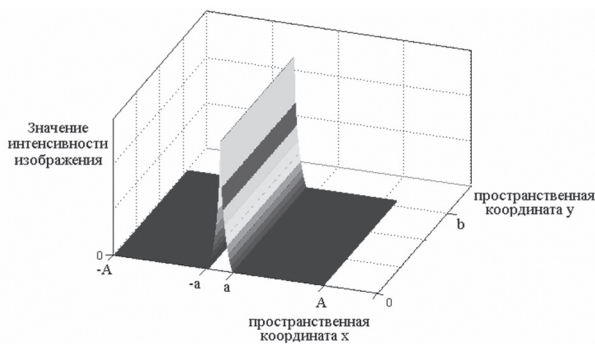
При оценке эффективной протяженности импульса, выбран уровень двукратного уменьшения амплитуды сигнала изображения в пространстве ПП, так как в случае оптимального обнаружения перепада интенсивности изображения как детерминированного сигнала по критерию максимума апостериорной вероятности в качестве порогового значения используется половина энергии сигнала изображения в пространстве ПП [5]. Последняя связана с амплитудой импульса, полученного в результате ПП изображения перепада интенсивности, неравенством Коши-Буняковского [6].

3. Оценка помехоустойчивости ПП

Статистический аспект информационных мер учитывает конкретные условия применения СКР-ЗО, а именно, помеховую ситуацию на изображении. Для ПП изображения, искаженного наличием шума, целесообразно применять линейную фильтрацию. Тогда в рамках статистического аспекта информационных мер помехоустойчивость результата ПП изображения оценим с помощью выигрыша фильтра, используемого в составе ПП. Для определения этого показателя заметим, что частотный коэффициент передачи $K(j\omega)$ линейного фильтра выбирается так, что значения $|K(j\omega)|$ велики в области частот, где сконцентрирована основная доля энергии изображения, и малы там, где велик спектр мощности шума. Тогда, подав на вход такого преобразования сумму сигнала и шума, на выходе получаем увеличение относительной доли сигнала изображения.



а



б

Рис. 1: а — определение эффективной протяженности импульса; здесь $A = \max_{x \in [a, b]} r(x, y)$; б — модуль результата ПП изображения вертикального перепада интенсивности

Пусть входной сигнал представляет собой значение интенсивности $I(x, y_i)$ i -го столбца изображения и является суммой полезного сигнала $I_0(x, y_i)$ и шума $N(x, y_i)$:

$$I(x, y_i) = I_0(x, y_i) + N(x, y_i).$$

Предполагается, что незашумленное изображение и шум некоррелированы в том смысле, что

$$\langle I_0(x, y_i), N(x, y_i) \rangle = \int_a^b I_0(x, y_i) N(x, y_i) dx = 0. \quad (1)$$

Интенсивность колебаний на выходе фильтра характеризуют средним квадратом входного изображения, который в силу равенства (1) является суммой средних квадратов незашумленного изображения и шума:

$$\langle I^2(x, y_i) \rangle = \langle I_0^2(x, y_i) \rangle + \langle N^2(x, y_i) \rangle = \langle I_0^2(x, y_i) \rangle + \sigma_{\text{вх}}^2,$$

где $\sigma_{\text{вх}}^2$ — дисперсия шума на входе линейного фильтра.

Относительный уровень сигнала изображения на входе фильтра описывается отношением сигнал/шум $Q_{\text{вх}}$:

$$Q_{\text{вх}} = \langle I_0^2(x, y_i) \rangle / \sigma_{\text{вх}}^2.$$

Так как линейная фильтрация подчиняется принципу суперпозиции, незашумленное изображение и шум обрабатываются независимо при помощи этой операции. Результирующее изображение

$$I_{\text{вых}}(x, y_i) = I_{0\text{вых}}(x, y_i) + N_{\text{вых}}(x, y_i)$$

имеет средний квадрат

$$\begin{aligned} \langle I_{\text{вых}}^2(x, y_i) \rangle &= \langle I_{0\text{вых}}^2(x, y_i) \rangle + \langle N_{\text{вых}}^2(x, y_i) \rangle = \\ &= \langle I_{\text{вх}}^2(x, y_i) \rangle + \sigma_{\text{вых}}^2, \end{aligned}$$

где $\sigma_{\text{вых}}^2$ — дисперсия шума на выходе фильтра.

Отношение сигнал/шум на выходе фильтра определяется формулой

$$Q_{\text{вых}} = \langle I_{0\text{вых}}^2(x, y_i) \rangle / \sigma_{\text{вых}}^2.$$

Выигрыш фильтра называется величина $M = Q_{\text{вых}} / Q_{\text{вх}}$. Если $M > 1$, то фильтрация зашумленного изображения приводит к повышению относительного уровня полезного сигнала изображения на выходе фильтра. Поэтому для оценки помехоустойчивости ПП предлагается использовать выигрыш фильтра преобразования, характеризующий в рамках статистического аспекта изменение количества информации на изображении после ПП.

4. Энтропийный показатель изображения

В качестве характеристики погрешности определения координат точек перепада значений признака сегментации также использовался энтропийный показатель, который определяется следующим образом [7].

Изображение рассматривается как плотность вероятности случайной величины в пространстве значений интенсивности. Каждая точка из области определения изображения Ω характеризуется зна-

чением интенсивности или соответственно (после нормирования) вероятностью освещенности. Тогда можно определить распределение вероятностей с плотностью

$$p(x,y) = \frac{I(x,y)}{\sum_{(x,y) \in \Omega} I(x,y)}. \quad (2)$$

Энтропийный показатель изображения как распределения вероятностей с плотностью вероятности $p(x,y)$ определен в [7] как

$$S(p(x,y)) = - \sum_{(x,y) \in \Omega} p(x,y) \ln p(x,y). \quad (3)$$

Эта величина интерпретируется как среднее содержание информации в распределении $p(x,y)$. Энтропийный показатель (3) инвариантен к умножению изображения на константу вследствие нормирования (2), однако значение этого показателя изменяется при изменении значений интенсивности изображения на аддитивную константу. Как характеристика погрешности определения координат точек перепада значений интенсивности изображения с помощью ПП энтропийный показатель монотонно изменяется при монотонном изменении параметра преобразования.

Этот результат справедлив для разных видов ПП (обычного ВП, ВП с использованием лифтинга и вейвлетов, полученных путем преобразований графика степенной функции [8, 9], преобразования с обобщенными вейвлет-функциями [8] и др.). Столбец изображения модуля результата ПП тестового изображения перепада представляет собой треугольный импульс (рис. 1, б), ширина которого зависит от параметра преобразования (масштаба для обычного ВП, показателя степени для репагулярного ВП). Энтропийный показатель изображения, вычисляемый по формуле (3), монотонно возрастает при возрастании ширины импульса и характеризует погрешность определения координат точек перепада значений признака сегментации.

5. Экспериментальные исследования

Для экспериментальных исследований использовалось тестовое полутоновое изображение, состоящее из 256×256 элементов. В центре этого изображения находился вертикально ориентированный перепад интенсивности. На изображение перепада был наложен независимый гауссовский шум со среднеквадратическим отклонением $\sigma_{\text{вх}}$. Отношение сигнал/шум $q_{\text{вх}}$ по амплитуде определялось как $q_{\text{вх}} = h^2 / \sigma_{\text{вх}}^2$, где h — высота перепада интенсивности.

Для вычисления эффективной протяженности импульса, полученного в результате ПП перепада интенсивности, строка незашумленного тестового изображения подвергалась свертке с импульсной

характеристикой фильтра ПП. Далее определялось максимальное значение A модуля результата преобразования. Число пикселей, результат преобразования интенсивности которых превысил по модулю $A/2$, принималось в качестве оценки эффективной протяженности импульса ПП.

Чтобы оценить выигрыш фильтра, входящего в состав ПП, сгенерированное зашумленное тестовое изображение подвергалось свертке с этим фильтром. Из полученного изображения вычитался результат свертки незашумленного тестового изображения с фильтром ПП. Вычислялось среднеквадратичное отклонение $\sigma_{\text{вых}}$ значений интенсивности изображения, полученного после вычитания. Отношение сигнал/шум $q_{\text{вых}}$ по амплитуде определялось как $q_{\text{вых}} = S_{\text{вых}}^2 / \sigma_{\text{вых}}^2$, где $S_{\text{вых}}$ — амплитуда незашумленного тестового изображения после свертки с фильтром ПП.

Выигрыш M фильтра ПП вычислялся по формуле $M = q_{\text{вых}} / q_{\text{вх}}$. Оценивание энтропийного показателя производилось по формуле

$$S(p(x,y)) = - \sum_{x=1}^M \sum_{y=1}^N \tilde{p}(x,y) \ln \tilde{p}(x,y),$$

где $\tilde{p}(x,y) = \begin{cases} p(x,y), & p(x,y) > 0 \\ 1, & p(x,y) = 0 \end{cases}$; $p(x,y)$ — результат ПП изображения, нормированный по формуле (2).

При проведении экспериментальных исследований использовались фильтры модифицированного гиперболического ВП [8], репагулярного ВП [8], ВП с использованием лифтинга и вейвлетов, полученных путем преобразований графика степенной функции [9], преобразования с обобщенными вейвлет-функциями [8], а также фильтры, применявшиеся в методе Канни сегментации изображений [10]. Последние получены путем дискретизации свертки гауссиана

$$G_{\sigma}(x) = \frac{1}{2\pi\sigma^2} \exp\left(-\frac{x^2}{2\sigma^2}\right)$$

с его производной, где σ — параметр гауссиана.

Фильтры модифицированного гиперболического ВП получают в результате дискретизации функций

$$\psi_1\left(\frac{x}{s}\right) = \frac{1}{\pi\alpha x \sqrt{s}} G_1\left(\frac{x}{s}\right),$$

где $G_1(x) = \begin{cases} \theta(x-\varepsilon) - \theta(x-\gamma), & |x| > \varepsilon, \\ x^2/\varepsilon, & |x| \leq \varepsilon; \end{cases}$; $\theta(x)$ — функция Хевисайда; $\varepsilon, \gamma, \alpha$ — константы; s — параметр масштаба.

Фильтры репагулярного ВП являются результатом дискретизации при $x \geq 1$, $x \leq 1$ функций

Фильтры репагулярного ВП являются результатом дискретизации при $x \geq 1$, $x \leq 1$ функций

$$\psi(x, a) = |x|^{-a} \operatorname{sgn}(x) w(x, a),$$

где $w(x, a)$ — функция окна; $w(x, a) = \begin{cases} 1, & |x|^{-a} \geq \varepsilon \\ 0, & |x|^{-a} < \varepsilon \end{cases}$;

ε — фиксировано. Эти фильтры рассчитывались по формуле

$$\{g_a(n)\}_{n=0}^{2N_a} = \left\{ -\frac{1}{N_a^a}, \dots, -\frac{1}{4^a}, -\frac{1}{3^a}, -\frac{1}{2^a}, -1, \right. \\ \left. 1, \frac{1}{2^a}, \frac{1}{3^a}, \frac{1}{4^a}, \dots, \frac{1}{N_a^a} \right\},$$

где a — параметр репагулярного ВП; $2N_a$ — количество коэффициентов фильтра.

Фильтры преобразования с обобщенными вейвлет-функциями были получены путем дискретизации функций

$$\Psi_2\left(\frac{x}{s}\right) = \tilde{\Psi}_2\left(\frac{x}{s}\right) \Phi(x),$$

где $\tilde{\Psi}_2\left(\frac{x}{s}\right) = \frac{1}{\left(\frac{x}{s}\right)^2 + \sqrt{\frac{x}{s}}}$; $\Phi(x)$ — регуляризирующая

функция, определенная в [11]; s — параметр масштаба.

ВП с использованием лифтинга и вейвлетов, полученных путем преобразований графика степенной функции, предполагает различные подходы для построения анализирующих вейвлетов в случае, когда параметр масштаба $s > 1$, $s < 1$. В первом случае для построения вейвлет-функций применялся лифтинг. Фильтры преобразования в этом случае имеют коэффициенты:

$$g_s = \left\{ -\frac{1}{s \cdot k}, \dots, -\frac{1}{s \cdot 3}, -\frac{1}{s \cdot 2}, -1, 1, \frac{1}{s \cdot 2}, \frac{1}{s \cdot 3}, \dots, \frac{1}{s \cdot k} \right\}.$$

Во втором случае построение вейвлет-функций осуществлялось путем геометрических преобразований графика степенной функции. Коэффициенты соответствующих фильтров имеют вид:

$$g_s^{new} = \left\{ -\frac{1}{s \cdot k + 1}, \dots, -\frac{1}{s \cdot 2 + 1}, -\frac{1}{s + 1}, -1, \right. \\ \left. 1, \frac{1}{s + 1}, \frac{1}{s \cdot 2 + 1}, \dots, \frac{1}{s \cdot k + 1} \right\}.$$

Заметим, что в данной работе проблема сегментации формируется как проблема обнаружения детерминированных сигналов на фоне помех. Поэтому при разработке метода оценки качества ПП учитывались принципы оптимальности обнаружения.

Фильтры перечисленных преобразований сравнивались с фильтрами преобразования Хаара [12]. Последние получают путем дискретизации функций

$$h_\gamma(x) = \begin{cases} \frac{1}{\sqrt{2\gamma}}, & -\gamma \leq x < 0 \\ -\frac{1}{\sqrt{2\gamma}}, & 0 \leq x < \gamma \\ 0, & x \in \left[-\frac{1}{2}, \frac{1}{2}\right] \setminus [-\gamma, \gamma] \end{cases},$$

где $\gamma \leq \frac{1}{2}$ — размер носителя функции Хаара. Эта функция представляет собой импульсную характеристику фильтра, согласованного с моделью перепада значений признака сегментации. Фильтры преобразования в этом случае имеют вид:

$$h_n(x) = (\underbrace{0, 0, \dots, 0}_n, \underbrace{1, 1, \dots, 1}_n),$$

где n — количество нулей и количество единиц среди коэффициентов фильтра. Случай $n = 1$ соответствует дифференциатору, $n = 128$ — согласованному фильтру для исследуемого тестового изображения.

6. Анализ полученных результатов

Замечено, что с ростом параметра ПП, а именно, масштаба обычного ВП или параметра репагулярного ВП, ширина треугольного импульса на изображении с рис. 1, б увеличивается, что обусловлено возрастанием помехоустойчивости преобразования. Вместе с этим, как было показано, увеличивается и энтропийный показатель такого изображения. Последнее вызвано увеличением погрешности определения координат точек на перепаде значений признака сегментации, для характеристики которой и применяют энтропийный показатель.

Охарактеризуем поведение показателей качества ПП в зависимости от свойств реализующих его фильтров. Наименьшую помехоустойчивость имеет дифференциатор, которому соответствует фильтр с коэффициентами $(-1, 1)$. Применение дифференциатора для подчеркивания границ однородных областей изображения дает наименьшую погрешность определения координат точек границ. Самую высокую помехоустойчивость получаем, используя в составе ПП фильтр, согласованный с моделью перепада значений признака сегментации. Однако погрешность определения координат точек на границах однородных областей у такого ПП также наиболее высокая. Поэтому выигрыш фильтра и эффективная протяженность импульса ПП имеют минимальное значение для дифференциатора, возрастают по мере роста эффективной ширины носителя импульсной характеристики фильтра преобразования и достигают максимального значения для согласованного фильтра. Аналогично ведет себя и энтропийный показатель изображения. Однако его значение зависит не только от эффективной ширины носителя

импульсной характеристики фильтра преобразования и эффективной протяженности импульса на выходе ПП, но и от формы последнего. Чем ближе по форме импульс на выходе ПП к треугольнику, тем ниже энтропийный показатель. Для импульса в форме треугольника, который получается при использовании фильтра, согласованного с перепадом значений признака сегментации, энтропийный показатель принимает самое низкое значение при фиксированном значении выигрыша фильтра преобразования. При сужении импульса, который по форме похож на график функции $y = |x|^\alpha$ при $\alpha \in (0, 1), \alpha \rightarrow 0$, энтропийный показатель возрастает.

На рис. 2 представлены кривые «эффективная протяженность импульса — выигрыш фильтра» для каждого вида ПП. При построении кривых варьировалось значение отношения сигнал/шум тестового изображения (оно выбиралось из множества $\{1, 2, 5, 10, 20, 50, 100\}$), а также изменялся параметр ПП в соответствии с табл. 1.

Таблица 1

Значения варьируемых параметров ПП

Вид ПП	Варьируемый параметр	Множество значений параметра
Регулярированное ВП	Показатель степени a	$\{0,0625; 0,125; 0,18; 0,25; 0,4; 0,5; 0,6; 0,65; 0,7; 0,8; 0,9; 1\}$
Модифицированное гиперболическое ВП	Параметр масштаба s	$\{0,8; 1; 1,5; 2; 2,5; 3; 4; 4,5; 4,7; 5,1; 5,3; 5,5; 5,7; 5,85\}$
Преобразование с вейвлетами, построенными путем преобразования графика функции и лифтинга	Параметр s	$\{0,0001; 0,001; 0,005; 0,01; 0,02; 0,03; 0,05; 0,1; 0,25; 0,5; 1; 2; 3\}$
Преобразование с обобщенными вейвлет-функциями	Параметр масштаба s	$\{0,05; 0,15; 0,25; 0,5; 0,75; 1; 1,25; 1,5; 2; 4; 8; 12; 16\}$
Преобразование из метода Канни	Параметр масштаба σ	$\{0,2; 0,6; 0,8; 1; 1,5; 2; 3; 4; 6; 7; 8; 10; 12; 16\}$
Преобразование Хаара	Размер носителя вейвлета Хаара	$\{2; 20; 40; 60; 80; 100; 120; 140; 160; 180; 200; 220; 240; 256\}$

Аналогичным образом строились кривые «энтропийный показатель — выигрыш фильтра» (рис. 3). Рассматриваемые характеристики только косвенно учитывают уровень ложных совпадений, так как перечисленные преобразования в большинстве случаев сглаживают повторные обнаружения.

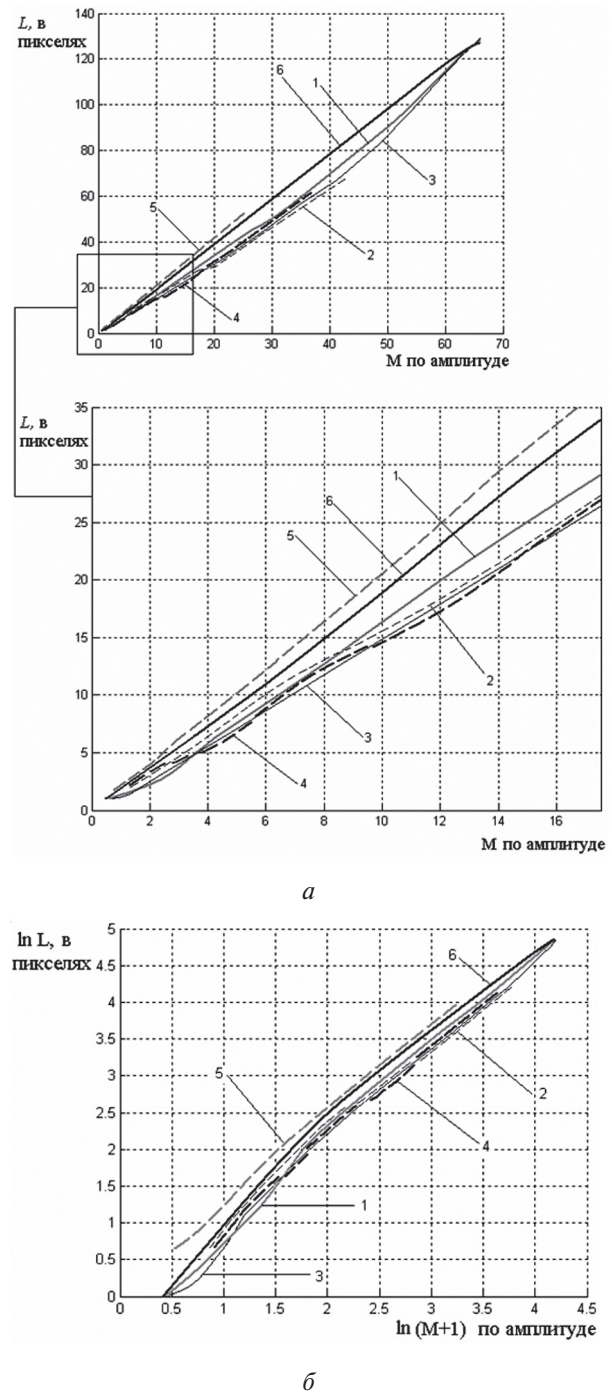
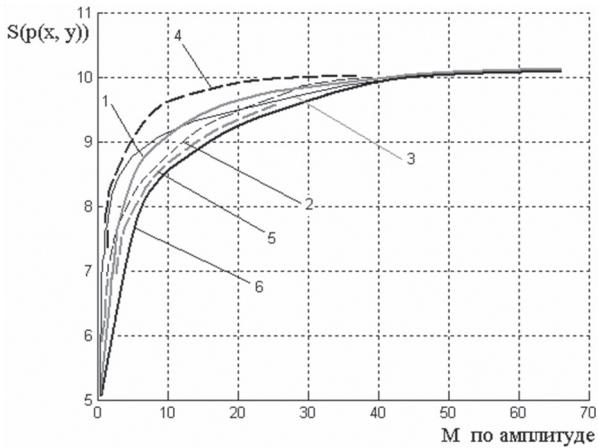


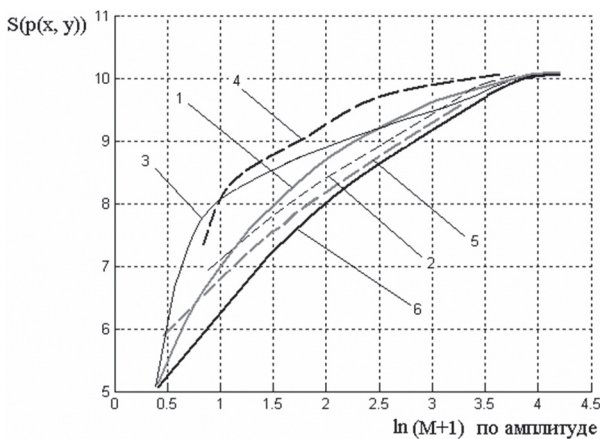
Рис. 2. Зависимость эффективной протяженности импульса от выигрыша фильтра (а — в линейном и б — в логарифмическом масштабе): 1 — регулярированное ВП; 2 — модифицированное гиперболическое ВП; 3 — преобразование с вейвлетами, построенными путем преобразования графика степенной функции и лифтинга; 4 — преобразование с обобщенными вейвлет-функциями; 5 — преобразование из метода Канни; 6 — преобразование Хаара

Анализируя полученные результаты, заметим, что левая нижняя точка на графиках соответствует дифференциатору (самая низкая помехоустойчивость и самая низкая погрешность определения координат точек на границах однородных областей).

Правая верхняя точка соответствует согласованному фильтру для исследуемого тестового изображения. Параметры репагулярного ВП, ВП с использованием лифтинга и преобразования Хаара можно подобрать таким образом, чтобы эти преобразования включали дифференциатор и согласованный фильтр. Поэтому кривые, соответствующие этим преобразованиям, пересекаются в крайней верхней и крайней нижней точках динамического диапазона значений показателей. Для остальных преобразований нельзя подобрать значения параметров таким образом, чтобы эти преобразования включали дифференциатор и согласованный фильтр. Тогда кривые, соответствующие этим преобразованиям, представляют собой отрезки, приблизительно параллельные фрагментам кривых, покрывающих весь динамический диапазон значений показателей для исследуемого тестового изображения.



а



б

Рис. 3. Зависимость энтропийного показателя изображения от выигрыша фильтра (а — в линейном и б — в логарифмическом масштабе): 1 — репагулярное ВП; 2 — модифицированное гиперболическое ВП; 3 — преобразование с вейвлетами, построенными путем преобразования графика степенной функции и лифтинга; 4 — преобразование с обобщенными вейвлет-функциями; 5 — преобразование из метода Канни; 6 — преобразование Хаара

На координатной плоскости «эффективная протяженность импульса — выигрыш фильтра» предпочтительнее считаются значения, расположенные ниже (им соответствует низкая погрешность определения координат точек границ однородных областей) и правее (с высокой помехоустойчивостью). На координатной плоскости «энтропийный показатель — выигрыш фильтра» низкой погрешности определения координат точек на границах однородных областей и высокой помехоустойчивости соответствуют точки, расположенные выше и правее. Тогда при фиксированном значении выигрыша фильтра наибольшую эффективную протяженность импульса ПП дают фильтры преобразования Хаара. Помехоустойчивость у преобразования Хаара выше, чем у остальных исследуемых ПП, т. к. это преобразование реализует принцип оптимальной фильтрации [6].

Наихудшую помехоустойчивость показали вейвлеты, построенные с использованием лифтинга и преобразований графика степенной функции. Однако соответствующее ПП позволяет получить наименьшую эффективную протяженность импульса результата преобразования.

Репагулярное ВП, модифицированное гиперболическое ВП, преобразование с обобщенными вейвлет-функциями занимают промежуточное положение между преобразованием с вейвлетами, построенными путем преобразования графика степенной функции и лифтинга, и преобразованием Хаара. Они имеют свою область применения в зависимости от помеховой ситуации и цели обработки.

Для исследуемых преобразований самого высокого значения энтропийного показателя при выигрыше фильтра более 2 по амплитуде достигает ПП с обобщенными вейвлет-функциями. На интервале менее 2 по амплитуде самые высокие значения энтропийного показателя получены для вейвлетов, построенных с использованием лифтинга и преобразований графика степенной функции. Похожее соотношение наблюдалось также для зависимостей эффективной протяженности импульса преобразования от выигрыша фильтра.

Заметим также, что для преобразования, использующего гауссовский фильтр, зависимость энтропийного показателя изображения от выигрыша фильтра преобразования практически совпадает с зависимостью, полученной для преобразования Хаара. Это соотношение отмечено на диапазоне значений выигрыша фильтра, вычисленных для преобразования, использующего гауссовский фильтр. Незначительное улучшение помехоустойчивости гауссовского фильтра по сравнению с оптимальным фильтром, используемым в преобразовании Хаара, объясняется погрешностью эксперимента и близостью последнего фильтра к дифференциатору при малых n .

При возрастании параметров s, n, σ ПП из табл. 1 и убывании параметра a репагулярного ВП возрастает помехоустойчивость соответствующих преобразований и растет погрешность определения координат точек границ однородных областей. На графиках эти свойства представлены возрастанием выигрыша фильтра преобразования, эффективной протяженности импульса и энтропийного показателя изображения.

Выводы

Таким образом в данной работе введен обобщенный векторный показатель качества ПП. Этот показатель позволяет обоснованно выбирать вид и параметры ВП при сегментации изображений иерархических объектов и анализе сцен в зависимости от помеховой ситуации, цели обработки и свойств объекта распознавания.

Приведены также результаты эксперимента и сравнительный анализ в качестве ПП модифицированного гиперболического ВП, репагулярного ВП, ВП с использованием лифтинга и вейвлетов, полученных путем преобразований графика степенной функции, преобразования с обобщенными вейвлет-функциями, а также преобразования из метода Канни сегментации изображений. Эти данные рекомендуются для использования при проектировании СКРЗО для выбора ПП при контурной, текстурной сегментации изображений и локализации объекта распознавания.

Список литературы: 1. *Абакумов, В.Г.* Повышение эффективности обработки образной информации в автоматизированных системах [Текст] / В.Г. Абакумов, В.Н. Крылов, С.Г. Антошук // *Электроника и связь: Темат. вып. «Проблемы электроники»*. — 2005. — Ч. 1. — С. 100–105. 2. *Полякова, М.В.* Иерархическое сигнально-семантическое преобразование в пространстве обобщенных функций для задачи сегментации изображений [Текст] / М.В. Полякова, В.Н. Крылов // *Труды Одес. политех. ун-та*. — Одесса, 2006. — Вып. 2 (26). — С. 161–167. 3. *Колмогоров, А.Н.* Алгоритм, информация, сложность [Текст] / А.Н. Колмогоров. — М.: Знание, 1991. — 48 с. 4. *Шилейко, А.В.* Введение в информационную теорию систем [Текст] / А.В. Шилейко, В.Ф. Кочетов, Ф.Ф. Химушин; Под редакцией А.В.Шилейко. — М.: Радио и связь, 1985. — 280 с. 5. *Крылов, В.Н.* Вторичные преобразователи сигналов изображений [Текст] / В.Н. Крылов, М.В. Максимов. — Одесса: Астропринт, 1997. — 176 с. 6. *Баскаков, С.И.* Радиотехнические цепи и сигналы [Текст] / С.И. Баскаков. — М.: Высш. шк., 1988. — 488 с. 7. *Sporring J.* The entropy of scale-space // *Proc. 13th Int. Conf. Pattern Recognition (ICPR '96)* (August 3—11, 1996). — Vienna, Austria, 1996. — V. 1. — P. 900–904. 8. *Polyakova M.V., Krylov V.N.* Classification of methods of the signal semantic wavelet transform for image contour segmentation // *Computing*. — 2008. — V. 7, № 1. — P. 51–57. 9. *Polyakova M.V., Krylov V.N.* Edge detection based on wavelets constructed by transforms of the graph of

power function // *Сучасні проблеми радіоелектроніки, телекомунікацій, комп'ютерної інтеграції: Матеріали Міжнародної наук. техн. конф. TCSET 2010* (23–27 лютого 2010 р.). — Львів-Славськ: Видавництво Національного ун-ту Львівська політехніка, 2010. — С. 270 — 273. 10. *Canny J.E.* A computational approach to edge detection // *IEEE Trans. on Pattern Analysis and Machine Intelligence*. — 1986. — № 8. — P. 679–698. 11. *Гельфанд, И.М.* Обобщенные функции и действия над ними [Текст] / И.М. Гельфанд, Г.Е. Шилев. — Вып. 1. — М.: Государственное изд-во физико-математической литературы, 1959. — 470 с. 12. *Akansu A.N., Haddad R.A.* Multiresolution signal decomposition. — New York: Academic Press, 2000. — 500 p.

Поступила в редколлегию 7.06.2010.

УДК 004.93

Оцінка якості підкреслюючих перетворень для задачі сегментації зображень / М. В. Полякова, В. М. Крылов, Н. А. Гуляева // *Біоніка інтелекту: наук.-техн. журнал*. — 2010. — № 3 (74). — С. 62–69.

Введено узагальнений векторний показник якості підкреслюючих перетворень. Цей показник дозволяє обґрунтовано обирати вигляд і параметри вейвлет-перетворення при сегментації зображень ієрархічних об'єктів і аналізі сцен залежно від перешкоджуючої ситуації, мети обробки і властивостей об'єкту розпізнавання. Приведені також результати експерименту і порівняльний аналіз у якості підкреслюючого перетворення модифікованого гіперболического вейвлет-перетворення, репагулярного вейвлет-перетворення, вейвлет-перетворення з використанням ліфтинга і вейвлетів, отриманих шляхом перетворень графіка степеневі функції, перетворення з узагальненими вейвлет-функціями, а також перетворення з методу Канні сегментації зображень. Ці дані рекомендуються для використання при проектуванні системи комп'ютерного розпізнавання зорових образів для вибору підкреслюючих перетворень при контурній, текстурній сегментації зображень і локалізації об'єкту розпізнавання.

Лл. 3. Бібліогр.: 12 найм.

UDC 004.93

Estimation of quality of underlining transforms for the image segmentation problem / M. V. Polyakova, V. N. Krylov, N. A. Gulyaeva // *Bionics of Intelligence: Sci. Mag.* — 2010. — № 3 (74). — P. 62–69.

The generalized vector index of quality of underlining transforms is offered. This index allows grounded to choose a type and parameters of wavelet transform during the segmentation of images of hierarchical objects and an analysis of the scene depending on a noise level, processing purpose and properties of the recognition object. The results of experiment and comparative analysis are presented for modified hyperbolic wavelet transform, repagular wavelet transform, wavelet transform with the use of lifting and wavelets got by transforms of the graph of power function, transform with the distributions, and also transform from the method of Canny of image segmentation. This information is recommended for using for synthesis of image processing systems for the choice of underlining transform during edge detection and texture segmentation of images and localization of object of recognition.

Fig. 2. Ref.: 12 items.

УДК 004.93



КЛАСТЕРНЫЙ АНАЛИЗ С ИСПОЛЬЗОВАНИЕМ ГИБРИДНОЙ МОДЕЛИ НА ОСНОВЕ ИСКУССТВЕННОЙ ИММУННОЙ СЕТИ

С.А. Зайцев¹, С.А. Субботин²¹ Запорожский национальный технический университет, г. Запорожье, zaitsev.serge@gmail.com² Запорожский национальный технический университет, г. Запорожье, subbotin@zntu.edu.ua

Проведен сравнительный анализ модели искусственной иммунной сети aiNet и модели расширяющегося нейронного газа в задаче кластер-анализа. Предложена гибридная модель искусственной иммунной сети на основе проанализированных моделей, позволяющая повысить скорость обучения сети. Проведены эксперименты, подтверждающие эффективность предложенной модели.

КЛАСТЕР-АНАЛИЗ, ИСКУССТВЕННАЯ ИММУННАЯ СЕТЬ, РАСШИРЯЮЩИЙСЯ НЕЙРОННЫЙ ГАЗ

Введение

Кластерный анализ является инструментом, позволяющим из выборки данных выделить скопления подобных наблюдений и обобщить их, заменив каждую группу компактно расположенных экземпляров эталоном кластера. Кластерный анализ находит широкое применение при построении диагностических и распознающих моделей, поскольку, с одной стороны, позволяет заменить обучающую выборку большой размерности множеством эталонов меньшей размерности, а, с другой стороны, результаты кластерного анализа позволяют представить в модели топологию обучающих данных (например, для нейро-нечетких моделей кластерный анализ позволяет выделить обобщенные правила, а также сформировать нечеткие термы).

Традиционно используемые методы кластер-анализа [1], как правило, сильно зависят от начального разбиения выборки (начальных значений координат центров кластеров) и количество кластеров должно быть известно заранее.

В качестве инструмента для решения задач кластер-анализа возможно использование искусственной иммунной сети (ИИС). Наиболее распространенной моделью ИИС является aiNet [2], которая автоматически определяет количество кластеров и выполняет своего рода компрессию при избыточном объеме входных данных. Тем не менее, для модели aiNet характерно слишком большое число управляющих параметров и высокие требования к вычислительным ресурсам [3]. Поэтому возникает необходимость разработки метода кластер-анализа на основе модели aiNet, свободного от отмеченных недостатков.

Применение метода расширяющегося нейронного газа [4] в кластерном анализе позволяет избежать этих недостатков. Данный метод не зависит от расположения начальных точек поиска, поскольку в его основе лежит вероятностный подход. Однако в методе расширяющегося нейронного газа отсутствует возможность динамического управления

параметрами, влияющими на скорость сходимости метода, что приводит к многократному перезапуску метода с целью нахождения оптимальной комбинации параметров и усложняет его использование в автоматизированных системах.

Перспективным направлением развития модели aiNet в задаче кластер-анализа является ее совмещение с методом расширяющегося нейронного газа. В частности, в гибридной модели целесообразным будет применение используемого в aiNet механизма адаптации к поступающим на вход сети данным и взятого из метода расширяющегося газа принципа определения области сети, которая, собственно, подлежит изменению в процессе адаптации.

Цель работы заключается в разработке метода кластер-анализа на основе гибридной модели ИИС, объединяющей высокую адаптивность модели aiNet и критерий отбора адаптируемой части сети метода расширяющегося нейронного газа, что призвано повысить скорость сходимости метода обучения ИИС и, соответственно, кластер-анализа.

1. Постановка задачи

В общем виде задача кластер-анализа выглядит следующим образом. Пусть X – множество объектов, Y – множество меток кластеров. Задана метрика – функция расстояния между двумя объектами $p(x, x')$, $x \in X$, $x' \in X$. Существует конечная обучающая выборка объектов $X^m = \{x_1, x_2, \dots, x_m\}$, $X^m \in X$. Необходимо разбить выборку X^m на непересекающиеся подмножества (кластеры) таким образом, чтобы каждый кластер состоял из объектов, близких по метрике p , а объекты разных кластеров существенно отличались. При этом каждому объекту $x \in X^m$ приписывается номер кластера $y \in Y$. Следует отметить, что рассматриваемые ниже методы лишь выделяют области кластеров, но не назначают им метки. Задача назначения меток кластерам является тривиальной и не рассматривается в данной статье.

2. Искусственная иммунная сеть

Ключевой идеей, лежащей в основе парадигмы иммунных сетей [5], является тот факт, что иммунная память формируется по причине взаимодействия иммунных клеток, которое происходит даже при отсутствии инородных антигенов. Воздействие одной иммунной клетки на другую может быть либо стимулирующим, либо подавляющим, при этом компоненты ИИС могут распознавать друг друга, а также быть распознанными другими компонентами.

ИИС aiNet можно представить в виде взвешенного графа (не обязательно полносвязного), узлы которого называют клетками, а ребра определяют связи между клетками. Каждому ребру графа соответствует некоторое числовое значение, называемое весом или силой связи. В терминах ИИС экземпляры обучающей выборки называются антигенами. Большинство моделей ИИС, в том числе и aiNet, не делают различий между В-лимфоцитами и антителами, а потому под антителом будем подразумевать узел иммунной сети.

Работа ИИС построена на постоянной адаптации структуры сети к поступающим на вход данным [6]. Принимая входное распределение как набор антигенов, сеть формирует иммунный ответ и создает множество антител, подвергая их мутации и отбору наилучших для дальнейшего использования. После обучения таким образом модель aiNet способна классифицировать поступающие на вход новые экземпляры, отсеивая по заданному порогу связи между антителами и выделяя тем самым группы антител, которые и представляют собой кластеры.

Метод обучения сети формирует иммунную память из множества клеток, способных распознавать экземпляры из обучающей выборки, и множества связей, представляющих структурную организацию данных. В конечном итоге сеть стремится снизить количество клеток в памяти и сделать их более специфичными, поэтому в таком случае сеть выполняет своего рода компрессию хранящейся в памяти информации.

В процессе обучения сети на вход модели поступают антигены, и сеть реагирует на них, вырабатывая иммунный ответ с помощью механизма клональной селекции. Для регуляции размера популяции клонов применяют подходы клональной супрессии и сетевой супрессии. Поскольку в качестве реакции на один и тот же входной вектор генерируется множество клонов, то часть из них будет реагировать друг на друга как на антигенные образцы. Для устранения таких клонов и обеспечения толерантности служит механизм клональной супрессии. В то же время, сетевая супрессия служит для подавления взаимного влияния клонов, произведенных в результате реакции различных антител на антигены.

Метод обучения модели aiNet включает следующие этапы:

1. Дана обучающая выборка антигенов Ag . Инициализировать иммунную память модели $M = \emptyset$. Для каждого антигена из выборки $ag_i \in Ag$, состоящей из n элементов, где $i = 1, \dots, n$, выполнять п. 1.1-1.10.

1.1. Вычислить аффинность антигена к каждому элементу из набора антител $ab_j \in Ab$, где $j = 1, \dots, m$; m – объем множества антител. Значение аффинности обратно пропорционально Евклидовому расстоянию между антигеном и антителом и вычисляется по формуле:

$$aff(ag_i, ab_j) = \frac{1}{\|ab_j - ag_i\|}.$$

1.2. Из набора антител Ab выбрать подмножество Ab^* , состоящее из q антител с наиболее высоким значением аффинности.

1.3. Сгенерировать множество клонов C из q антител пропорционально их уровню аффинности: чем выше значение аффинности антитела к антигену, тем большее количество клонов необходимо создать.

1.4. Клоны из множества C проходят этап управляемой мутации, результатом которого является множество мутировавших клонов C^* :

$$C_j^* = C_j + \alpha_j \cdot (ag_i - ab_j),$$

где α_j – коэффициент, определяющий степень мутации антитела ab_j . Степень мутации каждого клона обратно пропорциональна аффинности его родительской клетки к антигену.

1.5. Вычислить взаимную аффинность антигена ag_i к клонам антител из C^* по формуле:

$$d_{ij} = \frac{1}{\|C_j^* - ag_i\|}.$$

1.6. Отобрать некоторый процент δ клонов с наилучшей аффинностью d из C^* и поместить их в клональную память M_j .

1.7. Удалить из клональной памяти M_j те элементы, у которых аффинность d ниже некоторого порогового значения σ_d .

1.8. Произвести клональную супрессию – определить взаимную аффинность клонов по формуле:

$$s(ab_j, ab_k) = \|ab_j - ab_k\|,$$

где ab_j, ab_k – антитела из клональной памяти M_j .

1.9. Удалить те клоны, для которых значение аффинности s превышает заданный порог σ_d .

1.10. Добавить оставшиеся клоны в иммунную память $M = M \cup M_j$.

2. Определить взаимную аффинность каждой пары антител $s(ab_p, ab_r) = \|ab_p - ab_r\|$, где $p = 1, \dots, q$, $r = 1, \dots, q$.

3. Произвести сетевую супрессию — удалить те антитела, для которых аффинность s превышает порог σ_s .

4. Если критерий останова не достигнут — перейти к этапу 1, в противном случае — к этапу 5.

5. Останов.

Как видно из описания метода, на каждой итерации отбирается часть антител с наиболее высокой аффинностью к заданному антигену. Такие антитела подвергаются клонированию пропорционально значению аффинности. Затем клоны мутируют, и определенная их часть остается в иммунной памяти. При этом из сети удаляются клетки, чья взаимная аффинность превышает порог, тем самым регулируется размер популяции без существенного влияния на работу модели.

При использовании сетей aiNet на практике применяют следующие критерии останова:

- достижение максимального количества итераций;
- превышение максимальной популяции сети;
- возрастание ошибки на k последних последовательных итерациях.

Результатом обучения сети является набор антител, которые отображают внутреннюю структуру антигенов из обучающей выборки. С помощью обученной сети информацию о кластерах можно получить, проанализировав связи между антителами в сети. Количество кластеров можно изменять, варьируя порог отсечения связей между антителами в сети.

3. Расширяющийся нейронный газ

Метод расширяющегося нейронного газа представляет собой итеративный метод кластеризации с добавлением кластеров. Для заданного входного распределения (обучающей выборки) строится граф (сеть узлов), каждый узел которого описывает позицию экземпляра обучающей выборки в пространстве признаков. Метод расширяющегося нейронного газа является адаптивным, поскольку структура газа изменяется, реагируя на изменения во входных данных.

Ниже описан метод обучения модели расширяющегося нейронного газа. Изначально модель расширяющегося нейронного газа представлена в виде двух случайных узлов, соединенных связью с нулевым “возрастом”. Возраст определяется количеством итераций, прошедших с момента последнего взаимодействия узла сети с некоторым экземпляром из обучающей выборки.

1. Установить номер текущей итерации $N=0$. Задать значения коэффициентов уменьшения ошибок α и β . Установить интервал добавления новых узлов в сеть λ . Построить сеть, содержащую два случайных узла с нулевой локальной ошибкой ϵ и нулевым возрастом a .

2. Из обучающей выборки получить вектор x , описывающий некоторый экземпляр.

3. Среди узлов сети найти два узла s и t , представленных векторами w_s и w_t соответственно, таким образом, чтобы расстояние $\|w_s - x\|$ было минимальным, а $\|w_t - x\|$ — вторым минимальным.

4. Увеличить локальную ошибку в узле s :

$$\epsilon_s = \epsilon_s + \|w_s - x\|^2.$$

5. Переместить узел s и все связанные с ним узлы по направлению к x следующим образом:

$$w_s = w_s + \epsilon_w \cdot (x - w_s),$$

$$w_n = w_n + \epsilon_n \cdot (x - w_n),$$

где n — узел, соседний с s , ϵ_w и ϵ_n — коэффициенты, определяющие расстояние, на которое смещаются узлы.

6. Увеличить возраст всех связей между узлом s и его соседними узлами.

7. Если узлы s и t соединены связью, то обнулить ее возраст $a_{\{s,t\}} = 0$, в противном случае — создать связь с нулевым возрастом между ними.

8. Удалить все связи, возраст которых превышает заданное пороговое значение a_{max} .

9. Удалить все узлы, не имеющие связей с соседними узлами.

10. Если порядковый номер текущей итерации N кратен λ , и при этом не достигнуто заданное максимальное количество узлов, то выполнить пп. 10.1-10.5, в противном случае — перейти к этапу 11.

10.1. Найти узел u с максимальным значением ошибки ϵ_u .

10.2. Среди связанных с u узлов найти узел v с наибольшим значением ошибки ϵ_v .

10.3. Создать узел r , описанный вектором w_r между узлами u и v :

$$w_r = \frac{(w_u + w_v)}{2}.$$

10.4. Создать связи между r и v , r и u . Удалить связь между u и v .

10.5. Согласно значению коэффициента α изменить значение ошибок в узлах u и v :

$$\epsilon_u = \epsilon_u - \alpha \cdot \epsilon_u,$$

$$\epsilon_v = \epsilon_v - \alpha \cdot \epsilon_v,$$

$$\epsilon_r = \epsilon_u.$$

11. Уменьшить значения ошибок каждого узла сети соответственно коэффициенту β :

$$\epsilon_j = \epsilon_j - \beta \cdot \epsilon_j,$$

где j — узел сети.

12. Если критерий сходимости не достигнут — увеличить номер итерации и перейти к этапу 2, в противном случае — перейти к этапу 13.

13. Останов.

В расширяющемся нейронном газе локальные ошибки нейронов накапливаются, в то время как

метод обучения приводит к минимизации этой ошибки. Для того, чтобы новые ошибки имели большее влияние на структуру газа, а также для снижения темпов роста локальной ошибки, на каждой итерации выполняется глобальное уменьшение ошибки пропорционально ее значению.

Изменение структуры газа осуществляется за счет перемещения отдельных узлов. На каждой итерации узел, наиболее близкий к входному вектору, а также его соседние узлы, перемещаются в направлении, снижающем ошибку. Все эти узлы перемещаются по направлению к входному вектору на некоторое расстояние, определяемое коэффициентами ϵ_w и ϵ_n . Фактически эти коэффициенты имеют наибольшее влияние на скорость сходимости метода обучения модели ИИС.

Целесообразно использовать такие значения ϵ_w и ϵ_n , чтобы коэффициент ϵ_w значительно превышал ϵ_n . Это обеспечит плавность и высокую скорость сходимости метода обучения. Малые значения этих коэффициентов снижают эффективность и скорость работы метода, а высокие приводят к построению модели, неустойчивой к выбросам.

В [7] было установлено экспериментально, что оптимальное значение коэффициента $\epsilon_w \approx 0.05$, а значение ϵ_n должно быть на пару порядков ниже. Однако более точные значения ϵ_w и ϵ_n зависят от входных данных конкретной решаемой задачи.

Еще одним важным управляющим параметром модели расширяющегося нейронного газа является частота добавления новых узлов. Из описанного выше метода видно, что узлы сети добавляются на каждой λ -ой итерации, т.е. через определенные фиксированные интервалы. Частота добавления новых узлов оказывает существенное влияние на производительность метода обучения. Занижен-

ные значения коэффициента λ приводят к резкому снижению ошибки в начале работы метода, поэтому на ранних итерациях узлы сети будут иметь меньше шансов для распространения, и модели потребуется больше времени для достижения требуемой структуры, чем при более высоких значениях λ . Низкое значение λ также приводит к формированию неактивных узлов – узлов, которые не способны адаптироваться к входным данным из-за большой отдаленности от них. Такие узлы не влияют на эффективность модели, но повышают её потребность в вычислительных ресурсах. С другой стороны, слишком большие значения λ снижают темпы роста модели, что приводит к медленной сходимости метода обучения.

4. Сравнительный анализ моделей aiNet и расширяющегося нейронного газа

В таблице приведены результаты сравнительного анализа моделей aiNet и расширяющегося нейронного газа.

Как видно из таблицы, обе модели поддерживают разнообразие в популяции за счет различных механизмов адаптации. Модели обладают подобным внутренним представлением данных и общей схемой работы, что определяет их способность к обобщению и устойчивость к шумам.

Однако имеются существенные отличия на этапах отбора узлов и их модификации. Отличия заключаются не только в количестве модифицируемых узлов за одну итерацию, но и в способе их модификации. Тот факт, что рассмотренная выше модель ИИС использует клональную селекцию для создания новых узлов, определяет более высокую скорость сходимости метода ее обучения. С другой стороны, накопление локальной ошибки в узлах

Таблица

Результаты сравнительного анализа моделей aiNet и расширяющегося нейронного газа

Характеристики	Искусственная иммунная сеть	Расширяющийся нейронный газ
Структура	Взвешенный граф	Взвешенный граф
Узлы	Антитела	Нейроны
Ребра	Сила связи антител (степень подобия)	Количество итераций, прошедших с момента взаимодействия соединенных узлов с экземпляром из обучающей выборки
Представление данных	Вектор вещественных чисел	Вектор вещественных чисел и накопленная локальная ошибка
Метрика	Евклидово расстояние	Евклидово расстояние
Механизм адаптации	Клональная селекция	Линейное перемещение узлов
Отбор узлов	Отбираются несколько антител с наибольшей аффинностью к антигену	Отбираются два узла: ближайший узел и ближайший из связанных с ним узлов
Динамика популяции	Удаляются клетки, обладающие низкой аффинностью к антигенам и высокой взаимной аффинностью (расположенные достаточно близко друг с другом)	Удаляются клетки без связей и связи после определенного периода неактивности
Инициализация	Несколько случайно расположенных узлов	Два случайно расположенных узла
Добавление элементов	Происходит путем клонирования существующих антител как реакция на антиген	На некоторых итерациях с заданным заранее интервалом добавляется один узел между двумя другими узлами с наибольшей локальной ошибкой

сети в модели расширяющегося нейронного газа служит своего рода “памятью” нейронов и позволяет принимать решения о модификации узлов с учетом их реакции на предыдущие итерации, что делает изменение структуры модели более “плавным”.

Исходя из этого, предлагается создать гибридную модель ИИС, использующую преимущества расширяющегося нейронного газа.

5. Гибридная модель ИИС

В предлагаемой авторами модели антитела характеризуются локальной ошибкой, изменяющейся всякий раз, когда антитело обладает достаточной аффинностью к некоторому антигену из обучающей выборки. Это позволит использовать значение локальной ошибки в качестве управляющего параметра на этапе клональной селекции, т.е. создавать популяцию клонов, основываясь не на аффинности антитела к текущему антигену, а на его локальной ошибке. Локальная ошибка аккумулирует в себе результат предыдущих взаимодействий антитела с антигенами, и высокое ее значение свидетельствует о том, что данное антитело находится в области с высокой концентрацией антигенов, а, следовательно, необходимо провести клональную селекцию для более точного покрытия этой области. С другой стороны, низкая локальная ошибка говорит о том, что в окрестности данного антитела либо находится небольшое количество антигенов, либо существует достаточное количество антител, чтобы их распознавать. Ниже приведен метод обучения гибридной модели.

1. Инициализировать иммунную память M двумя случайными антителами с нулевой локальной ошибкой.

2. Для каждого антигена из обучающей выборки $ag \in Ag$ выполнить п. 2.1-2.7.

2.1. Отобрать из иммунной памяти M некоторый процент δ антител Ab^* с наибольшей аффинностью к антигену ag .

2.2. Накопить для отобранных антител Ab^* локальную ошибку

$$\varepsilon_{ab} = \varepsilon_{ab} + \|ab - ag\|,$$

где $ab \in Ab^*$, ab – рассматриваемое антитело, ε_{ab} – его локальная ошибка.

2.2. Сгенерировать множество клонов C для каждого из отобранных антител Ab^* пропорционально их локальной ошибке. Антитела с большей локальной ошибкой произведут большее число клонов, вычисляемое по формуле:

$$N_C = N_{max} \cdot \left(1 - \frac{1}{\varepsilon_{ab}}\right),$$

где N_{max} – максимально возможное количество клонов одного антитела на одной итерации.

2.3. Клоны из множества C подвергаются мутации. Степень мутации также прямо пропорциональна локальной ошибке родительского антитела.

2.4. Выполнить клональную супрессию и сетевую супрессию, как описано выше в методе обучения ИИС (п. 1.5-1.9).

2.5. Оставшиеся после этапа супрессии клоны C^* поместить в иммунную память $M = M \cup C^*$.

2.6. Установить уровень локальной ошибки полученных клонов C^* равным нулю.

2.7. Если не достигнут критерий останова – перейти к этапу 2, в противном случае – к этапу 3.

3. Останов.

Для антител с низкой локальной ошибкой (очень близких к текущему антигену) вполне возможно, что не будет произведено ни одного клона. Таким образом, когда все антитела будут достаточно близко находиться к антигенам из обучающей выборки, суммарная ошибка сети будет оставаться практически неизменной. Суммарная ошибка рассчитывается по формуле:

$$\varepsilon^* = \sum_{ab \in Ab} \varepsilon_{ab}.$$

Предлагается использовать это свойство в качестве критерия останова для приведенного выше метода обучения.

Рассмотрим влияние основных управляющих параметров метода на его работу. Под управляющими параметрами будем понимать количество антител, допускаемых к клонированию, максимальное количество клонов, максимальную степень мутации. Коэффициенты, отвечающие за проведение клональной и сетевой супрессии подробно рассмотрены в [8] и не имеют специфичного воздействия на предложенную модель.

Важно отметить, что количество антител, допускаемых к клонированию, должно превышать количество клонов, поскольку в противном случае на каждой итерации будут производиться новые антитела с нулевой ошибкой, которые не будут успевать ее накапливать. Большое значение δ целесообразно использовать на начальных итерациях метода, поскольку это приведет к быстрому распространению антител. В то же время, когда модель будет уже сформирована, отбор большого количества антител для клонирования дестабилизирует иммунную сеть и ухудшает ее характеристики.

Увеличить скорость работы метода обучения модели можно за счет изменения способа мутации клонов. Так, если мутирующие клоны расположены на одном отрезке, соединяющем родительское антитело и текущий антиген, то поведение модели очень напоминает расширяющийся нейронный газ. Если же клоны создаются в некотором радиусе от родительского антитела, и радиус зависит от степени мутации клонов, то модель имитирует поведение иммунной сети.

6. Эксперименты и результаты

Предложенный метод был программно реализован на языке Java. Разработанное программное обеспечение позволяет осуществлять разбиение

на кластеры множества входных данных. Помимо разделения на кластеры, программное обеспечение предоставляет в текстовом и графическом виде информацию о динамике локальной ошибки модели.

Предложенная модель тестировалась на синтетических выборках данных, взятых из каталога SPAETH (<http://people.sc.fsu.edu/~jburkardt/datasets/spaeth>). Оценивались такие параметры модели, как скорость обучения, динамика популяции, скорость снижения ошибки и точность работы. Выборки содержали небольшое число экземпляров (50-100), описанных векторами вещественных чисел. Большинство выборок содержали двумерные данные, однако экспериментально было установлено, что предложенный метод также хорошо работает при высоких размерностях данных.

На рисунке представлены графики динамики суммарной ошибки ϵ^* предложенной гибридной модели и модели расширяющегося нейронного газа.

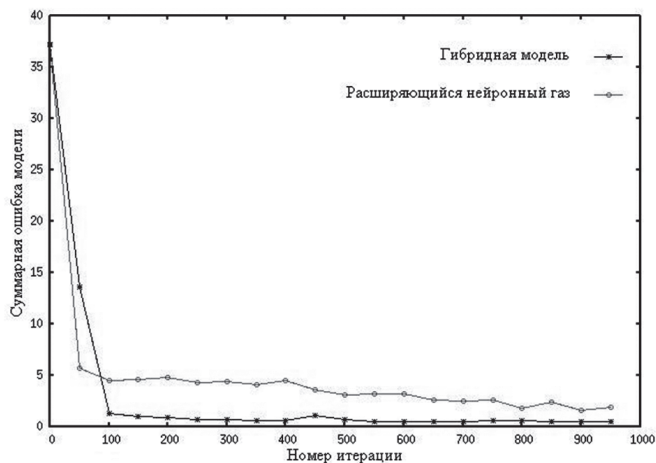


Рис. 1. Динамика суммарной ошибки гибридной модели и модели расширяющегося нейронного газа

Более низкие значения суммарной ошибки гибридной модели объясняются тем, что в ней минимизируется локальная ошибка с помощью механизма клональной селекции, но это не приводит к избыточному количеству клеток из-за механизмов супрессии, сокращающих популяцию.

Выводы

С целью решения актуальной задачи автоматизации кластер-анализа выборки данных разработано математическое обеспечение, позволяющее проводить кластеризацию на основе методов иммунокомпьютинга.

Научная новизна работы заключается в том, что впервые предложен метод кластеризации, использующий гибридную модель, основанную на модели ИИС aiNet и принципах расширяющегося нейронного газа, что позволяет повысить скорость сходимости метода обучения модели.

Практическая ценность работы заключается в том, что было разработано программное обеспе-

чение для проведения кластер-анализа с помощью предложенной гибридной модели ИИС.

Тестирование предложенной гибридной модели ИИС показало высокую скорость сходимости метода обучения модели, что позволяет рекомендовать ее использование для решения практических задач.

Дальнейшие исследования могут быть сосредоточены на анализе влияния вида мутации клонов антител на эффективность работы метода обучения модели ИИС.

Список литературы: 1. Jain A.K., Murty M.N., Flynn P.J. Data clustering: a review // ACM Computing Surveys. 1999. №31(3). P. 264-323. 2. Tang N., Rao Vemuri V. An artificial immune system approach to document clustering // Proceedings of the 2005 ACM symposium on Applied computing (2005). New York, USA: ACM, 2005. P. 918-922. 3. De Castro L. N., Von Zuben F.J. aiNet – an artificial immune network for data analysis // Data Mining: A Heuristic Approach. Eds. H. A. Abbass, R. A. Sarker, C. S. Newton. – USA, Miami: Idea Publishing Group, 2001. – P. 231-259. 4. Jirayusakul I A., Auwatanamongkol S. A supervised growing neural gas algorithm for cluster analysis // International Journal of Hybrid Intelligent Systems. 2007. №4(4). P. 217-229. 5. Jerne N.K. Towards Network Theory of the Immune System // Ann. Immunol. 1974. №125C. P. 373-389. 6. De Castro L.N., Von Zuben F.J. An evolutionary immune network for data clustering // Brazilian Symposium on Artificial Neural Networks: Proceeding of the IEEE SBRN (22-25 November 2000). Zurich: Springer-Verlag, 2000. P. 84-89. 7. Holmstrom J. Growing neural gas: experiments with GNG, GNG with utility and supervised GNG // Thesis: M.Sc. in Computer Science. – Sweden, Uppsala: Uppsala University, 2002. – 42 p. 8. De Castro L.N. The immune response of an artificial immune network (aiNet) // Proceedings of the 2003 Congress on Evolutionary Computation (8-12 December 2003). Chicago, USA: IEEE Press, 2003. P. 146-153.

Поступила в редакцию 8.06.2010.

УДК 004.93

Кластерний аналіз з використанням гібридної моделі на основі штучної імунної мережі / С. О. Зайцев, С. О. Субботін // Біоніка інтелекту: наук.-техн. журнал. – 2010. – № 3 (74). – С. 70–75.

Проведено порівняльний аналіз моделі штучної імунної мережі aiNet та моделі нейронного газу, що поширюється, в задачі кластер-аналізу. Запропоновано гібридну модель штучної імунної мережі на основі проаналізованих моделей, що дозволяє підвищити швидкість навчання мережі. Проведено експерименти, що підтверджують ефективність запропонованої моделі.

Табл. 1. Іл. 1. Бібліогр.: 7 найм.

UDC 004.93

Using a hybrid model based on the artificial immune network in cluster-analysis / S. A. Zaitsev, S. A. Subbotin // Bionics of Intelligence: Sci. Mag. – 2010. – № 3 (74). – P. 70–75.

The role of the artificial immune network model (aiNet) and growing neural gas model in cluster-analysis has been analyzed. A hybrid model based on the artificial immune network has been proposed, which allows to perform faster network training. The experiments have been carried to approve the efficiency of the suggested model.

Tab. 1. Fig. 1. Ref.: 7 items.

УДК 004.89



РАЗРАБОТКА КОМПЛЕКСНОЙ ТЕХНОЛОГИЧЕСКОЙ МОДЕЛИ ИЗДЕЛИЯ В ВИДЕ СЕМАНТИЧЕСКОЙ СЕТИ

А.С. Котов

ХНУРЭ, г. Харьков, Украина

Рассматривается проблема выбора технологических решений при вариантном синтезе и оценке технологических процессов производства. На основе концепции единого информационного пространства предприятия и знаниеориентированной поддержки принятия технологических решений предлагается комплексная технологическая модель изделия, которая включает классификационную модель детали и модель технологического процесса с учетом оборудования и другой дополнительной информации. Модель представлена в виде семантической сети. При этом определены виды объектов, семантические типы узлов, их атрибуты, определены виды связей между объектами.

ПРЕДПРОЕКТНЫЙ ЭТАП, ТЕХНОЛОГИЧЕСКОЕ РЕШЕНИЕ, ТЕХНОЛОГИЧЕСКИЙ ПРОЦЕСС, ИЗДЕЛИЕ, СЕМАНТИЧЕСКАЯ СЕТЬ, ОБЪЕКТЫ, АТРИБУТЫ, СВЯЗИ

Введение

Предприятия, стремясь сохранить свои позиции на рынке в условиях конкуренции и нестабильности, вынуждены непрерывно улучшать производственный процесс, осваивая новые технологии и оборудование. Особую остроту приобретают вопросы улучшения методов планирования и управления процессом развития, разработкой инновационной продукции и прогнозирования мер по адаптации к изменяющимся требованиям внешней среды. В современных рыночных условиях на смену крупносерийному производству приходит в основном мелкосерийное, постоянно модернизируемое и обновляемое производство, способное оперативно реагировать на колебания рыночной конъюнктуры. Это обстоятельство в свою очередь вынуждает оперативно разрабатывать и внедрять новые технологические процессы, оборудование, материалы.

1. Постановка задачи исследования

На современном этапе развития информационных технологий управления предприятием и прогнозирования его развития проблемы интеграции информационных ресурсов занимают одно из центральных мест. Поэтому интеграция информационных систем и распределенных баз данных на общей методологической основе является актуальной проблемой.

Конкретное управленческое решение требует анализа специально отобранной информации [1]. Обмен информацией является необходимым элементом управления предприятием, подверженным влиянию динамики постоянно изменяющегося рынка. Частью концепции интегрированного управления информационными ресурсами является разработка ряда информационных моделей основных производственных объектов и их элементов. На этой основе осуществляется поддержка принимаемых планово-оптимизационных решений.

Эффективность разработки новых изделий в первую очередь определяет технологический уро-

вень производства. Сегодня в конкурентной борьбе выигрывает не то предприятие, которое выводит на рынок идеальный продукт, а то, которое делает это быстрее, отвечая тем самым на быстро изменяющиеся потребности. Новой концепцией сокращения «времени до рынка» является fast new product development (FPD) — концепция «быстрой разработки новых продуктов», которая приобрела статус одной из самых перспективных управленческих доктрин.

Основным направлением повышения технологического уровня производства является разработка формализованных процедур оценивания существующих технологических решений и предлагаемых новых технологий, их влияния на общий уровень и затраты производства. Основные научно-технические вопросы, которые необходимо решать, связаны с выбором типового технологического процесса (ТП), определением состава стандартного оборудования и оснастки.

Разработка моделей структуризации и методов быстрого поиска и обработки информации о существующих технологических решениях при проектировании технологических процессов позволит существенно сократить сроки проектирования и модификации ТП.

Существующие подходы в использовании технологического информационного обеспечения основаны на традиционной структуризации информации в виде баз данных [2]. В соответствии с концепцией информационной интеграции модель изделия, созданная на стадии конструкторской подготовки производства, передается на дальнейшие стадии, где дополняется технологической информацией [3].

В последнее время развиваются методы и модели поиска решения на основе структурной аналогии. Основное внимание уделяется методу структурной аналогии свойств, позволяющему учитывать контекст [4].

Использование элементов синтаксического анализа призвано приблизить процесс описания

предметной области к нормам естественного языка. Развитие этих элементов привело к появлению двух программных продуктов, Knowlegist и GoBrain, которые позволяют подвергать семантическому анализу большие объемы текстовой информации, трансформируя их в базы знаний, существенно расширяющие поисковые возможности.

Сегодня возникла задача системного управления знаниями предприятия, которая решается путем накопления структурированных, формализованных знаний, позволяющих сделать знания доступными и повторно используемыми. Основная цель управления знаниями состоит в эффективном представлении данных и информационных ресурсов для принятия решений.

В системах управления знаниями (СУЗ) в качестве знаний выступает вся доступная информация (конструкторская, технологическая и техническая документация, описание продукции и т. д.), а также закономерности предметной области, полученные из практического опыта или внешних источников. На этапе структурирования необходимо разрабатывать модели представления знаний, которые обеспечивали бы автоматизированную обработку информации на семантическом уровне в системах управления знаниями (СУЗ).

Сформулируем *требования* к системе выбора и оценки технологических решений.

1. Система должна быть ориентирована на работу с информационными моделями технического изделия, ресурсами производственной системы, с информационной моделью ТП.

2. В системе должны присутствовать накопленные знания по технологическому проектированию.

3. Для обеспечения интеграции систем проектирования, подготовки и управления производством необходимо единство представления модели технического изделия и методов работы с ней в едином информационном пространстве.

Для обеспечения этих требований следует решить такие задачи:

- описать информационные ресурсы формируемого техпроцесса;
- создать формализованную информационную модель изделия;
- сформировать структуру знаний по реализации техпроцесса.

Целью данной статьи является разработка формализованной технологической модели изделия с применением аппарата семантических сетей.

2. Решение задачи

Рассмотрим задачу организации информационного взаимодействия процессов предпроектной подготовки производства. Предпроектные работы включают в себя первоначальный анализ выполнения проекта. Укрупненная схема информацион-

ной поддержки предпроектных работ с использованием накопленного опыта предприятия приведена на рис. 1.

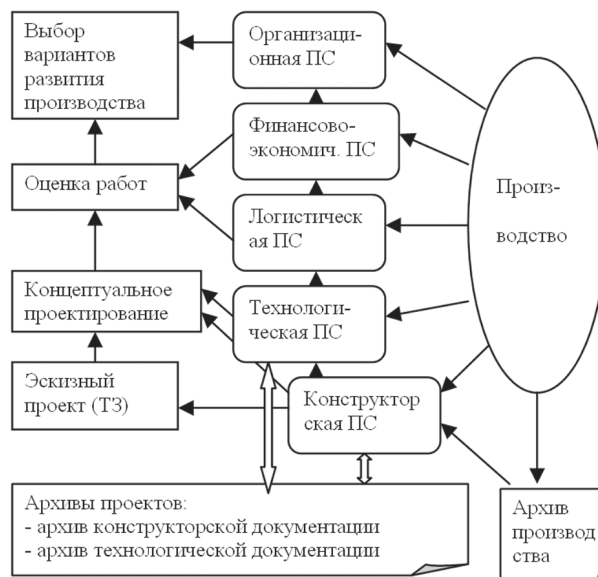


Рис. 1. Укрупненная схема информационной поддержки предпроектных работ с использованием накопленного опыта

Каждой подсистеме предприятия доступны знания всего предприятия — опыт его лучших профессионалов, сведения о последних научных разработках, технологиях, методах.

На этапе описания ресурсов в систему заносятся данные, необходимые при формировании информационной модели техпроцесса:

- технологические ресурсы - основное оборудование, оснастка, материалы, комплектующие изделия и т.п.; каждый из объектов обладает своим набором параметров;
- технологический процесс — совокупность технологических операций, расположенных во временной последовательности, составляющих пооперационный технологический процесс.

Для формализованного описания технологических и производственных объектов исходные данные представляют в виде многоуровневой фреймовой структуры, абстрактную модель предметной области — в виде организованной совокупности множества понятий, а конкретную модель ситуации — в виде совокупности взаимосвязанных экземпляров этих понятий; используются механизмы семантического описания элементов [5]. Семантические правила определяют типовые фразы для записи этапов, групп операций и операций [6]. Структура правил записи и взаимосвязей между элементами соответствует нормативно-технической документации.

Общий принцип представления исходного объекта описания заключается в том, что он состоит из множества элементов. Поэтому для представления объекта описания следует:

- определить объекты понятий;
- описать свойства объектов понятий;
- описать элементы.

Выявление понятий для описания объекта начинается с выделения центрального понятия, затем понятий, отражающих их ассоциации и второстепенных понятий. Для выделения центральных понятий следует использовать следующие рекомендации:

- центральное понятие является обобщающим понятием его составных частей;
- центральное понятие каким-либо образом управляет деятельностью или состоянием других понятий;
- обособленное центральное понятие может не иметь связей с другими понятиями предметной области, но являться неотъемлемой частью описания ситуации.

Получив список центральных понятий, необходимо перейти к поиску их связей с другими понятиями. Зная о наличии таких связей и действуя подобно поиску центральных понятий, можно выявить контекст этих понятий. Для этого, рассматривая представление предметной области и изучив положение центрального понятия в проблемной ситуации и его взаимосвязь с другими элементами модели, следует провести классификацию полученных связей по соответствующим категориям.

Так, работая над объектом S , получаем два множества кандидатов: cE – множество понятий, являющихся сущностями, и cR – множество понятий, выражающих отношение между другими понятиями. Первым шагом является выделение из множества кандидатов cE множества центральных понятий – cC . Затем поиск центральных понятий проводится среди остальных элементов множества кандидатов cE .

Следующим шагом будет выделение полученных кандидатов в центральные понятия, а затем – выделение контекста. Для этого из множества кандидатов cE удаляются элементы множества центральных понятий cE , в результате чего получится множество отношений и второстепенных понятий cA – cE/cC . Задача состоит в выделении подмножества отношений (ассоциаций) $cR \subset cA$.

Ассоциация центрального понятия с другими понятиями может выражать следующие типы активных связей между этими понятиями:

- зависимость,
- подчиненность,
- отношение агрегирования.

Таким образом, руководствуясь списком типов активных и дополнительных связей, получим элементы множества отношений cR . Каждый из элементов множества отношений $r_i \in cR$ будет связан с определенным центральным понятием $c_i \in cC$, формируя множество описаний элементов $s_i = \langle c_i, R_{c_i} \rangle$, в котором элементы множеств R_{c_i} являются элементами множества cR . На данном

этапе полученные описания элементов еще не являются полными, поскольку элементам множества cR не поставлены в соответствие второстепенные понятия. Для этого необходимо провести поиск второстепенных понятий.

В качестве второстепенных понятий могут выступать любые элементы множества кандидатов cE независимо от того, попали ли они во множество центральных понятий cC , множество отношений cR или ни в одно из них. В большинстве случаев второстепенные понятия получаются из дополнений к глагольным формам. В общем случае второстепенные понятия – это те понятия, на которые ссылаются ассоциации.

Полученное таким образом технологическое описание изделия (ТОИ) будет состоять из множества элементов, описывающих общее представление состояния предметной области без свойств понятий, отражающих сущности и отношения. Этого описания уже достаточно для поиска описаний аналогичных разработок, похожих по структуре. Получив значения атрибутов для понятий ТОИ, результаты поиска можно еще приблизить к наиболее подходящему результату.

Входной информацией для построения семантической модели является эскизный проект на изготовление изделия. Эскизный проект разрабатывают с целью установления принципиальных решений изделия, дающих общее представление о принципе работы и устройстве изделия, степени новизны, технологичности и т.д. На стадии разработки эскизного проекта рассматривают варианты изделия и (или) его составных частей.

В комплект документов эскизного проекта включают конструкторские документы, предусмотренные техническим заданием и протоколом рассмотрения технического предложения [7]. Эскизный проект включает текстовую, графическую и количественную информацию [8]. Кроме того, большинство подразделов документов являются взаимосвязанными: одна информация обуславливает (уточняет) другую или является ее составляющей частью. Поэтому для структуризации информации о разрабатываемом изделии целесообразно применение аппарата семантических сетей.

Для построения семантической модели технологического описания объекта на основе эскизного проекта на изделие следует:

- проанализировать содержание эскизного проекта с целью выделения основных элементов семантической сети;
- определить виды элементов, типы их возможных значений;
- сформировать структуру – выделить уровни иерархии элементов, образовать связи, указать их направленность;
- определить виды объектов – элементов семантической сети;

- определить семантические типы узлов, указать их атрибуты;
- указать виды связей (их атрибуты).

Построение семантической сети начинается с выделения составных элементов, выступающих в виде объектов описания. Исходя из содержания эскизного проекта разработки нового изделия и задачи выбора и оценки технологического способа изготовления изделия, определим две основных составляющих комплексной технологической модели изделия (на примере детали заготовительной фазы производства) (рис. 2):

- классификационная модель детали;
- модель технологического процесса с учетом оборудования и другой дополнительной информации.

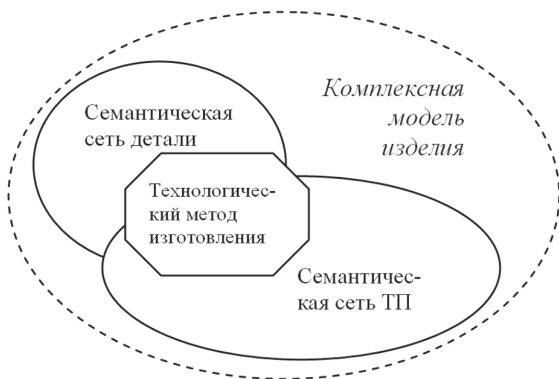


Рис. 2. Концептуальное представление составляющих комплексной технологической модели изделия

Структура классификационной модели детали представлена на рис. 3. Отметим, что объекты сети образуют семь уровней иерархии. Данная структура использована при разработке семантической модели классификационных элементов детали, которая включает описание атрибутов узлов сети и

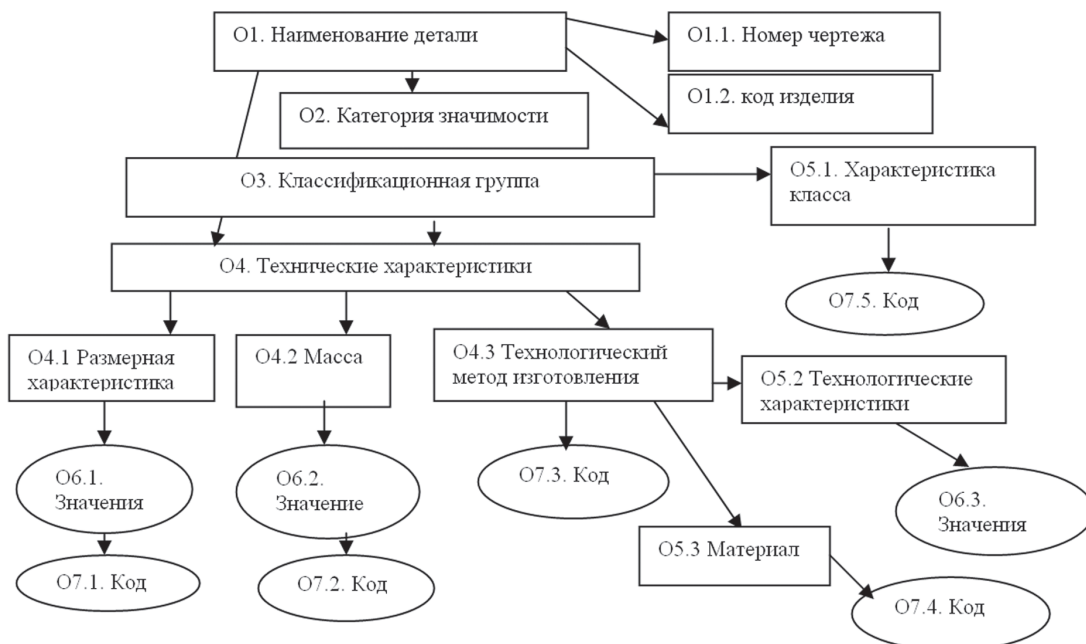


Рис. 3. Структурные элементы семантической сети описания детали

указание видов связей (рис. 4).

Узлы и связи в предложенной семантической сети имеют набор следующих основных атрибутов:

- ObjectType — вид объекта, соответствующего узлу: *центральное понятие (Central)*, *понятие (Ass)*, *группа понятий (Group)*, *контекст (Context)*, *второстепенное понятие (Second)*;
- SemanticType — семантический разряд референта узла; основные выделяемые разряды: *наименование (Name)*, *действие (Act)*, *состояние (State)*, *предмет (Item)*, *свойство (Qva)*, *число (Num)* и *пр.*;
- Name — строка текста, соответствующего узлу; для именных групп может иметь несколько значений, которые представляют словосочетания, образованные от ключевого существительного;
- Relation — тип синтактико-семантической связи между узлами: *агрегирование (или принадлежность классу) (Agr)*, *подчиненность (Sub)*, *зависимость (Corr)*, *атрибут (Atr)*.

Далее выделим следующие виды объектов: центральное понятие (наименование детали)

– ObjectType = “Centr”[O1] ;

– второстепенное понятие (номер чертежа, код изделия)

ObjectType = “Second”[O1.1, O1.2] ;

– контекст (категория значимости, классификационная группа, технологический метод, масса, материал, размерная характеристика)

ObjectType =

= “Context”[O2, O3, O4, O4.1-O4.3, O5.3];

– группа понятий (характеристики группы, технические, технологические характеристики)

ObjectType = “Group”[O4, O5.1, O5.2] .

При описании элементов были введены следующие семантические типы (SemanticType):

1) наименование (Name) – наименование детали, номер чертежа, код изделия. Эти элементы являются текстовыми, представляют собой сугубо информационные характеристики технологического объекта. Номер чертежа содержит цифровые элементы, но его значение рассматривается в целом как текстовое;

2) свойство (Qva) – категория значимости, классификационная группа, технические и технологические характеристики. Эти элементы являются характеризующими, они описывают свойства технологического объекта;

3) число (Num) – размерная характеристика, масса, некоторые технологические характеристики;

4) состояние (State) – значения отдельных технологических характеристик; чаще всего представляют собой числовые и текстовые значения;

5) предмет (Item) – предмет, материал; описывает элемент материального (или технического) обеспечения изготовления объекта;

б) действие (Act) – технологический метод изготовления.

Определим типы связей, используемых для построения семантической сети (Relation):

1) атрибут (Atr) – наиболее часто присутствующий в семантической сети тип связи. При этом

объект нижнего уровня иерархии является атрибутом (уточняющим, раскрывающим суть) объекта верхнего уровня. Например, сама деталь, определенная своим наименованием имеет атрибут «классификационная группа». Для поиска аналогичных изделий более существенными являются атрибуты, а не сами понятия;

2) зависимость (Corr) – выражает зависимость между объектами. Чаще эта связь возникает между объектами одного уровня иерархии и является двунаправленной. При этом зависимость может быть явно выражена только для составляющих элементов данных объектов (для сети второго уровня). Например, технологический метод изготовления иногда зависит от массы. Анализ связей зависимости в сети позволяет:

- определить непротиворечивость сети;
- оптимизировать (минимизировать) ее содержание путем удаления дублирующей информации (сильно коррелированной);

3) агрегирование (Agr) – выражает агрегирование нескольких элементов нижнего уровня одним элементом верхнего уровня. Или наоборот: декомпозицию одного элемента верхнего уровня на несколько однородных элементов нижнего уровня.

Поиск аналогичных изделий осуществляется с

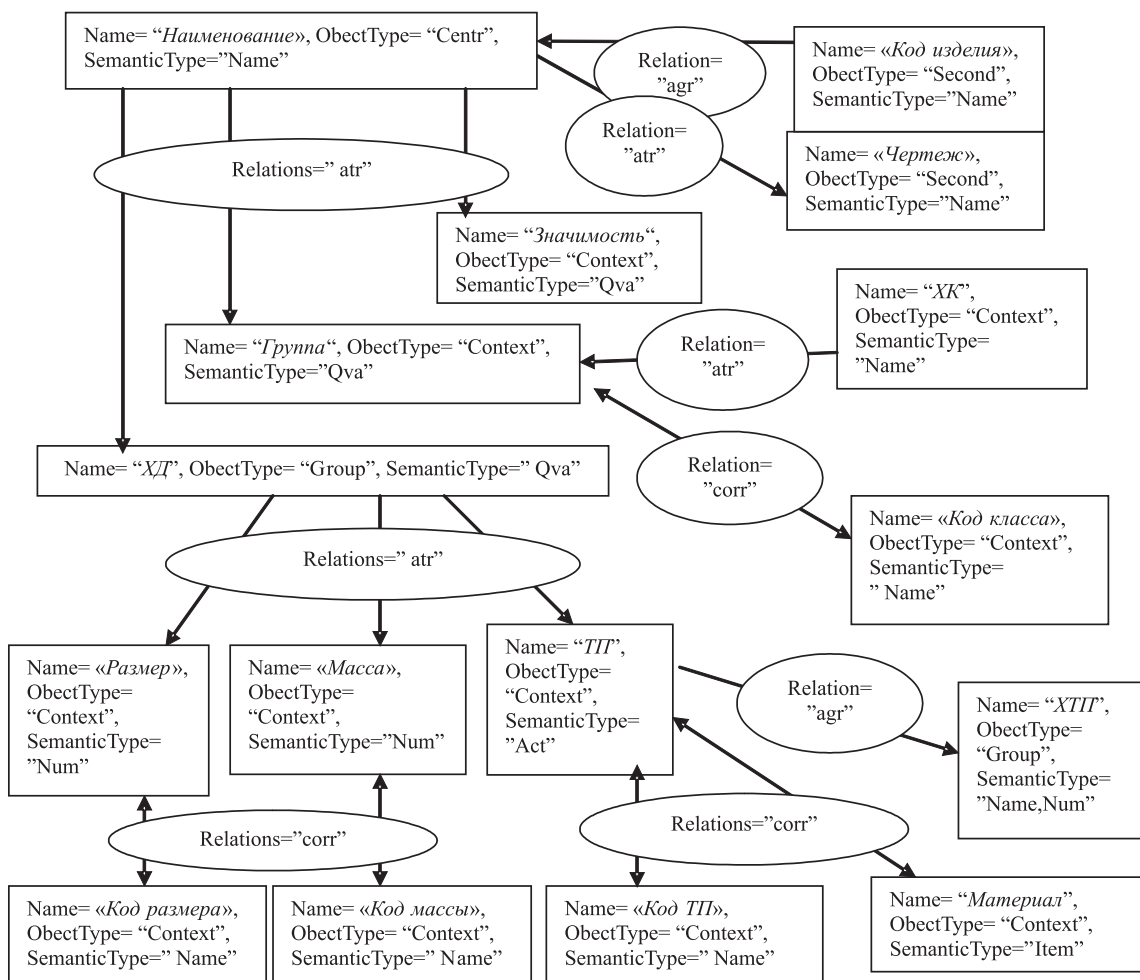


Рис. 4. Семантическая модель описания классификационных характеристик детали

использованием отдельных (наиболее существенных) объектов сети. При этом описание аналога также должно быть структурировано – в нем выделяются элементы, соответствующие основным элементам эскизного проекта на изготовление нового изделия, и затем производится сравнение.

Некоторые структурные элементы могут быть описаны с помощью семантической сети второго уровня представления. На втором уровне структурируются центральные понятия (ObjectType = “Centr”) с текстовым описанием (SemanticType = ”Name”), при необходимости контекстные (ObjectType = “Context”), такие как «технологический метод изготовления» и «материал», и объекты, представляющие собой группу элементов (ObjectType = “Group”).

Сформирована структура модели технологического процесса, которая представлена на рис. 5. Здесь объекты сети образуют пять уровней иерархии. Данная структура использована при разработке семантической модели технологического процесса, которая включает описание атрибутов узлов сети и указание видов связей (рис. 6).

Выделены следующие виды объектов:

– центральные понятия (типовой технологический процесс)

ObjectType = “Centr”[O1];

– группа понятий (технологические операции, оборудование, технические параметры и эксплуатационные данные оборудования, оснастка, приспособления, вспомогательные материалы)

ObjectType = “Group”[O2., O3, O4.3 - 7];

– контекст (категория оборудования, разряд работы на данном оборудовании, значения технических параметров и эксплуатационных данных, нормы расхода вспомогательных материалов)

ObjectType = “Context”[O4.1, O4.2, O5.1 - 3].

Выводы

На современном этапе развития информационных технологий управления предприятием и прогнозирования его развития проблемы интеграции информационных ресурсов занимают одно из центральных мест. Частью концепции интегрированного управления информационными ресурсами является разработка ряда информационных моделей основных производственных объектов и их элементов. На этой основе осуществляется поддержка принимаемых планово-оптимизационных решений. Поэтому тема данной статьи является актуальной.

Предложенная в статье комплексная технологическая модель является основой для структуризации информации при разработке базы данных архива технологических решений и базы знаний информационной поддержки процедур поиска в архиве ТП-аналогов.

Список литературы: 1. *Зильбербург, Л.И.* Реинжиниринг и автоматизация технологической подготовки производства в машиностроении [Текст] / Л.И. Зильбербург, В.И. Молочник, Е.И. Яблочников. – СПб.: Компьютербург, 2003. – 152 с. 2. *Серебряков, А.А.* Библиотека технологических решений как инструмент автоматизации технологической подготовки производства [Текст] / А.А. Серебряков // САПР и графика. – 2008. – № 5. – С. 70-74. 3. *Куликов, Д.Д.* Моделирование объектов и процессов в автоматизированных системах технологической подготовки производства в машино- и приборостроении (методы и инструментальные средства): автореф. дис. ... д-ра техн. наук / Д.Д. Куликов. – СПб. ин-т точной механики и оптики. – СПб., 1992. – 38 с. 4. *Варшавский, П.Р.* Поиск решения на основе структурной аналогии для интеллектуальных систем поддержки принятия решений [Текст] / П.Р. Варшавский, А.П. Еремеев. // Известия РАН. Теория и системы управления. – № 1. – 2005. – С. 97-109. 5. *Сироджа, И.Б.* Состояние проблемы знаниеориентированной поддержки принятия технологических решений в условиях современной автоматизации труда технолога [Текст] / И.Б.

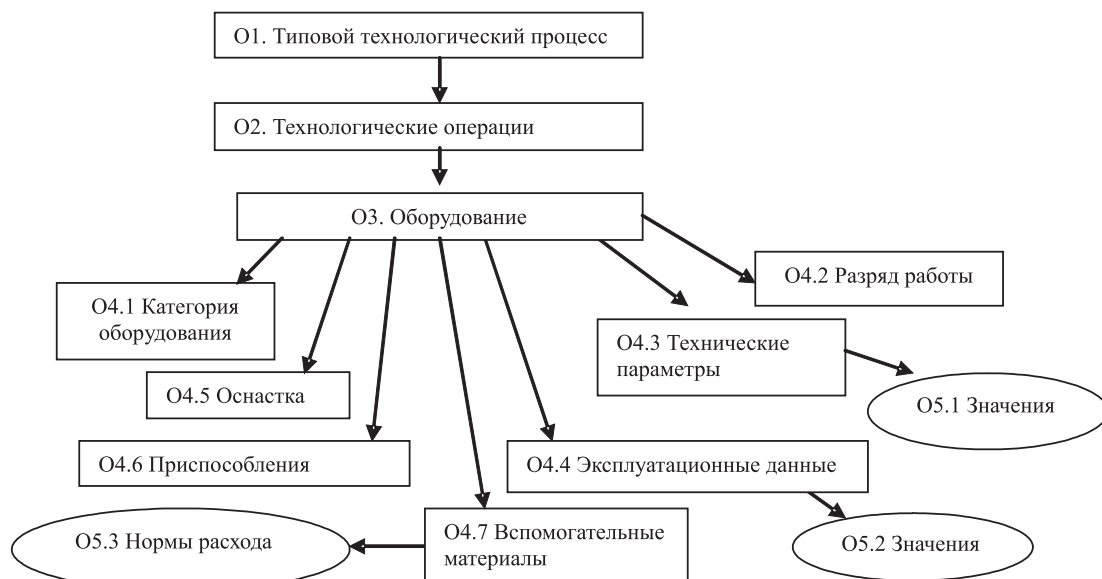


Рис. 5. Структурные элементы семантической сети технологического процесса

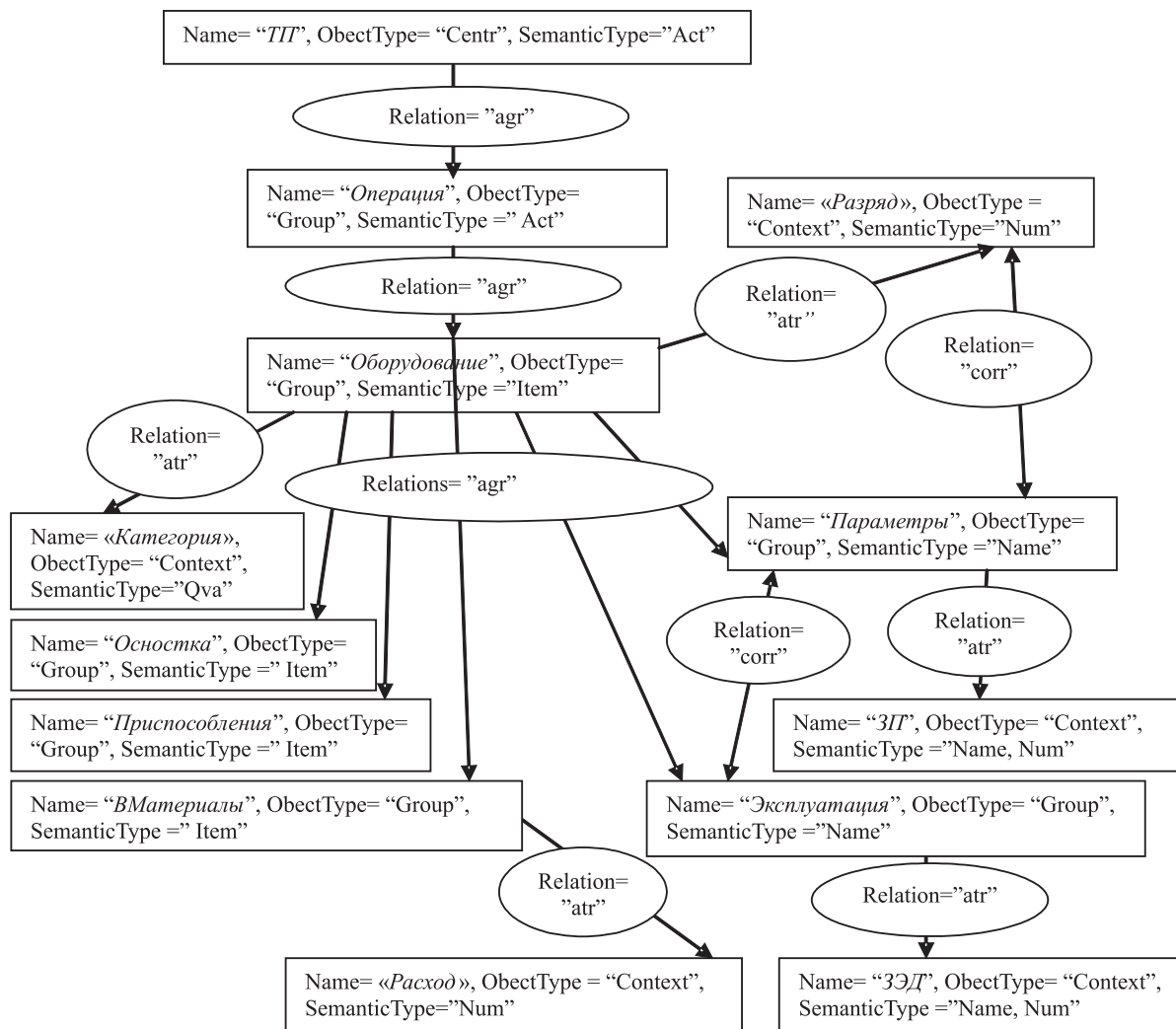


Рис. 6. Семантическая модель описания основных объектов технологического процесса

Сироджа, Л.С. Молодых // Открытые информационные технологии. – 2008 – №39. – С. 253 – 259. 6. Некрасов, А.Б. Семантическая модель онтологии технической документации для разработки научно-технической продукции [Текст] / А.Б. Некрасов, Н.А. Соколова, Д.Э. Лысенко // Системы управления, навигации та зв'язку. Зб. Наук. праць Центрального науково-дослідницького інституту навігації і управління. – Київ: ЦНДІНУ, 2008. – Вип. 2(6). – С. 96-101. 7. Единая система конструкторской документации. Виды и комплектность конструкторских документов: ГОСТ 2.102-68. Єдина система конструкторської документації. Основні положення. Довідник: Укр. та рос. мовами [Текст] // За заг. ред. В.Л. Иванова. – Львів: НТЦ "Леонорм-стандарт", 2001. – 272 с. 8. Единая система конструкторской документации. Эскизный проект: ГОСТ 2.119-73. Єдина система конструкторської документації. Основні положення. Довідник: Укр. та рос. мовами [Текст] // За заг. ред. В.Л. Иванова. – Львів: НТЦ "Леонорм-стандарт", 2001. – 272 с.

Поступила в редколлегию 9.06.2010.

УДК 004.89

Розробка комплексної технологічної моделі виробу у вигляді семантичної мережі / А.С. Котов // Біоніка інтелекту: наук.-техн. журнал. – 2010. – № 3 (74). – С. 76–82.

На сучасному етапі розвитку інформаційних технологій управління підприємством і прогнозування його

розвитку проблеми інтеграції інформаційних ресурсів займають одне з центральних місць.

У статті розглядаються питання побудови моделей семантичної мережі для опису та багатокритеріальної оцінки оптимізації варіантів технологічних процесів на передпроектній стадії прийняття стратегічних рішень. Запропонована модель є основою для структуризації інформації при розробці бази даних архіву технологічних рішень і бази знань інформаційної підтримки процедур пошуку в архіві ТП-аналогів.

Л. 6. Бібліогр.: 8 найм.

UDC 004.89

Development of complex technological model of a product as the semantic network / A.S. Kotov // Bionics of Intelligence: Sci. Mag. – 2010. – № 3 (74). – P. 00–00.

At the current stage of development of information technologies of enterprise control and prediction of its development, a problem of information resources integration occupies one of central places.

In article questions of models construction of the semantic network for the description and a multicriteria estimation of optimization of technological processes of variants at a pre-design stage of acceptance of strategic solutions are considered. The model offered in article is a basis for structurization of the information at development database of the of technological solutions archive and knowledge bases of information support of procedures of search in archive of TП-clones.

Fig. 6. Ref.: 8 items.

УДК 004.93



АГЕНТНЫЕ ТЕХНОЛОГИИ В СИСТЕМАХ ГРУППОВОЙ ВИДЕОКОММУНИКАЦИИ

Ан.А. Олейник¹, А.А. Ключкова², Ал.А. Олейник³

¹ Запорожский национальный технический университет, г. Запорожье, Украина, olejnik@zntu.edu.ua

² Запорожский национальный технический университет, г. Запорожье, Украина

³ Запорожский национальный технический университет, г. Запорожье, Украина, lexaol@gmail.com

Рассмотрены программные средства групповой видеоконференции. Проанализированы системы Skype, Gizmo, Google Talk. Предложена мультиагентная модель управления потоками данных при видеоконференции. Разработана мультиагентная программная система видеоконференции, которая не требует доступа в интернет и позволяет работать в локальных корпоративных сетях и выполнять передачу видеоданных, текстовой информации, а также битовых данных.

АГЕНТ, ВИДЕОКОММУНИКАЦИЯ, КОРПОРАТИВНАЯ СЕТЬ, ИНТЕРНЕТ, ПЕРЕДАЧА ДАННЫХ

Введение

Технологии видеоконференцсвязи (ВКС) предоставляют компаниям, имеющим территориально удаленные офисы, мощные инструменты повышения эффективности делового общения. Видеоконференцсвязь дает собеседникам, находящимся на значительном расстоянии друг от друга, возможность интерактивного общения и обмена документами, практически равную по результативности личной встрече. Системы ВКС позволяют решить такие задачи как увеличение скорости принятия решений, уменьшение потерь времени ключевых сотрудников, сокращение командировочных и накладных расходов [1–4].

1. Постановка задачи

Каждая отдельно взятая организация обладает своей специфической структурой, характеризуется особенностью деятельности, что вызывает необходимость разработки методов, моделей и программных средств, позволяющих решать специфические задачи, возникающие при построении систем организации видеоконференций. Помимо того, необходимо достигать максимальной эффективности организации взаимосвязи, что может быть обеспечено путём организации видеосвязи в виде мультиагентной системы.

Целью данной работы является разработка мультиагентных программных средств многопоточной видеоконференции корпоративного предприятия, позволяющие выполнять передачу видеоданных, текстовой информации, а также битовых данных.

Для достижения поставленной цели необходимо решить следующие задачи:

- анализ существующих средств обеспечения групповой видеоконференции с целью выделения основных характеристик, которыми должна обладать разрабатываемая система;

- разработка мультиагентной концепции организации управления потоками данных при видео-

коммуникации для обеспечения соответствия качества передаваемых данных текущей пропускной способности канала;

- разработка программной системы видеоконференции.

2. Анализ известных систем видеоконференции

Для выделения основных возможностей, которыми должна обладать разрабатываемая система, необходимо рассмотреть существующие и наиболее известные системы, обеспечивающие видеоконференцию: Skype, Gizmo, Google Talk.

Основной особенностью программной системы Skype является то, что для передачи голоса используются идеи пиринговых сетей, широко используемых для обмена файлами (Napster, KaZaA, Gnutella и другие). Основа успеха Skype определяется возможностями, предоставляемым данным сервисом:

- передача речи, мгновенных сообщений и файлов;

- высокое качество речи;

- проведение конференций;

- связь с традиционной телефонной сетью (через сервисы SkypeIn и SkypeOut);

- дополнительные возможности (интеграция с MS Outlook, Firefox, работа с сетями wi-fi хот-спотов, возможности «корпоративного» управления счетами и другие);

- шифрация всех передаваемых данных (алгоритмы AES, RC4);

- отличная работа из-за NAT и бранд-мауэров;

- поддержка различных платформ;

- наличие API для разработки собственных расширений программы;

- множество различных языков в интерфейсе клиента Skype.

Сеть Skype состоит из программных клиентов Skype, работающих на компьютерах пользователей. В зависимости от условий работы один и тот же клиент может выступать либо как «обычный узел», либо как «суперузел» сети [1]. У последних есть по меньшей мере два характерных признака:

- они всегда имеют реальный интернет-адрес;
- они принимают непосредственное участие в поддержке работы сети Skype.

Благодаря наличию «суперузлов», Skype относят к децентрализованным сетям. Поиск абонентов, отслеживание их текущего состояния (активен/не активен), обмен мгновенными сообщениями и файлами производится с участием «суперузлов». Самой важной функцией суперузлов является то, что каждый из них может выступать посредником при передаче голосового трафика между другими клиентами сети Skype, находящимися за серверами NAT. Теоретически подход Skype позволяет бороться абсолютно со всеми видами NAT. Однако такая борьба производится за счет ресурсов абонентов, поскольку физически «суперузлы» находятся на компьютерах обычных пользователей сети. Так, один день работы Skype в режиме «суперузла» может стоить владельцу компьютера до 800 Мб трафика других пользователей Skype [1, 3]. Ситуация усугубляется тем, что на данный момент не существует официального способа запретить клиенту Skype становиться «суперузлом».

Таким образом, самым большим недостатком Skype является то, что он использует ресурсы компьютеров своих пользователей для потребностей всей сети в целом. В результате, некоторые образовательные учреждения США и Европы официально запретили своим работникам и студентам использовать Skype.

Помимо того, в основе программы лежат закрытые протоколы. Никто до конца точно не знает, что на самом деле делает Skype, какие данные и куда он может передавать (посредством тех же «суперузлов»). Исходный код клиента также закрыт, поэтому остается лишь «верить», что он не содержит Spyware или других подобных средств.

Из-за использования закрытого протокола интеграция Skype с другими сетями на данный момент крайне затруднительна.

В отличие от Skype, архитектура сети SipPhone, используемой программным клиентом Gizmo [2, 4], является более традиционной.

В центре сети находятся серверы сети SipPhone. Благодаря использованию открытых протоколов, в сети возможна работа не только клиентов Gizmo Project, но и любых других, поддерживающих протокол SIP, в том числе и аппаратных (список рекомендуемого оборудования можно найти на сайте проекта).

Основными возможностями программного продукта Gizmo являются:

- передача голоса и мгновенных сообщений;
- связь с традиционной телефонной сетью через службы Call In и Call Out;
- связь с абонентами Gizmo через выделенные городские номера доступа (набор добавочного SIP-номера);

- проведение конференций, в том числе и с абонентами телефонных сетей общего пользования (ТФОП);

- бесплатная голосовая почта;
- возможность записи разговоров;
- отображение участников разговора на карте мира;
- использование эмодиконов и звуковых вставок;
- поддержка открытых стандартов (SIP, Jabber);
- связь с другими VoIP-сетями, использующими SIP;
- возможность работы с использованием аппаратных SIP-шлюзов и IP-телефонов.

Всем участникам сети SipPhone доступны бесплатные голосовые звонки абонентам других SIP-сетей (несколько VoIP-сетей университетов, VoIP-сети проекта SipBroker по всему миру). Кроме того, относительно просто должна достигаться и организация связи с обычными телефонными сетями, что достигается за счёт использования протокола SIP.

Для передачи мгновенных сообщений используется открытый протокол XMPP, на основе которого работают сети Jabber. На данный момент Gizmo Project является участником IM-конфедерации [3], что автоматически означает обмен сообщениями с абонентами других jabber-сетей.

Голосовая служба Google Talk (Gtalk), разработанная компанией Google, предоставляет такие основные возможности [4]:

- передача голоса и мгновенных сообщений;
- отображение качества разговора в реальном времени;
- уведомления о непрочитанных письмах в почтовом ящике;
- быстрое открытие ящика Gmail.

С технической точки зрения выбор Google особенно интересен. Если Skype использует собственный закрытый протокол, а Gizmo Project – стандартный для IP-телефонии SIP, то в основе Gtalk лежит протокол Jabber/XMPP, широко используемый Open Source-сообществом для передачи мгновенных сообщений. Благодаря этому, совместно с Google Talk гипотетически можно использовать абсолютно любой Jabber-клиент. Список гарантированно работающих альтернатив можно найти на соответствующем разделе сайта Google [6]. Клиенты будут работать с сервером Google, поэтому создавать учетную запись в Gmail все равно придется.

Исходя из вышеописанных характеристик и особенностей систем видеосвязи, можно выделить их основные особенности, представленные в табл. 1.

Таким образом, рассмотренные программные средства позволяют обеспечивать видеокommunikацию пользователей глобальной компьютерной сети Интернет. Однако данные средства работают

в рамках сети Интернет, не позволяя при этом осуществлять взаимодействие в локальных сетях. Это приводит к необходимости наличия доступа к Интернету во всех подразделениях предприятия, что часто на практике является нецелесообразным в силу политики безопасности компании и возможности снижения эффективности деятельности рядовых сотрудников.

Таблица 1

Сравнительный анализ возможностей систем видеокommunikации

Возможности	Skype	Gizmo Project	Google Talk
Возможность проведения конференции	Да	Да	Нет
Мгновенные сообщения	Да	Да	Да
Голосовая почта	Да	Да	Нет
Передача файлов	Да	Нет	Нет
Возможность работы с другими сетями	Нет	Да	Нет
Прямые звонки между абонентами	Нет	Да	Да
Используемый протокол	Собственный	SIP	XMP
Возможность контроля качества речи в реальном времени	Нет	Да	Да

В связи с этим требуется разработать программные средства многопоточной видеокommunikации, которые должны обеспечивать взаимодействие сотрудников предприятия вне зависимости от выхода в Интернет, в рамках корпоративной компьютерной сети.

Исходя из проведенного анализа существующих программных средств коммуникации, выявлены основные возможности, которыми должна обладать система, обеспечивающая видеокommunikацию.

1. Возможность проведения конференций: возможность создания одним пользователем видеотрансляции, к которой может подключаться множество других пользователей. В контексте разрабатываемой системы конференции, а соответственно, и вещание должны создать только сотрудники определённого иерархического уровня, чтобы исключить возможность неэффективного использования разрабатываемой системы.

2. Мгновенные сообщения: мгновенные сообщения в разрабатываемой системе должны служить средством обсуждения проводимой конференции.

3. Голосовая почта: в рамках разрабатываемой системы данная функция обретает свою специфику — сотрудник, создающий трансляцию, должен иметь возможность записи трансляции на сервер, чтобы сотрудники, которые должны наблюдать

трансляцию, но по тем или иным причинам отсутствовали при вещании, могли просмотреть её снова.

4. Передача файлов: в процессе видеоконференции может возникнуть необходимость передачи некоторой информации сотрудникам в виде отдельных файлов, при этом передаваемые файлы должны храниться некоторое время на сервере, чтобы сотрудники, не смотревшие трансляцию, со временем могли их также получить. В связи с этим необходимо обеспечить передачу файлов на сервер и закрепление переданных файлов к соответствующей трансляции.

5. Возможность работы с другими сетями: данная функция не является необходимой в контексте разрабатываемой системы, поскольку разрабатываемые программные средства являются единственной системой видеокommunikации организации. Однако в дальнейшем требуется обеспечить возможность расширения и адаптивности системы: должна быть возможность расширения системы в рамках протокола SIP.

6. Прямые звонки между абонентами: с учётом особенностей разрабатываемой системы (видеотрансляции могут создавать сотрудники определённого иерархического уровня) все трансляции требуется вести через сервер. Поскольку необходима их регистрация, то данная функция не должна быть учтена при разработке средств многопоточной видеокommunikации.

7. Используемый протокол: так как разрабатываемые программные средства должны работать в рамках любого браузера, то в соответствии со средствами веб-телефонии необходимо использовать протокол RTMP либо RTMFP. Однако в связи с лучшим развитием, стабильностью работы, а также существующими серверами, поддерживающими протокол RTMP, предлагается использовать протокол RTMP. Также возникает вопрос выбора медиа-сервера: Adobe Flash Media Server, Red5 или Wowza. Исходя из анализа, единственным бесплатным медиасервером является — Red5. При этом на основании данного медиасервера разработаны успешно и стабильно функционирующие веб-сервисы (наиболее известным из которых является Facebook — мировой лидер среди социальных сетей). В пользу медиасервера Red5 также следует отметить, что в настоящее время активно развивается проект Red5phone, позволяющий объединять системы, разработанные на базе Red5 с SIP-телефонами (что также является важным моментом и должно быть учтено в соответствии с возможностью 5: возможность работы с другими сетями). Таким образом, в качестве медиасервера предлагается использовать Red5.

8. Возможность контроля качества речи в реальном времени: поскольку пропускная способность

отдельных сегментов корпоративной компьютерной сети может отличаться, данная функция должна быть учтена при разработке, поскольку это позволяет снизить передаваемый трафик.

Таким образом, на основании проведенного анализа средств, методов и технологий видеокommunikации, выделены основные возможности, которыми должна обладать разрабатываемая система многопоточной видеокommunikации корпоративного предприятия, а также выделены основные программные и сетевые решения, которые должна использовать разрабатываемая система: протокол RTMP и медиасервер Red5.

Исходя из проведенного анализа существующих программных средств коммуникации, выявлены основные возможности, которыми должна обладать система, обеспечивающая видеокommunikацию.

3. Мультиагентные системы

При агентно-ориентированном программировании достигается наибольшая инкапсуляция – инкапсуляция кода, переменных и инициализации, за счёт чего достигается ряд преимуществ по сравнению с традиционными подходами. Например, если система состоит из ста агентов и для каждого агента необходимо предусмотреть десять возможных вариаций поведения, то в случае агентно-ориентированного программирования необходимо реализовывать тысячу поведений, в то время как при традиционных подходах необходимо реализовывать 10^{100} различных поведений. Также применение программных агентов позволяет снизить два важнейших параметра, которые следует учитывать при разработке программного обеспечения – стоимость внедрения и обслуживания. К тому же использование мультиагентных систем позволяет достичь большей адаптивности и ошибкоустойчивости по сравнению с системами, разработанными на основании традиционных подходов.

Мультиагентную систему (МАС) в общем виде можно представить в виде множества из трёх элементов [5, 6]: *Агенты*, *Среда*, *Связи между Средой и Агентами*:

$$\text{МАС} = \langle \text{Агенты}, \text{Среда}, \text{Связи} \rangle$$

Каждый *Агент_i* описывается с помощью множества четырёх элементов: *Состояние_i*, *Вход_i*, *Выход_i*, *Процесс_i*:

$$\text{Агент}_i = \langle \text{Состояние}_i, \text{Вход}_i, \text{Выход}_i, \text{Процесс}_i \rangle,$$

где *Состояние_i* – это множество переменных, полностью определяющих агента; *Вход_i* и *Выход_i* – подмножества *Состояния_i*, элементы которых связаны со средой.

Процесс_i – автономный метод, выполняющий соответствующие изменения над *Состоянием*. “Автономный метод” подразумевает, что данный метод вызывается без какого-либо внешнего участия.

Средой является множество из двух элементов:

$$\text{Среда} = \langle \text{Состояние}_c, \text{Процесс}_c \rangle$$

Важной особенностью такого представления *Среды* является то, что *Среда* является сама по себе активной, поскольку она содержит свой собственный *Процесс_c*, который может изменять *Состояние_c*, независимо от входящих в эту *Среду* агентов.

Вход_i и *Выход_i* различных агентов связаны с элементами *Состояния_c*, но среда не различает, какие из элементов *Состояния_c* находятся в зависимости с ними. Отсутствие *Входа* и *Выхода* у *Среды* означает, что она, в отличие от агентов, является неограниченной. Если бы в *Среду* был добавлен *Вход* и *Выход*, то это означало бы, что среда ограничена и является фактически высокоуровневым агентом. Таким образом, может моделироваться взаимодействие высокоуровневых агентов с низкоуровневыми.

4. Введение мультиагентности для управления потоками данных при видеокommunikации

Пропускная способность каналов, используемых для передачи аудио- и видеопотоков при проведении видеокommunikации, может быть различной или меняться динамически в зависимости от различных внешних (проблемы маршрутизаторов, перегрузка сетевых каналов и другие) и внутренних (предоставление большей приоритетности другим соединениям, проблемы с сетевой картой и другие) условий. В связи с этим может возникать проблема адекватной передачи аудио- и видеопотоков данных. Данная проблема может решаться различными способами: путём повышения качества технического обеспечения (использование надёжных маршрутизаторов, каналов связи, повышение пропускной способности), а также путём введения программного контроля за потоками данных, на основании которого производится изменение качества сжатия аудио- и видеопотоков в соответствии с текущей пропускной способностью канала.

В данной работе предлагается использовать мультиагентную систему управления потоками видеокommunikации. В предложенной мультиагентной системе каждый агент отвечает за отдельный канал передачи потоков данных, который он анализирует, и на основании проводимого анализа принимает решение об изменении качества сжатия передаваемых данных.

Предложенную мультиагентную систему контроля потоков данных при видеокommunikации можно представить в следующем виде:

$$\text{SMAS} = \langle \text{SAgents}, \text{SEnvironment}, \{\text{SharedObjects}, \text{Streams}\} \rangle,$$

где *SAgents* – агенты, анализирующие потоки данных соответствующего подключения; *SEnvironment* – среда взаимодействия агентов; *SharedObjects* – общие объекты, за счёт которых обеспечивается

взаимодействие агентов; *Streams* – потоки передаваемых данных.

Отдельный агент *SAgent* может быть представлен в виде следующей модели:

$$SAgent = \langle \{Waiting, Analyzing\}, NetStream, \{CAudioStream, CVideoStream\}, \{AnalyzeTraffic, AnalyzeConnection, CompressAudioStream, CompressVideoStream\} \rangle,$$

где *Waiting* и *Analyzing* – описывают возможные состояния агента: ожидание и анализ потоков данных соответственно; *NetStream* – вход агента, сетевой поток данных; *CAudioStream* и *CVideoStream* – выходы агента: описывают сжатый аудио- и видеопоток данных в соответствии с текущими возможностями канала; *AnalyzeTraffic*, *AnalyzeConnection*, *CompressAudioStream*, *CompressVideoStream* – процессы агента, которые описывают следующие действия: анализ трафика, анализ соединения, сжатие аудиопотока и сжатие видеопотока, соответственно.

Среда взаимодействия агентов *SEnvironment* описывается следующей моделью:

$$SEnvironment = \langle \{Waiting, Receiving, Sending\}, \{CreateAgent, DestroyAgent\} \rangle,$$

где *Waiting*, *Receiving*, *Sending* – состояния среды: ожидание, получение, отправка данных, соответственно; *CreateAgent*, *DestroyAgent* – процессы среды: создание агента для нового канала, уничтожение агента для закрытого канала.

Таким образом, за счёт использования предложенной мультиагентной системы на серверной стороне разработанного программного обеспечения достигается эффективная передача данных, которая заключается в динамическом изменении качества сжатия передаваемых потоков, что позволяет обеспечить соответствие объёмов передаваемых данных и текущей пропускной способности канала.

5. Программная система групповой видеокommunikации

Разрабатываемая программная система состоит из серверной и клиентской части. Клиентская часть отвечает за создание и просмотр видеотрансляций. Серверная часть в свою очередь отвечает за проверку пользователя, регистрацию трансляции, передачу трансляции по протоколу RTMP, передачу дополнительной информации о трансляции, получение и передачу файлов, передачу мгновенных сообщений, хранение записанных видеотрансляций, хранение базы пользователей, которые могут создавать трансляции. Таким образом, программные средства многопоточной видеокommunikации корпоративного предприятия должны делиться на пакет клиентских средств и пакет серверных средств. В соответствии с особенностями выбранных программных средств, а также с фун-

кциями, которые должна выполнять серверная и клиентская часть разрабатываемого программного обеспечения, была разработана структурная схема системы.

Предлагаемая структурная схема системы представлена на рис. 1.



Рис. 1. Структурная схема разрабатываемой системы

Функционирование системы предполагает, что разрабатываемая система имеет архитектуру клиент-сервер. В связи с этим следует рассматривать функционирование как системы в целом, так и функционирование отдельных компонентов. В первую очередь следует рассмотреть функционирование системы в целом.

Основное взаимодействие в системе происходит между двумя основными обобщающими элементами: Flash-приложение и Red5-приложение.

В контексте рассматриваемой системы основными задачами, какие выполняет Flash-приложение, являются: регистрация на сервере видеотрансляции (*publish*), просмотр видеотрансляции (*play*), а также дополнительное взаимодействие с сервером. Решение этих задач при взаимодействии с сервером Red5, Red5-приложением и мультиагентной системой *SMAS* представлено на рис. 2.

На рис. 2 представлены основные потоки данных: данных с микрофона, данные на динамики и другие данные (текстовые и бинарные). Как видно из рисунка, данные с микрофона из Flash-приложения передаются по команде *Publish*, после чего эти данные передаются через мультиагентную систему *SMAS* во внешний Internet, в другие приложения, по команде *Play*. Соответственно, данные на динамики приходят по команде *Play*, при этом эти данные передаются на сервер по команде *Publish*. Также следует отметить, что данные с микрофона и с веб-камеры передаются непосредственно через сервер Red5, без дополнительной нагрузки на отдельное Red5-приложение, развёрнутое на сервере. На Red5-приложение передаются дополнительные, текстовые и бинарные, данные, обеспечивающие корректную работу системы, связанную с подключением к серверу, передачей текстовых сообщений и т.п.

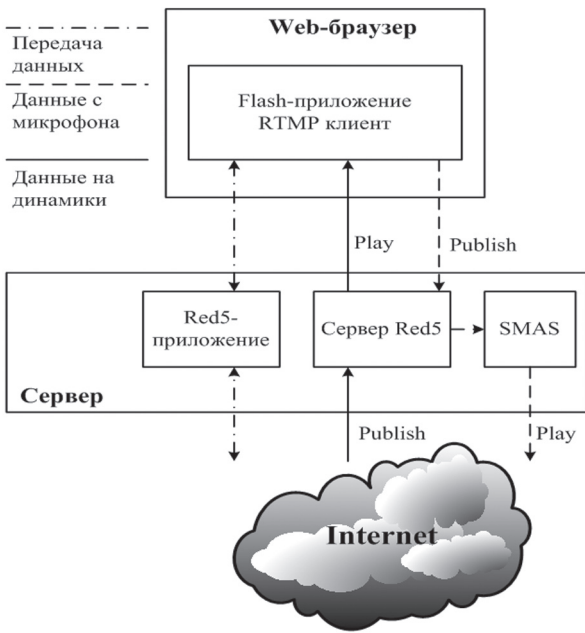


Рис. 2. Схема функционирования системы

Основными элементами, участвующими в функционировании серверной части являются: непосредственно сервер Red5, разработанное Red5-приложение MyVideoConf, СУБД MySql, а также создаваемый агент для анализа текущего соединения Agent. Диаграмма последовательности, отражающая функционирование данных элементов, находящихся на сервере, представлена на рис. 3.

Как видно из представленной диаграммы, работа начинается с соединения с сервером (connect). При соединении сервер вызывает метод appConnect разработанного приложения, который в свою очередь вызывает метод regTranslation, который обращается к БД, чтобы выяснить, обладает ли соединяющийся пользователь необходимыми правами. Исходя из привилегий пользователя, соединение разрешается или нет. Если соединение

произошло, то создаётся агент для анализа работы соответствующего соединения (analyze(stream)), а также удалённый пользователь может либо зарегистрировать трансляцию (publish), либо просмотреть (play). В случае создания трансляции сервер Red5 сам выполняет необходимые действия по регистрации трансляции и приёму потока аудио- и видеоданных. Когда пользователь хочет просмотреть трансляцию, то управление передаётся соответствующему агенту, который передаёт пользователю поток данных (stream) в сжатом виде. Когда пользователь отключается от сервера, происходит вызов disconnect. В этом случае Red5-приложение должно сообщить всем подключённым пользователям о том, что определённый пользователь отключился, что выполняется в методе remUserFromList, а также уничтожить соответствующего агента Agent.

На клиентской части может функционировать одно из двух Flash-приложений: Flash-приложение вещания видеотрансляции и Flash-приложение просмотра видеотрансляции.

Элементы клиентской части функционируют по схожей схеме: вначале создаётся соединение с удалённым сервером Red5. После того, как удалось подключиться к серверу, выполняется либо создание трансляции, либо подключение к уже существующей трансляции. Все эти действия выполняются вне зависимости от действий пользователя. После этого, когда происходит вещание или просмотр трансляции, пользователь может выполнять определённые действия в рамках Flash-приложения: может настраивать параметры аудио- и видеопотоков, передавать мгновенные сообщения другим пользователям, загружать файлы и т.п. После окончания сеанса работы с сервером, выполняется отключение от сервера.

Разработанная схема соответствует архитектуре клиент-сервер и позволяет выполнять все пос-

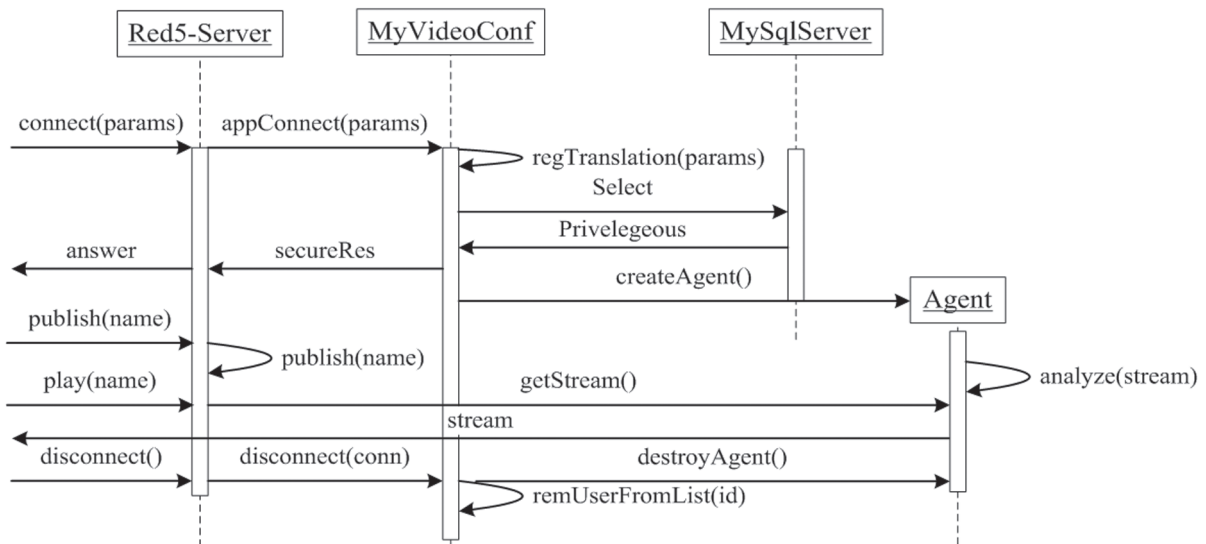


Рис. 3. Диаграмма последовательности функционирования серверной части

тавленные перед системой функции, основными из которых являются: создание и просмотр видеотрансляций.

Предложенные способы функционирования как системы в целом, так и частей отдельно, позволяют обеспечить логически верную и программно рациональную работу программных средств видеокommunikации корпоративного предприятия.

Выводы

С целью решения актуальной задачи обеспечения взаимодействия сотрудников корпоративного предприятия выполнен анализ существующих программных средств видеокommunikации, выделены их основные особенности, разработано программное обеспечение, свободное от недостатков существующих систем групповой коммуникации.

Научная новизна работы заключается в том, что предложена мультиагентная модель управления потоками данных при видеокommunikации, позволяющая выполнять анализ передаваемых потоков данных для каждого соединения, на основании которого производится динамическое сжатие данных в соответствии с текущими возможностями соединения.

Практическая ценность работы заключается в том, что разработана программная система, позволяющая осуществлять взаимодействие между различными сотрудниками корпоративного предприятия посредством конференцсвязи, передачи аудио- и битовой информации, а также мгновенных сообщений. Разработанная программная система не требует доступа в Интернет и позволяет работать в локальных корпоративных сетях, что понижает риск несанкционированной передачи и использования корпоративной информации в силу отсутствия необходимости использования внешних серверов. Кроме того, предложенная мультиагентная система управления потоками данных позволяет достичь более эффективного использования сетевых и технических ресурсов пользователей за счёт обеспечения соответствия объёма

передаваемых данных и текущей пропускной способности канала.

Список литературы: 1. *Goralski W.* IP Telephony / W. Goralski, M. C. Kolon. — Columbus: McGraw-Hill, 1999. — 416 p. 2. *Росляков, А. В.* IP-телефония [Текст] / А. В. Росляков, М. Ю. Самсонов, И. В. Шибаева. — М.: Радио и связь, 2003. — 252 с. 3. *Harte L.* Introduction to IP-telephony / L. Harte. — NC: Althos, 2003. — 80 p. 4. *Гольдштейн, Б. С.* IP-телефония [Текст] / Б. С. Гольдштейн, А. В. Пинчук, А. Л. Суховицкий. — М.: Эко-трендз, 2001. — 336 с. 5. *Ferber J.* Influences and reactions: A model of situated multiagent systems / J. Ferber, J. P. Muller // Proceedings of the 2nd International Conference on Multi-Agent Systems (ICMAS-96). — Menlo Park: IEEE Computer Society Press, 1996. — P. 72-79. 6. *Muller J. P.* The Design of Intelligent Agents: A Layered Approach (Lecture Notes in Computer Science) / J. P. Muller. — Berlin: Springer, 1996. — 227 p.

Поступила в редакцию 10.06.2010.

УДК 004.93

Агентні технології в системах групової відеокommunikacji. / А. О. Олійник, О. О. Клочкова, О. О. Олійник // Біоніка інтелекту: наук.-техн. журнал. — 2010. — № 3 (74). — С. 83–89.

Розглянуто програмні засоби групової відеокommunikacji. Проаналізовано системи Skype, Gizmo, Google Talk. Запропоновано мультиагентну модель управління потоками даних при відеокommunikacji. Розроблено мультиагентну програмну систему відеокommunikacji, яка не вимагає доступу до інтернет і дозволяє працювати в локальних корпоративних мережах та виконувати передачу відеоданих, текстової інформації, а також бітових даних.

Табл. 1. Лл. 3. Бібліогр.: 6 найм.

UDC 004.93

Agent technologies in systems of group video communication / A. O. Oliinik, O. O. Klochkova, O. O. Oliinik // Bionics of Intelligence: Sci. Mag. — 2010. — № 3 (74). — P. 83–89.

Software for group video communication are considered. Systems Skype, Gizmo, Google Talk are analysed. Multi-agent model of control of a data stream during video communications is offered. The multiagent program system of video communications which does not demand access to the Internet, allows to work in local corporate networks and to carry out transfer of the video data, the text information, and the bit data is developed.

Tab. 1. Fig. 3. Ref.: 6 items.

УДК 519.234.7



РОБАСТНЫЕ МЕТОДЫ ОЦЕНИВАНИЯ ЧИСЛОВЫХ ХАРАКТЕРИСТИК ВЫБОРКИ

В.Л. Шергин¹, Т.А. Мирошниченко²¹ ХНУРЕ, г. Харьков, Украина, shergin@kture.kharkov.ua² ХИНЭМ, г. Харьков, Украина.

Проведён анализ устойчивости числовых характеристик статистической выборки по отношению к выбросам. Предложено использовать робастные характеристики, основанные на порядковых статистиках – медиану выборки и медиану абсолютных отклонений. Проведено моделирование методом Монте-Карло, которое проиллюстрировало и подтвердило целесообразность предложенного подхода.

ЧИСЛОВЫЕ ХАРАКТЕРИСТИКИ ВЫБОРКИ, РОБАСТНОСТЬ, МЕДИАНА АБСОЛЮТНЫХ ОТКЛОНЕНИЙ

Введение

Одними из основных методов количественного анализа данных являются статистические методы. Основной предпосылкой применения классических методов математической статистики является гипотеза об однородности выборки. При использовании параметрических методов дополнительно предполагается, что элементы выборки следуют некоторому известному закону распределения. Так, если выборка получена из генеральной совокупности, подчиняющейся нормальному закону $N(\mu, \sigma)$, то эффективными оценками параметров сдвига (m) и масштаба (σ) являются среднеарифметическое значение и стандартное отклонение

$$m_x = \frac{1}{n} \sum_{i=1}^n x_i, \quad s_x = \sqrt{\frac{1}{n-1} \sum_{i=1}^n x_i^2} \quad (1)$$

соответственно, а для оценки меры корреляционной связи применяют выборочный коэффициент корреляции Пирсона:

$$r = \frac{\sum (x_i - m_x)(y_i - m_y)}{\sqrt{\sum (x_i - m_x)^2} \cdot \sqrt{\sum (y_i - m_y)^2}} = \frac{\sum (x_i - m_x)(y_i - m_y)}{(n-1) \cdot s_x \cdot s_y}. \quad (2)$$

В силу сложившейся традиции при проведении количественного анализа данных параметры сдвига, масштаба и силу связи оценивают именно по формулам (1)-(2). При этом гипотезы, лежащие в их основе, либо игнорируются, либо проверяются апостериорно с помощью критериев согласия, которые на коротких выборках имеют малую мощность. В результате получаемые оценки обладают низкой устойчивостью по отношению к выбросам в выборке и к отклонению закона распределения от нормального.

1. Анализ устойчивости числовых характеристик выборки

Традиционный подход, использующий оценки вида (1)-(2), имеет ряд существенных методологических недостатков. Во-первых, как отмечают специалисты в области прикладной статистики

[1], данные реальных измерений или наблюдений редко бывают распределены по нормальному закону. Гораздо чаще встречаются выборки с «тяжёлыми хвостами», то есть с существенно большими, чем при нормальном законе, частотами появления наблюдений, заметно отклоняющихся от средних. Для моделирования таких выборок целесообразно использовать другие законы распределения (например, распределение Стьюдента, Коши, логистическое), а для оценивания параметров (θ) применять метод максимального правдоподобия, согласно которому их находят из условия

$$\sum_i \frac{\partial \ln(P_i)}{\partial \theta} = 0, \quad (3)$$

где $P_i = f(x_i, \theta)$ – плотность вероятности закона распределения $f(x, \theta)$ при выборочном значении $x = x_i$.

Недостатками метода максимального правдоподобия являются как сложность получения ОМП-оценок (3), так и их жесткая «привязка» к выбранному закону, а также к гипотезе об однородности.

Во-вторых, выборки реальных данных зачастую бывают «засорены» как вследствие ошибок наблюдения, так и в силу внутренней неоднородности, то есть наличия наблюдений, подчиняющихся другому распределению. При этом наличие в выборке даже небольшого числа резко выделяющихся наблюдений (выбросов) способно фатально повлиять на результат статистического исследования, и значения оценок, подобных (1)-(3), могут потерять какую-либо практическую ценность.

В-третьих, существует ряд задач анализа данных, сама постановка которых отвергает гипотезу об однородности выборки. К таким задачам относятся, например, задачи классификации и распознавания, решаемые методами кластерного и дискриминантного анализа. При этом неотъемлемой частью большинства из этих методов являются процедуры стандартизации данных, то есть центрирования и нормирования:

$$z_i = \frac{x_i - m_x}{s_x}. \quad (4)$$

По инерции для этих целей чаще всего используют оценки (1), хотя их применение может быть обусловлено разве что простотой получения.

При проведении количественного анализа данных экономического характера исследователь сталкивается со всеми из перечисленных ситуаций: во-первых, реальные выборки могут быть коротки и обладать «тяжёлыми хвостами», во-вторых, выборки зачастую содержат не вполне достоверные элементы, и в-третьих, обычно нет оснований считать выборки однородными.

Таким образом, основой при проведении статистического анализа выборочных данных должны быть *робастные* методы оценивания. Под *робастностью* в статистике понимают нечувствительность к различным отклонениям и неоднородностям в выборке, связанным с теми или иными, в общем случае неизвестными, причинами [2].

Основными характеристиками робастности оценки являются точка разрыва *BP* (*breakdown point*) и функция влияния *IF* (*influence function*). *BP* определяется как максимальная доля в выборке экстремальных выбросов (равных бесконечности), при которых оценка параметра остаётся конечной. Легко видеть, что оценки (1) не являются робастными, так как для них $BP = 0$.

2. Построение робастных числовых характеристик выборки

К основным методам получения робастных оценок относятся: группировка данных, усечение выборки, взвешивание наблюдений. Усечение заключается в отбрасывании некоторой доли наблюдений. Очевидно, что если отбросить p процентов наблюдений, то $BP = p/100$. При группировке данных отдельные наблюдения не отбрасываются. Разбиение на интервалы не представляет особых трудностей и даёт весьма ощутимый результат в случае, когда выборка не очень короткая, причиной неоднородности являются немногочисленные резкие выбросы, а остальные элементы выборки действительно соответствуют выбранному закону распределения. Недостатками данного метода являются неработоспособность на коротких выборках и параметрический характер метода, то есть использование закона распределения. Так, если причиной отклонения выборки от нормального является «тяжелый хвост», то группировка данных не позволяет улучшить качество оценок параметров [3]. Подобными же свойствами в смысле робастности обладает и метод усечения.

Взвешивание наблюдений заключается в преобразовании данных выборки x_i с помощью некоторой весовой функции $\psi(z)$, где z_i является стандартизованным значением i -го элемента выборки, вычисляемым согласно (4). Целью взвешивания является установление барьера на влияние

резко выделяющихся (далеко отстоящих от предполагаемых значений параметров) наблюдений. В качестве весовой функции чаще всего используются биквадратная функция Тьюки $\psi(z) = \frac{z}{2}$, и функция гиперболического тангенса $\psi(z) = \frac{z}{1+z^2}$. Усечение также можно рассматривать как взвешивание усечённой линейной функцией.

При использовании взвешивания эффективными (и в то же время робастными) оценками параметров являются *M*-оценки, являющиеся обобщёнными *ОМП*-оценками:

$$\sum_i \psi(x_i, \theta) = 0. \quad (5)$$

Для получения *M*-оценок, то есть для решения системы уравнений (5), необходимо применять численные методы. В этом заключается недостаток взвешивания. При этом необходимо задать начальные приближения этих оценок. В общем случае они, конечно, могут быть любыми, однако желательно подбирать их как можно точнее, то есть ближе к искомым конечным значениям.

Наиболее логичным будет взять такие начальные приближения оценок параметров, которые сами являются робастными оценками. По этой причине рекомендуется [1], [4] использовать в качестве начального приближения для параметра сдвига *медиану* выборки

$$h_x = \text{median}(X). \quad (6)$$

Медиана является *наилучшей* из возможных оценок параметра сдвига по критерию *BP*, поскольку $BP(h_x) = \frac{\lfloor (n+1)/2 \rfloor - 1}{n} = \frac{1}{2} - \epsilon$, (где n – объём выборки), и, в то же время, эффективной *M*-оценкой. Для параметра масштаба такой робастной оценкой является значение, кратное медиане отклонений от медианы, называемое *MAD* – *median absolute deviation*:

$$MAD_x = k \cdot \text{median}|x_i - h_x|. \quad (7)$$

Для удобства практического применения оценок (7) коэффициент пропорциональности (k) обычно выбирают таким, чтобы оценка параметра масштаба для выборки X , следующей нормальному закону $N(\mu, \sigma)$, равнялась σ . В этом случае числовое значение коэффициента равняется

$$k = 1/u_{0,75} \approx 1,4826, \quad (8)$$

где $u_{0,75}$ – квантиль нормального распределения на уровне $p = 0,75$.

Дополнительным преимуществом использования медианной оценки является наличие у неё содержательной (а не только статистической) интерпретации, что особенно полезно при анализе экономических данных. Так, медиана выборки (то есть значение, находящееся посередине отсортиро-

ванной выборки) равна значению показателя для среднего наблюдения: расход ресурсов среднего по частоте встречаемости процесса; доход средней семьи; рентабельность среднего предприятия и т.п. Очевидно, что такая оценка не только робастна, но и гораздо более информативна, чем среднее арифметическое тех же величин.

3. Численное моделирование робастности числовых характеристик выборки

Для проверки робастности различных оценок сдвига и разброса было проведено численное моделирование методом Монте-Карло. Исследовались две оценки меры сдвига: среднее арифметическое значение и медиана (m_x и h_x), вычисляемые согласно (1) и (6) и обозначенные как mean и median соответственно, и две оценки меры разброса: стандартное отклонение s_x и медиана абсолютных отклонений MAD_x (1) и (7), обозначенные как stdev (standard deviation) и MAD.

Рассматривались пять различных законов распределения, а именно:

(А) – стандартный нормальный закон $N(0,1)$;

(В) – вероятностная смесь двух нормальных законов $N(0,1)$ и $N(0,3)$ с весами 0,9 и 0,1 соответственно;

(С) – логистическое распределение $L(0,1)$;

(D) – вероятностная смесь двух логистических законов $L(0,1)$ и $L(0,3)$ с весами 0,9 и 0,1 соответственно;

(Е) – распределение Коши $C(0,1)$.

Плотности логистического распределения $L(\alpha,\beta)$ и распределения Коши $C(\alpha,\beta)$ имеют вид

$$f(x) = \frac{e^{(x-\alpha)/\beta}}{(1 + e^{(x-\alpha)/\beta})^2} \quad \text{и} \quad f(x) = \frac{1}{\pi} \frac{\beta}{\beta^2 + (x - \alpha)^2}$$

соответственно.

Распределения (В) и (D) можно трактовать как “зашумлённые”, в которых 90% наблюдений следуют основному закону ($N(0,1)$ и $L(0,1)$ соответственно), а 10% наблюдений являются выбросами.

Для каждого из законов распределения методом Монте-Карло генерировалось по 100 выборок объёмом $n = 10, 20, 30, 50, 100, 200, 500, 1000, 2000, 5000$ каждая. Вычислялись оценки числовых характеристик выборки $\theta = \{m_x, h_x, s_x, MAD_x\}$ и дисперсии этих оценок $D(\theta_n, \theta^*)$ относительно математических ожиданий θ^* . Для всех пяти законов $M(m_x) = M(h_x) = 0$. Математические ожидания от s_x для рассматриваемых законов равны соответственно 1, $\sqrt{1,8}$, $\pi/\sqrt{3}$, $\pi\sqrt{0,6}$ и $+\infty$. Математические ожидания от MAD_x определяются из условий $k \cdot F^{-1}\left(\frac{3}{4}\right)$, где $F(x)$ – функция распределения, а $F^{-1}(p)$ – квантиль.

Поскольку $D(\theta_n, \theta^*)$ убывает обратно пропорционально n , то мерой качества полученных оценок может служить статистика [3]

$$T(n, \theta_n) = \sqrt{n \cdot D(\theta_n, \theta^*)}. \quad (9)$$

Для примера на рис. 1-2 приведены графики статистики (9), полученные для оценок меры сдвига в зависимости от объёма выборки (n) в случае порождающих распределений (А) и (В). По оси абсцисс используется логарифмический масштаб.

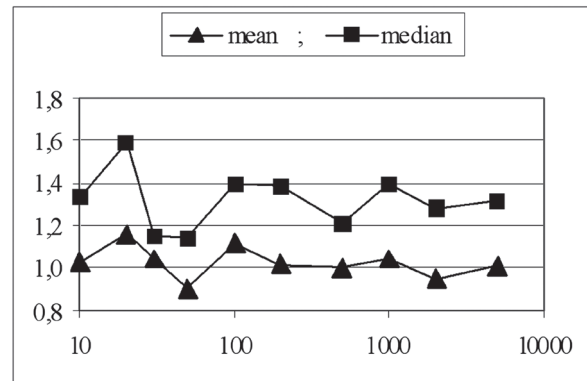


Рис. 1. Зависимость статистики (9) от размера выборки при порождающем распределении (А)

Итоговые значения этой статистики для параметров сдвига и разброса, полученные путём усреднения 100 выборок объёмом n каждая, приведены в табл.1-2 соответственно. Для сравнения в табл.1 представлены асимптотические оценки этой статистики при $n \rightarrow \infty$, полученные теоретически.

Анализ полученных результатов показывает, что робастные оценки параметров сдвига и разброса (6)-(7), как и следует из теории, уступают по эффективности классическим оценкам (1) в случае, когда выборка соответствует “чистым” нормальным или логистическим законам (А) и (С). В то же время, в случае наличия выбросов, т.е. при законах (В) и (D) робастные оценки оказываются существенно эффективнее. Кроме того, робастные оценки могут успешно применяться и для тех законов, для которых моменты распределения не определены, таких, как закон Коши.

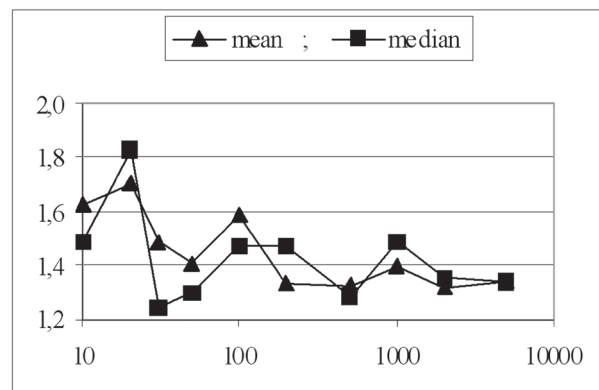


Рис. 2. Зависимость статистики (9) от размера выборки при порождающем распределении (В)

Таблица 1

Значения статистики (9) для оценок параметра сдвига

Порождающий закон распределения	Значения статистики (9) при $n = 20$		Значения статистики (9) при $n = 200$	
	mean	median	mean	median
A	1,159	1,593	1,015	1,383
B	1,699	1,824	1,330	1,475
C	1,541	1,757	1,667	1,994
D	1,805	1,859	2,376	2,105
E	297,528	1,468	415,320	1,577

Порождающий закон распределения	Значения статистики (9) при $n = 5000$		Асимптотические значения статистики (9)	
	mean	median	mean	median
A	1,005	1,316	1,000	1,253
B	1,338	1,339	1,342	1,343
C	1,770	2,051	1,814	2,000
D	2,490	2,218	2,433	2,143
E	1073,801	1,611	$+\infty$	1,571

Таблица 2

Значения статистики (9) для оценок параметра разброса

Порождающий закон распределения	Значения статистики (9) при $n = 20$		Значения статистики (9) при $n = 5000$	
	stdev	MAD	stdev	MAD
A	0,731	1,219	0,656	1,055
B	1,503	1,289	1,656	1,181
C	1,818	2,112	1,722	1,929
D	3,180	2,244	3,828	2,156
E	1323,9	2,734	75154,8	2,227

Выводы

Проведён анализ устойчивости числовых характеристик статистической выборки по отношению к выбросам. Выявлено, что среднее арифметическое значение и среднеквадратичное отклонение не являются устойчивыми. Предложено использовать устойчивые (робастные) характеристики сдвига и разброса – медиану выборки и медиану абсолютных отклонений (MAD). Проведено моделирова-

ние методом Монте-Карло, которое подтвердило целесообразность предложенного подхода.

Остаётся лишь удивляться, что такой полезный и широко применяемый для количественного анализа данных программный продукт как MS Excel, не содержит встроенной функции (6), позволившей бы получать робастные оценки параметра масштаба выборки. Более того, даже в таких пакетах для профессионального статистического анализа данных как Statistica и StatGraf хоть и реализованы методы робастного оценивания, но при стандартизации данных используются классические оценки (1), не являющиеся робастными.

Список литературы: 1. Орлов, А. И. Прикладная статистика [Текст] / А.И. Орлов. — М.: Экзамен, 2006. — 672 с. 2 Estimation of scale [Электронный ресурс]. — Режим доступа: http://en.wikipedia.org/wiki/Robust_statistics#Estimation_of_scale — 02.09.2010 г. — Загл. с экрана. 3. Кокс, Д. Теоретическая статистика [Текст] / Д. Кокс, Д. Хинкли. — М., Мир, 1978. — 640 с. 4. Гайдышев, И. Анализ и обработка данных [Текст]: специальный справочник / И. Гайдышев. — СПб: Питер, 2001. — 784 с.

Поступила в редколлегию 15.06.2010.

УДК 519.234.7

Робастні методи оцінювання числових характеристик вибірки / В.Л. Шергін, Т.О. Мірошніченко // Біоніка інтелекту: наук.-техн. журнал. — 2010. — № 3 (74). — С. 90–93.

Аналізується стійкість числових характеристик статистичної вибірки щодо викидів. Запропоновано використовувати робастні характеристики зсуву та розкиду вибірки — медіану вибірки та медіану абсолютних відхилень (MAD). Наведені результати чисельного моделювання, які підтверджують доцільність запропонованого підходу.

Л. 2. Бібліогр.: 4 найм.

UDK 519.234.7

Robust methods for sampling characteristics estimation // V.L. Shergin, T.A. Miroshnichenko // Bionics of Intelligence: Sci. Mag. — 2010. — № 3 (74). — P. 90–93.

Stability of numerical characteristics of the sampling concerning to the outliers is analyzed. Robust characteristics such as median and median absolute deviation (MAD) are proposed to use instead mean and standard deviation as the characteristics of the shift and scale. The presented results of the numerical simulation confirms usability and appropriateness of the median and MAD.

Fig. 2. Ref.: 4 items.

УДК 616-07:004.032.26



СИСТЕМА ПЕРВИЧНОЙ ДИАГНОСТИКИ МЕЛАНОМЫ КОЖИ

И.В. Новосельцев¹, Н.Г. Аксак²

¹ ХНУРЭ, . Харьков, Украина, grank@mail.ru;

² ХНУРЭ, Харьков, Украина, ahak@kture.kharkov.ua

Разработана система медицинской первичной диагностики меланомы кожи, которая позволяет повысить точность постановки диагноза рака кожи, проводить наблюдение за кожными образованиями удаленно. Предложен метод контроля изменения размеров кожного образования для наблюдения за процессом развития меланомы из невуса.

МЕЛАНОМА, ДИАГНОСТИКА КОЖНЫХ ОБРАЗОВАНИЙ, СИСТЕМА ДИАГНОСТИКИ

Введение

Кожа человека представляет собой естественный живой барьер, предохраняющий организм, и поэтому она более чем другие ткани подвержена воздействию неблагоприятных экологических и профессиональных факторов. Не случайно в общей структуре злокачественных новообразований опухоли кожи в большинстве стран становятся наиболее часто встречающимися онкологическими заболеваниями [1].

Сложная структура кожи включает в себя клетки различных тканей, каждая из которых может подвергнуться злокачественной трансформации. Этим объясняется большое количество различных типов опухолей, которые могут возникать в коже. Однако подавляющее большинство из них представлено тремя видами злокачественных новообразований: меланома, плоскоклеточный рак кожи и базальноклеточный рак.

Меланома является наиболее серьезным типом рака кожи, потому что в отличие от других видов рака кожи может распространяться на лимфатические узлы и другие части тела. Наиболее частому воздействию подвергается печень, легкие, кости и мозг. Так же, как существует много различных типов рака, существует много различных типов меланомы.

Несмотря на то, что доля меланомы кожи в структуре всех опухолевых заболеваний кожи составляет только 3 – 5%, именно это новообразование является главной причиной смерти больных с онкопатологией кожи. За последние 20 лет практически всеми исследователями отмечен устойчивый рост заболеваемости населения меланомой кожи во всех регионах мира. Так, в большинстве стран мира стандартизированные показатели заболеваемости ежегодно возрастают на 7% [2].

1. Существующие методы первичной диагностики меланомы кожи

Одним из методов первичной диагностики меланомы является позитронно-эмиссионная томография (ПЭТ). ПЭТ – это развивающийся диагностический и исследовательский метод ядерной

медицины. Позитронно-эмиссионная томография представляет собой радионуклидный томографический метод исследования внутренних органов. Этот метод основан на том, что в ткани вводят особый радиофармпрепарат. В его состав входят радионуклиды, для которых характерен так называемый позитронный бета-распад. После введения радиофармпрепарата проводится регистрация так называемых «гамма-квантов». Для опухолевых клеток характерен повышенный обмен веществ, это приводит к тому, что они быстрее и сильнее захватывают из крови введенный радиофармпрепарат. После того, как радиоактивное вещество оказывается в опухолевой клетке, начинается его распад. Во время распада образуются особые частицы (кванты), которые и регистрируются с помощью специальной аппаратуры. Данный метод позволяет определить область подозрительной активности раковых клеток [3]. Недостатком данного метода является бедная анатомическая информация изображений, что сильно затрудняет локализацию выявленных патологических очагов. Кроме того, данный метод может применяться лишь для выявления опухолей малого размера. ПЭТ – довольно дорогой метод диагностики, он имеется далеко не во всех медицинских центрах. Несмотря на высокую точность и диагностическую ценность ПЭТ является достаточно дорогим и не всегда доступным для рядового человека методом. Зачастую это связано с высокой стоимостью самого ПЭТ-томографа и сложностями с обеспечением радиофармпрепаратами, используемыми в процессе исследования.

Недавно появился новый перспективный неинвазивный метод диагностики меланомы кожи – эпилюминесцентная дерматоскопия, который позволяет повысить частоту выявления ранних форм заболевания и значительно сократить число случаев неоправданного хирургического удаления доброкачественных пигментных новообразований кожи [4]. Дерматоскопия позволяет выявить и оценить дифференциально-диагностические структурные признаки пигментных новообразований кожи в пределах эпидермиса и верхних слоев дермы, характерные для меланомы и доброкачественных

твенного невуса. По сравнению с другими методами клинического и инструментального обследования дерматоскопия, по данным ряда авторов, повышает качество диагностики меланомы кожи на 10—30% [4]. Однако сложность распознавания и интерпретации дерматоскопических признаков требует от врача, проводящего исследование, специальной подготовки и накопления определенного клинического опыта в онкодерматологии.

В дерматологическом отделении израильской больницы Хадасса недавно открылась лаборатория компьютерного наблюдения за родимыми пятнами. При обращении пациента проводится осмотр кожных покровов для выявления родинок, подлежащих наблюдению. Они фотографируются с помощью цифровой камеры, также проводится измерение диаметра, площади и цвета родимого пятна. Все эти данные заносятся в компьютерную базу данных и сохраняются. Последующие осмотры проводятся раз в полгода, и их результаты сравниваются с предыдущими измерениями. При необходимости производится биопсия.

2. Постановка задачи

Меланома очень часто возникает из пигментных невусов (рис. 1). Сам факт наличия у человека невусов представляет определенный риск возникновения меланомы кожи. Известно, что примерно от 30 до 66,5% меланом развиваются из предшествующего пигментного образования. Невус представляет собой пигментированное образование нейроэктодермального происхождения на коже, в состав которого входят невусные клетки, содержащие меланин. Далеко не все невусы имеют риск трансформации в меланому, фоном для ее возникновения служат лишь невусы особых гистологических форм.

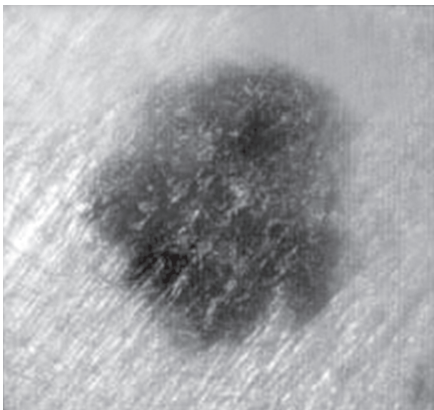


Рис. 1. Меланома из вырожденного невуса

В подавляющем большинстве случаев (примерно у 70% больных) меланома развивается на месте врожденного или приобретенного поражения кожи и только у 28—30% — на неизменной коже.

Меланома так же может возникать из лентиго — лентигомеланомы (рис. 2). Лентиго — плоское

равномерно пигментированное пятно от коричнево-бурого до черного цвета.



Рис. 2. Лентигомеланома

По сравнению с веснушками лентиго темнее и располагаются более редко; кроме того, их окраска не усиливается и число не возрастает под действием солнечного излучения. Лентиго относят к пигментным невусам. Лентигомеланомы составляют около 4-10% случаев. Появляются узелки в виде пятен темно-синего, темно-коричневого или светло-коричневого цвета диаметром 1,5-3 мм. Они поражают лицо, шею и другие открытые участки тела. Эти меланомы могут развиваться из доброкачественных меланотических веснушек Хатчинсона [6].

Несмотря на широкие исследования макро- и микроскопической морфологической структуры меланомы кожи, эффективность ее клинической диагностики остается невысокой. По данным исследований точность установления диагноза меланомы кожи врачами, специализирующимися в области дерматоонкологии, не превышает 65,6% [6]. Более того, на ранней стадии развития меланома может быть не замечена или определена как доброкачественное новообразование. Гипердиагностика меланомы кожи при визуальном обследовании достигает 30% [8].

Существующий уровень диагностики меланомы в странах СНГ сегодня трудно признать удовлетворительным, учитывая наружную локализацию опухоли и возможность её визуального распознавания. До 45,7% больных меланомой кожи поступают в онкологические стационары с далеко зашедшими метастатическими формами заболевания, что, естественно, не позволяет надеяться на успешное их излечение.

К сожалению, в настоящее время отечественными исследователями недостаточное внимание уделяется вопросам этиологии и патогенеза меланомы кожи. Вместе с тем, изучение именно этих ас-

пектов проблемы способствует определению групп риска возникновения этой опухоли, разработке адекватных программ скрининга меланомы кожи и созданию рациональной системы профилактики этого заболевания.

Таким образом, решение проблемы первичной диагностики меланомы кожи является актуальной задачей.

Целью работы является разработка диагностической системы, помогающей врачу в постановке диагноза рака кожи на основании опроса больного и снимков участков пораженной кожи, что позволяет осуществить своевременное лечение и производить диагностику удаленно.

Задача медицинской диагностики кожных образований состоит в определении возможных диагнозов больного на основе знаний предметной области и данных его обследования, к которым относят значения признаков (в моменты их наблюдения), значения анатомо-физиологических особенностей (постоянные во времени) и значения произошедших событий (в моменты, когда они происходили).

Для разработки медицинской системы первичной диагностики меланомы необходимо решить следующий ряд задач:

- разработать метод контроля изменения размеров кожного образования, позволяющего с большой точностью отслеживать изменения размеров наблюдаемых пигментных образований кожи для ранней диагностики меланомы;
- выделить основные симптомы, по которым можно определить тип кожного образования;
- разработать базу данных для регистрации медико-диагностической информации, включающей высококачественные цифровые фотографии кожных образований и набор клинических данных пациента, позволяющую накапливать материал для более точной постановки диагноза.

3. Диагностика меланомы

В связи с частотой возникновения меланом из доброкачественных пигментных образований необходимо знание клинических проявлений их малигнизации. К таковым относятся: рост невуса, его уплотнение или изъязвление; изменение окраски (усиление или ослабление); появление гиперемии или застойного ореола вокруг его основания; развитие лучистых разрастаний пигментного или непигментного характера вокруг первичного образования; возникновение экзофитного компонента на поверхности невуса; частые кровотечения; наличие увеличенных регионарных лимфатических узлов независимо от степени и характера изменений пигментного пятна; образование вблизи невуса пигментированных или непигментированных дочерних узелков-сателлитов.

Выявление нескольких из перечисленных симптомов позволяет клинически с большей долей вероятности установить правильный диагноз. При этом отдельные проявления активизации невуса (повышение пролиферации, дисплазии клеток) имеют разное диагностическое значение. Так, увеличение лимфоузлов и появление сателлитов указывают не только на озлокачествление, но и на генерализацию опухолевого процесса. Это является бесспорным доказательством наступившего перерождения. Наличие только одного признака малигнизации недостаточно для постановки диагноза, так как первые признаки озлокачествления нередко трудноотличимы от обычных воспалительных изменений.

На основе проведенного анализа выделены основные симптомы, по которым можно определить тип кожного образования (табл. 1).

Большую роль в постановке диагноза играет обработка снимков участков пораженной кожи. Имеется два снимка – первоначальный снимок заболевания, второй – сделанный в день обследования.

Таблица 1

Симптомы кожных образований

Симптом	Тип кожного образования		
	Невус	Меланома	Линтиго
А. Пятно появилось в детстве?	+	-	-
Б. Пятно появилось после того, как сошел загар?			+
В. Окрас неоднородный?		+	-
Г. Окрас с возрастом темнеет?	+		
Д. Окрас светло-коричневый?			+
Е. Присутствуют белые и розовые тона?	+		
Ж. Границы четкие?	+		
З. Размер меняется?	-	+	-
И. У родственников была меланома?	-	+	
К. Фототип кожи 1 или 2?		+	+
Л. Имеются ли кровотечения?	-	+	-

В результате анализа снимков определяется, прогрессирует ли болезнь или нет [9]. Производятся фотосъемки шаблона и области интереса кожного образования (КО). Основываясь на свойствах преобразования подобия, вычисляется реальная площадь наблюдаемого КО на основе рассчитанной площади цифрового изображения КО и шаблона, что позволяет идентифицировать прогрессирующие изменения в области интереса. На основании формулы (1), вычисляется реальная площадь наблюдаемого КО, которая считается базовой.

$$S_{об.р.} = \frac{S_{об.ф.} \cdot S_{ш.р.}}{S_{ш.ф.}}, \quad (1)$$

где $S_{об.р.}$ – реальная площадь наблюдаемого КО; $S_{ш.р.}$ – реальная площадь шаблона; $S_{об.ф.}$ – площадь КО на фотографии; $S_{ш.ф.}$ – площадь шаблона на фотографии.

Имея базовую площадь КО и соответствующую площадь КО при повторном наблюдении, вычисляется разница Δ_1 на 2-ом наблюдении:

$$\Delta_1 = S_{об.р.}^{(1)} - S_{об.р.}^{(2)}, \quad (2)$$

В случае положительной разницы можно сделать вывод, что КО в размерах уменьшилось, в случае отрицательной разницы – КО в размерах увеличилось, а в случае нулевой разницы – КО в размерах не изменилось.

Фотосъемку поверхности кожи и шаблона производят в одной плоскости, соблюдая перпендикулярность оптической оси к наблюдаемому объекту. Допустимый угол отклонения от перпендикуляра при фотографировании α определяется по формуле:

$$\alpha = \arcsin \frac{S_{ш.ф.}}{\pi \times r^2}, \quad (3)$$

где r – радиус шаблона.

На рис. 3 представлен график зависимости допустимого угла отклонения от перпендикуляра при фотографировании, на основании которого определяется погрешность вычисления площади (табл. 2).



Рис. 4. Схема функционирования системы первичной диагностики меланомы

Таблица 2

Погрешность вычислений

Угол отклонения	Абсолютная погрешность	Точность вычислений
80°	0,04	98,73%
74,3°	0,13	95,86%
68,6°	0,23	73,25%

Результаты обработки изображения являются одним из компонентов входного вектора поступающего на вход нейронной сети.

Функционирование системы первичной диагностики меланомы представлено на рис. 4.

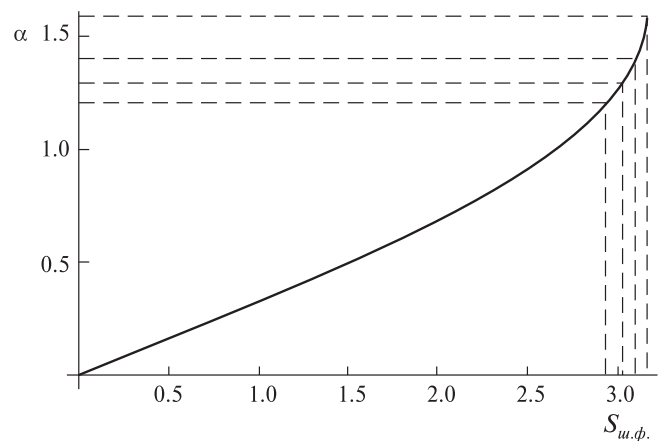


Рис. 3. График зависимости допустимого угла отклонения от перпендикуляра

В процессе консультации пациент отвечает на вопросы, представленные в таблице 1. На основании ответов формируется вектор опроса, который вместе с результатом обработки изображения поступает на вход нейронной сети. Состав и взаимосвязь компонентов нейросетевой экспертной системы медицинской диагностики S выражается как преобразование входных значений X в выходные величины Y :

$$S \subset X \times Y, \quad (4)$$

где $X(i) = [x_1(i), x_2(i), \dots, x_n(i)]$ – вектор опроса; $Y = \{y_1, y_2, \dots, y_Q\}$ – множество классов.

Таким образом, симптомы заболевания пациента на основе вектора опроса $X(i)$ классифицируются благодаря системе S при заданном числе эталонных образов Q , к которым относится заболевание.

Для выходных величин Y построено множество заданий, решение которых принадлежит множеству $D_X = D_{опрос} \cup D_{фото} \cup D_{НС}$, где $D_{опрос}$ – задача составления вектора опроса, $D_{фото}$ – задача обработки изображений кожного образования, которая сводится к определению изменения его размера, $D_{НС}$ – задача постановки диагноза с помощью нейронной сети. Отображение $T: X \rightarrow Y$ позволяет для каждого $X(i)$ найти такое $y_j \in Y$ ($j = \overline{1, Q, Q}$ – количество классов), которое является решением задачи D_X .

База данных (БД) цифровых фотографий представляет собой реляционную БД, позволяющую производить добавление, удаление, редактирование и поиск снимков кожных образований.

В табл. 3 представлены результаты проведенного тестирования. Сначала врачу предлагалось поставить диагноз самостоятельно, затем с использованием разработанной системы.

Тестирование показало, что разработанная система медицинской диагностики кожных образований позволяет повысить точность диагностики на 9.3%.

Выводы

Научная новизна полученных результатов заключается в следующем:

- предложенный метод контроля изменения размеров кожного образования для наблюдения за процессом развития рака кожи или меланомы из невуса (родинки), позволяет с большой точностью отслеживать изменения размеров наблюдаемых пигментных новообразований кожи, и помогает врачу при постановке диагноза;

- разработанная система медицинской диагностики меланомы позволяет контролировать увеличение размеров кожного новообразования и повысить точность диагностики на 9.3%;

Предлагаемая система позволит также проводить наблюдение за кожными образованиями на удалении и использовать ее для постановки диагноза в телемедицине.

Таблица 3

Результаты тестирования

Номинальный класс	Количество примеров	Результат диагностики					
		без использования системы			с использованием системы		
		поставленный диагноз			поставленный диагноз		
		лентиго	невус	меланома	лентиго	невус	меланома
лентиго	50	40	7	3	44	4	2
невус	50	2	41	7	3	45	2
меланома	50	0	10	40	0	5	45

Список литературы: 1. *Иванов, О.Л.* Кожные и венерические болезни [Текст] / О.Л. Иванов. – Москва: Шико, 2006. – 219 с. 2. *Анисимов, В.В.* Диагностика и профилактика меланом кожи: методические рекомендации [Текст] / В.В. Анисимов, А.С. Барчук. – Санкт-Петербург: Научно-исследовательский институт онкологии им. проф. Н.Н. Петрова Минздрава РФ, 2000. – 40 с. 3. *Барчук, А.С.* Комплексное лечение меланомы кожи с использованием физических методов [Текст] / А.С. Барчук, М.Л. Гельфонд, М.Л. Гершанович, М.А. Акимов. – Санкт-Петербург: Научно-исследовательский институт онкологии им. проф. Н.Н. Петрова Минздрава РФ. – 2003. – 15 с. 4. *G. Argenziano, H.P. Soyer.* Dermoscopy of pigmented skin lesions—a valuable tool for early diagnosis of melanoma. – *Lancet Oncol* 2001, vol 3. 443-449p. 5. *Соколов, Д.В.* Опыт автоматического распознавания меланомы кожи на основе цифровой эпилюминесцентной дерматоскопии [Текст] / Д.В. Соколов, Н.Н. Потехаев, Л.В. Демидов, Г.Н. Ворожцов. // Клиническая дерматология и венерология. – Москва: Медиа Сфера, 2010. – № 3. – С. 72-75. 6. *Курс пластической хирургии: Руководство для врачей.* Т.1: Общие вопросы. Кожа и ее поражения. Голова и шея [Текст] / Под ред. К.П. Пшениснова. Ярославль; Рыбинск: Изд-во ОАО Рыбинский дом печати, 2010. – 754 с. 7. *D.N. Chiacchio, S.H. Hirata, M.Y. Enokihara, N.S. Michalany, G. Fabbrocini, A. Tosti.* Dermatologists' accuracy in early diagnosis of melanoma. // *Arch Dermatol.* 2010 Apr, 146(4). P. 431-434. 8. *Залуцкий, И.В.* Меланома кожи: практи-

ческое пособие для врачей [Текст] / И.В. Залуцкий, С.З. Фрадкин. – Минск: Беларусь, 2000. – 215 с. 9. *Аксак, Н.Г.* Спосіб контролю зміни розмірів об'єкта за допомогою шаблону [Текст] / Н.Г. Аксак, М.В. Кушнарів, И.В. Новосельцев, А.Ю. Тыхун // Патент на винахід № 84380 А по МПК (2006) G01C 11/00 от 10.10.2008. Бюл. № 19.

Поступила в редколлегию 16.06.2010.

УДК 616-07:004.032.26

Система первинної діагностики меланоми шкіри / И.В. Новосельцев, Н.Г. Аксак // Біоніка інтелекту: наук.-техн. журнал. – 2010. – № 3 (74). – С. 94–98.

У статті вирішується актуальна задача первинної діагностики меланоми шкіри. Запропонована діагностична система шкірних утворень, що дозволяє підвищити точність постановки діагнозу раку шкіри і контролювати зміни розмірів області інтересу.

Л. 4. Бібліогр.: 9 найм.

UDK 616-07:004.032.26

Diagnostic system for malignant melanoma / I.V. Novoseltsev, N.G. Axak // Bionics of Intelligence: Sci. Mag. – 2010. – № 3 (74). – P. 94–98.

In article the diagnostic system which helps doctor to provide diagnosis at an early stage is considered. The new methods for monitoring of the size's changes in the skin disease were proposed. The diagnostic system allows increasing diagnostics accuracy of skin cancer on early stage.

Fig. 4. Ref.: 9 items.

УДК 615.47



ДИАГНОСТИЧЕСКИЕ ВОЗМОЖНОСТИ ЭЛЕКТРОМИОГРАФИЧЕСКОГО МЕТОДА ПРИ ИССЛЕДОВАНИИ ФУНКЦИИ НОСОВОГО КЛАПАНА

О.Г. Аврунин, Т.В. Жемчужкина, Т.В. Носова

ХНУРЭ, г. Харьков, Украина, gavrun@list.ru

Рассмотрены возможности электромиографического метода для функциональной диагностики носового клапана. Предлагаются методы анализа электромиографического сигнала мышц носового клапана. Обсуждаются возможности и перспективы этих методов для диагностики патологий носового клапана.

ЭМГ-СИГНАЛ, НОСОВОЙ КЛАПАН, СПЕКТРАЛЬНЫЙ АНАЛИЗ, ПЕРИОДОГРАММА, КО-ЭФФИЦИЕНТ УПРУГОСТИ, ГИСТОГРАММА

Введение

В настоящее время во всем мире наблюдается значительный рост заболеваний верхних дыхательных путей. По данным медицинской статистики последних пяти лет число различных видов ЛОР-патологии возросло в 3,2 раза. Наиболее частыми заболеваниями ЛОР органов являются острые и хронические воспаления слизистой оболочки дыхательного тракта, придаточных пазух носа, а также воспаления наружного, среднего и внутреннего уха. Причинами большинства ЛОР заболеваний являются осложнения инфекционных болезней, при которых поражаются верхние дыхательные пути (грипп, ангина, острые респираторные заболевания, корь, скарлатина и др.) [1]. Точная дифференциальная диагностика и распознавание сложных функциональных нарушений верхних дыхательных путей на современном этапе невозможны без применения сложной аппаратуры и методов интеллектуального анализа снимаемых биосигналов. Верхние дыхательные пути выполняют большое количество функций, основными из которых являются респираторная, согревательная, обонятельная и фильтрующая. Соответственно анатомически нос человека (см. рис. 1) представляет собой сложное структурное образование, в состав которого входят костные и хрящевые образования, мышцы и слизистая оболочка носовой полости. В преддверии носа располагается носовой клапан, регулирующий поток воздуха через носовые ходы.

1. Актуальность темы исследования

В настоящее время при реконструктивной хирургии полости носа (эндопластика) ведущую роль играет понимание функционирования носового клапана. Носовой клапан является структурой, регулирующей поток воздуха в носовой полости. Известные методы исследования носового клапана основаны на качественной оценке его функционирования. Методы интроскопической визуализации, такие как спиральная компьютер-

ная томография, позволяют оценить его анатомические размеры, но не позволяют проводить функциональную диагностику. Поэтому актуальной задачей является создание критериев его функционирования на основе измерения его физиологических показателей.

2. Постановка задачи

В данной работе проводится исследование функции носового клапана по данным электромиографии. На основе анализа электромиографических сигналов мышц носового клапана предлагается определить количественные критерии оценки его функционирования.



Рис. 1

3. Анатомические и физиологические особенности носового клапана

Под клапаном подразумевают мобильную структуру, регулиующую поток жидкости и газа. Нос человека имеет несколько клапанов. Пещеристая ткань носовых раковин (турбинальный клапан) и перегородка носа (септальный клапан) могут регулировать поток воздуха в полости носа. Таким образом, эти две тканевые структуры могут выполнять функцию клапана носа. Но обычно под носовым клапаном подразумевают пространство между каудальным краем верхнего латерального хряща и перегородкой носа (см. рис. 2).

Носовой клапан — треугольное щелевидное пространство между каудальным краем верхнего латерального хряща и перегородкой носа. Угол носового клапана измеряется в градусах и представляет собой угол между верхним латеральным хрящом и перегородкой носа. Эта область по форме напоминает каплю или перевернутый угол, щелевидная верхушка которого образует угол клапана носа, в норме равный 10 — 15 градусам. Область носового клапана, возможно, является первичным регулятором объема воздуха, проходящего через полость носа, в большей степени определяющая сопротивление воздушной струе [2].

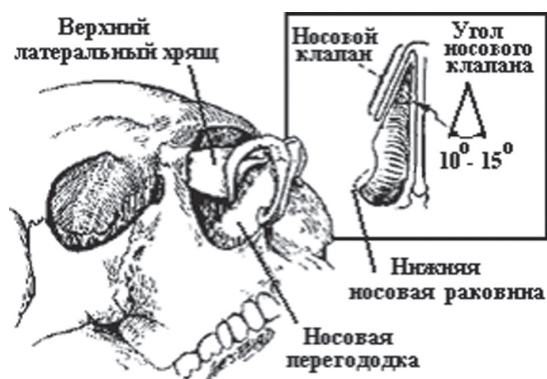


Рис. 2

4. Регистрация электромиографического сигнала в области носового клапана

Авторами предлагается исследование функций носового клапана по данным электромиографии. Электромиография (ЭМГ) — метод регистрации биотоков мышц. Он широко используется для диагностики нервно-мышечных заболеваний. Электромиограмма отражает электроактивность мышечных волокон.

Отводимые от мышцы токи действия дают сведения о функциональном состоянии мышцы и иннервирующего ее нерва.

Периферическую двигательную единицу составляет совокупность двигательной клетки переднего рога спинного мозга или ядра черепно-мозгового нерва, выходящего из нее аксона и иннервируемых им мышечных волокон.

С мышцы, находящейся в покое, потенциал действия не отводится. При незначительном мышечном сокращении уже появляются биоэлектрические волны с частотой колебаний 5—19 в секунду и амплитудой 100 мкВ. Поскольку отводимые потенциалы действия относятся не к единичному мышечному волокну, а ко всем мышечным волокнам, иннервируемым моторным нейроном переднего рога спинного мозга или ядром двигательного черепного нерва, при более сильном сокращении мышцы потенциалы действия становятся продолжительнее и интенсивнее, интерферируют и достигают 3000 мкВ

Электромиография является ценным методом исследования, позволяющим дифференцировать различные уровни поражения нервной системы. Электромиограммы имеют разную картину при двигательных нарушениях, обусловленных поражением центральной, периферической нервной систем и мышечного аппарата. Изменения биоэлектрической активности мышц зависят от локализации поражения нервной системы, тяжести и стадии болезни. Электромиография помогает диагностировать центральные, спинномозговые (сегментарные), невральные и мышечные двигательные нарушения. Электромиография позволяет также обнаружить типичные нарушения биоэлектрической активности в ранней стадии заболевания и наблюдать за динамикой процесса и эффективностью лечения.

Электромиография позволяет оценить активность мышц носового клапана, противодействующих его закрытию при разрежении во время форсированного дыхания. Используемая нами электромиографическая установка представляет собой поверхностный электрод, подключенный к микрофонному входу звуковой карты ПК, с питанием от интерфейса USB.

Методика исследования состоит в закреплении электрода на крыле носа обследуемого, регистрации электромиографического сигнала, поступающего со звуковой карты с помощью программы CoolEdit и последующей обработке и анализе электромиограмм в пакете MatLab. При этом электромиографический сигнал регистрируется в процессе форсированного дыхания.

5. Определение упругости мышц носового клапана

Для определения упругости мышц носового клапана необходимо проанализировать зависимость деформации мышц носового клапана от силы напряжения мышцы. Уровень напряжения мышц предлагается определять по ЭМГ-сигналу как максимальное значение напряжения в момент форсированного вдоха, при котором проводится измерение степени деформации мышц носового клапана по известной методике [2].

Известно, что форма огибающей ЭМГ-сигнала межреберных мышц соответствует объему воздуха, поступающего в легкие (рис. 3). На рис. 3 сверху вниз: электромиографический сигнал за два цикла дыхания с парастернальной межреберной мышцы собаки; огибающая ЭМГ, полученная с помощью модифицированного фильтра Бесселя с постоянной времени 100 мс; поток вдыхаемого воздуха [3]. Можно предположить, что форма огибающей электромиограммы мышц носового клапана будет соответствовать расходу воздуха, проходящего через носовую полость. Известно, что зависимость

силы мышечных сокращений от среднеквадратичных значений электромиографического сигнала носит практически линейный характер [4]. Следовательно, максимальное значение расхода воздуха будет соответствовать максимальному значению напряжения мышцы, т.е. максимуму огибающей ЭМГ-сигнала, а значит и максимальному значению амплитуды самого ЭМГ-сигнала.

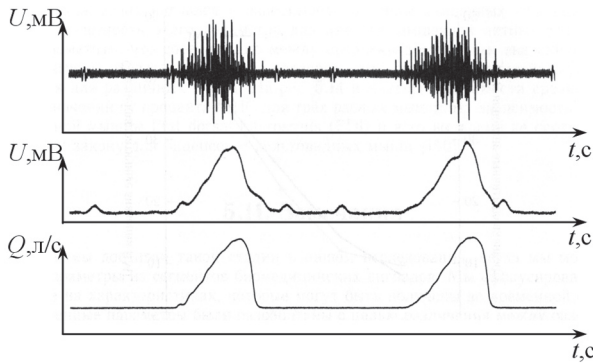


Рис. 3

Коэффициент упругости мышц носового клапана определим по закону Гука. При малых деформациях x сила упругости $F_{упр}$ пропорциональна деформации тела и направлена в сторону, противоположную направлению перемещения частиц тела при деформации:

$$F_x = F_{упр} = -kx, \quad (1)$$

где k – жесткость тела.

Согласно данным, приведенным в [3], сила упругости мышц прямо пропорциональна амплитуде U ЭМГ-сигнала:

$$-kx = f(U) = aU, \quad (2)$$

где a , Н/В – размерный коэффициент пропорциональности между амплитудой ЭМГ-сигнала и силой мышц носового клапана.

Отсюда относительный коэффициент упругости k' , В/м:

$$k' = \frac{k}{a} = -\frac{U}{x}. \quad (3)$$

Среднее значение коэффициента k' равно 0,1 В/м.

6. Анализ электромиографического сигнала в области носового клапана

Для анализа электромиографических сигналов в частотной области рассчитали периодограмму (оценку спектральной плотности мощности, полученную по N отсчетам одной реализации случайного процесса) по следующей формуле:

$$\hat{W}(w) = \frac{1}{Nf_{д}} \left| \sum_{k=0}^{N-1} x(k)e^{-j\omega kT} \right|^2,$$

где N – количество отсчетов электромиографического сигнала; f – частота дискретизации;

$x(k)$ – отсчеты электромиографического сигнала; T – интервал дискретизации.

На рассчитанной периодограмме выделили три частотных диапазона: очень низкие частоты (5–150 Гц), низкие частоты (150–300 Гц) и высокие частоты (300–500 Гц). Рассчитали суммарный вклад периодограммы в каждом из частотных диапазонов и построили столбцовые диаграммы соотношения спектральных мощностей по частотным диапазонам (рис. 4). На рис. 4 *а, б* приведены диаграммы для правой и левой мышц носового клапана соответственно для обследуемого с искривлением носовой перегородки, и как следствие деформацией носового клапана, на рис. 4 *в, г* – правой и левой мышц для обследуемого с нормально функционирующим носовым клапаном. Для удобства сравнительного анализа диаграммы распределения суммарного вклада периодограммы по частотным диапазонам приведены в одинаковом масштабе.

Аналогичные расчеты были проведены для спектрального состава ЭМГ-сигнала и рассчитаны мощности по тем же трем частотным диапазонам.

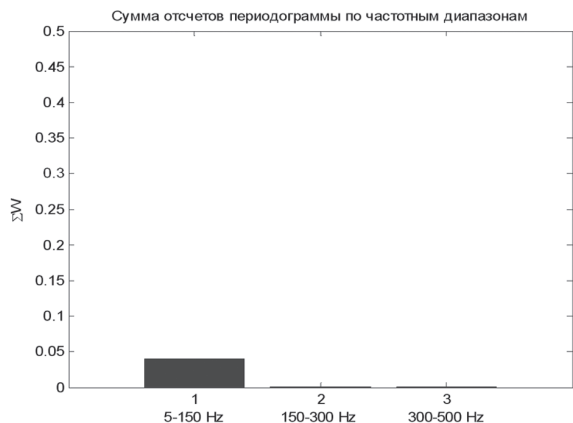
Спектральную функцию рассчитывали с помощью дискретного преобразования Фурье:

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

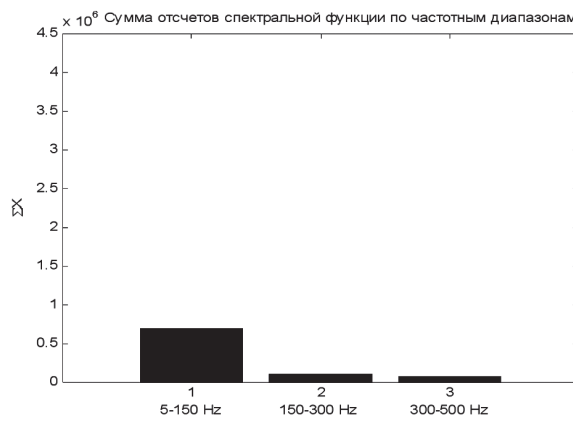
где N – количество отсчетов электромиографического сигнала; $x(k)$ – отсчеты электромиографического сигнала.

Результаты распределения суммарного вклада спектральной функции по трем частотным диапазонам приведены на рис. 5 в одинаковом масштабе. На рис. 5 *а, б* приведены диаграммы для правой и левой мышц носового клапана соответственно для обследуемого с искривлением носовой перегородки, на рис. 5 *в, г* – правой и левой мышц для обследуемого с нормально функционирующим носовым клапаном.

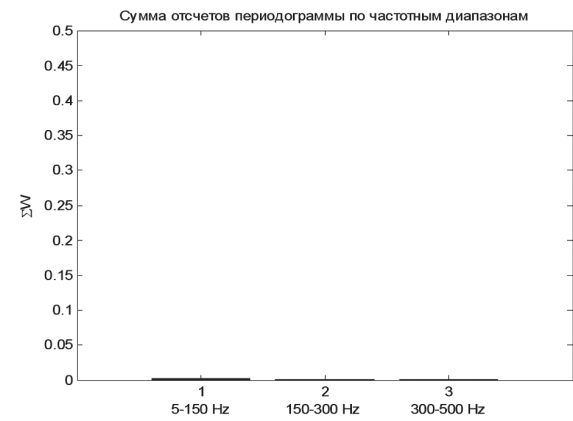
При анализе столбцовых диаграмм распределения суммарного вклада периодограммы и спектральной функции по трем частотным диапазонам было выявлено, что наибольшее отличие можно заметить во вкладе очень низких частот (столбец 1 (5–150 Гц) на (рис. 4–5) в электромиографический сигнал разных обследуемых. Для обследуемого с искривлением носовой перегородки (рис. 4 *а, б*, рис. 5 *а, б*) было отмечено значительное снижение (на порядок) амплитуды столбцов диаграммы по сравнению со значениями для обследуемых без патологии (рис. 4 *в, г*, рис. 5 *в, г*). Кроме того, нами было отмечено, что у некоторых обследуемых наблюдается значительная асимметричность функционирования правого и левого клапанов, которая выражается в разной амплитуде столбцов диаграммы для правого и левого клапанов (различие на порядок), что значительно всего проявляется в амплитуде первого столбца (5–150 Гц) (рис. 4–5).



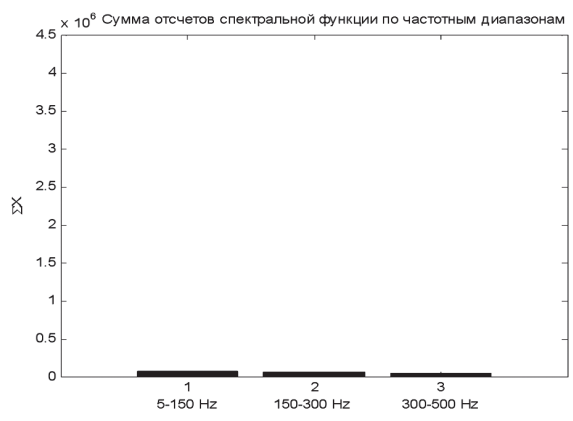
a



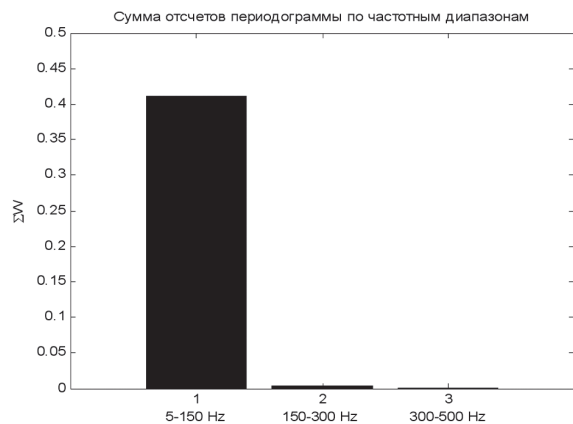
a



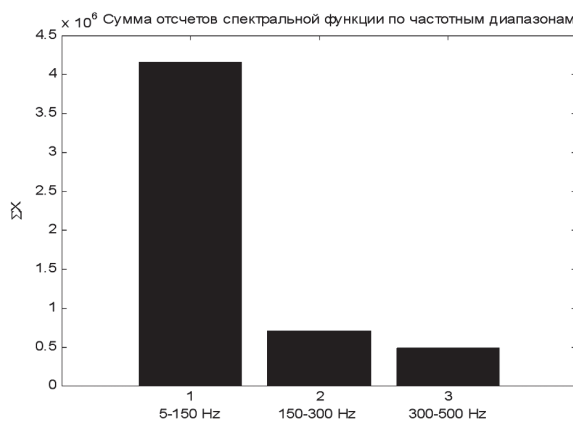
б



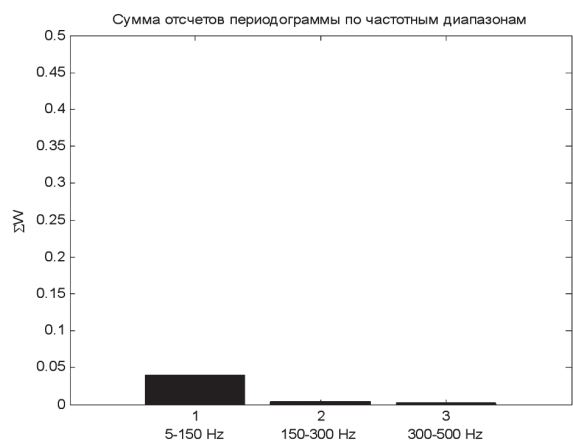
б



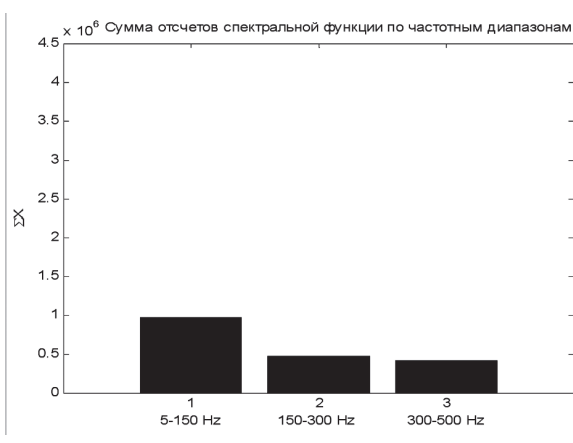
в



в



з



з

Рис. 4

Рис. 5

Проанализировали гистограммы распределения амплитуд электромиографического сигнала. Для построения разбили весь диапазон изменения амплитуд ЭМГ на 100 непересекающихся интервалов, каждому из которых ставим в соответствие количество амплитуд, попавших в этот интервал. По оси абсцисс откладываем значения амплитуд (A) для рассчитанных интервалов, по оси ординат – количество амплитуд в интервале (N).

Форма гистограммы распределения амплитуд соответствует нормальному закону распределения. Для определения соответствия экспериментальной гистограммы распределения амплитуд ЭМГ-сигнала нормальному закону вычислим параметр отклонения экспериментальной гистограммы от теоретической $\Delta S_{\text{норм}}$ (%) следующим образом:

$$\Delta S_{\text{норм}} = \frac{\sum (H_A(X_A) - N(X_A))^2}{\sum N(X_A)^2} \cdot 100\%,$$

где H_A – значения экспериментальной гистограммы распределения амплитуд ЭМГ сигнала; X_A – значения амплитуд ЭМГ в интервалах, рассчитанных для экспериментальной гистограммы; N – значения теоретической нормальной гистограммы.

На рис. 6 приведена аппроксимация нормальным законом распределения экспериментальной гистограммы распределения амплитуд ЭМГ-сигнала.

На рис. 6 *а, б* приведены аппроксимация нормальным законом распределения экспериментальной гистограммы распределения амплитуд ЭМГ-сигнала для правой ($\Delta S_{\text{норм}} = 45,7\%$) и левой ($\Delta S_{\text{норм}} = 5,5\%$) мышц носового клапана соответственно для обследуемого с искривлением носовой перегородки, на рис. 6 *в, г* – правой ($\Delta S_{\text{норм}} = 32,8\%$) и левой ($\Delta S_{\text{норм}} = 19,9\%$) мышц для обследуемого с нормально функционирующим носовым клапаном.

При дальнейших экспериментальных наблюдениях и статистической обработке результатов может быть получен количественный показатель функционирования носового клапана, расчет которого может быть произведен автоматически при снятии электромиографического сигнала.

Выводы

Полученные результаты исследований позволяют оценить степень ригидности носового клапана по амплитудному и частотному анализу

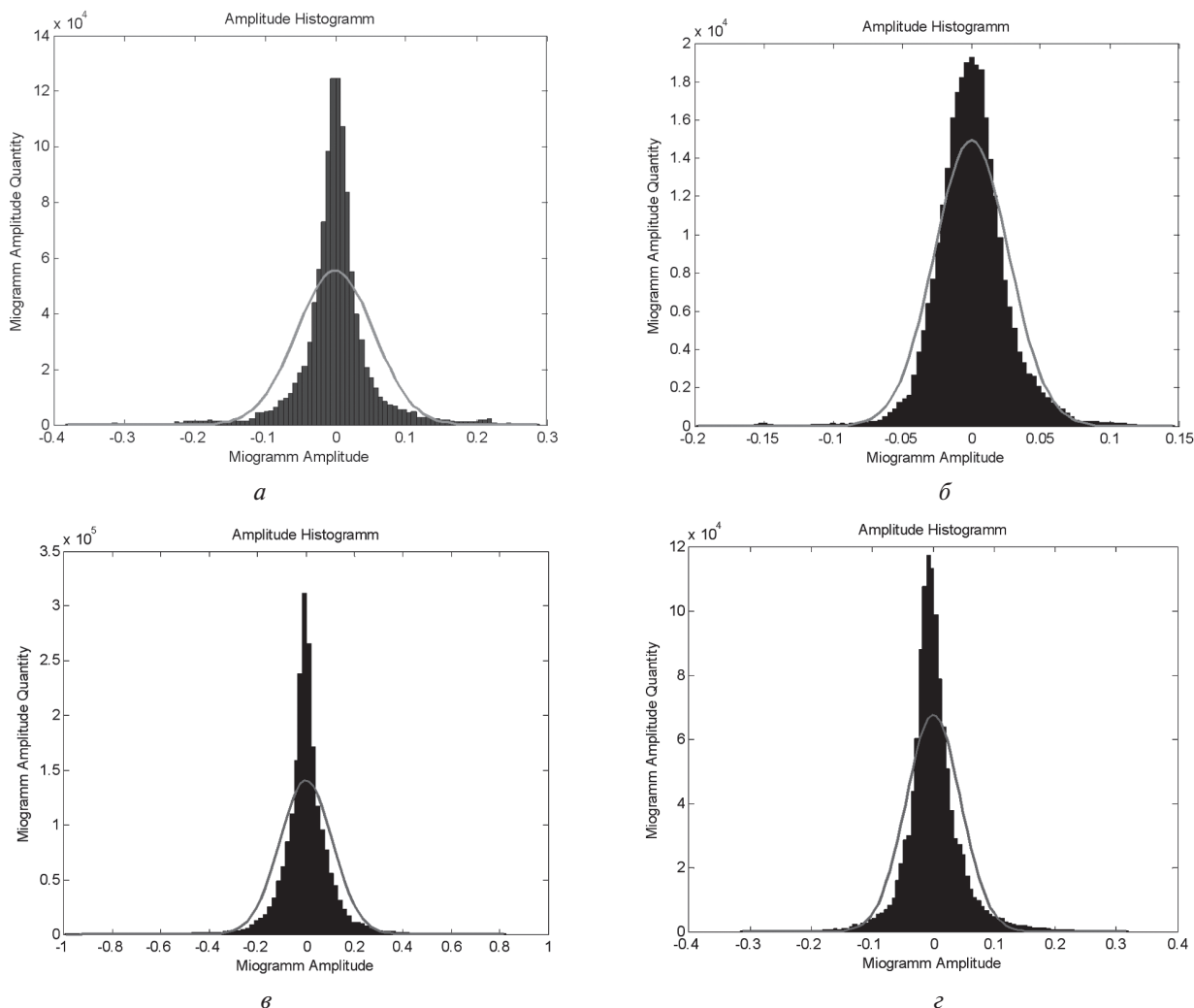


Рис. 6

электромиографического сигнала, который показывает взаимосвязь между уровнем напряжения ЭМГ-сигнала и коэффициентом упругости мышц носового клапана. Перспективой работы являются клинические испытания предложенной методики и формирование статистически значимых количественных критериев оценки функционального состояния носового клапана.

Список литературы: 1. Козлов, В.С. Синуситы: современный взгляд на проблему / В.С. Козлов, В.В. Шиленкова, А.А. Шиленков // *Consilium medicum*. – 2003. – Т5, № 4. – С. 212-219. 2. Юджин, Б.Керн. Хирургия носового клапана [Текст] / Юджин Б.Керн, Том Д.Уонг // *Российская ринология*. – 1995. № 1. – С. 4-33. 3. Platt R.S., Hajduk E.A., Hulliger M/ and Easton P.A. A modified Bessel filter for amplitude demodulation of respiratory electromyograms // *J. Appl. Physiol.* – 1998. –V.84, No. 1. – P. 378–388. 4. Lawrence J.H. and de Luca C.J. Myoelectric signal versus force relationship in different human muscles // *J. Appl. Physiol.* – 1983. – V. 54, No. 6. – P. 1653–1659.

Надійшла до редколегії 24.06.2010.

УДК 615.47

Діагностичні можливості електроміографічного методу для дослідження функції носового клапана / О.Г. Аврунін, Т.В. Жемчужкіна, Т.В. Носова // *Біоніка інтелекту: наук.-техн. журнал*. – 2010. – № 3 (74). – С. 99–104.

В статті розглянуто можливості електроміографічного методу для функціональної діагностики носового клапана. Запропоновано методи аналізу електроміографічного сигналу м'язів носового клапана. Обговорюються можливості та перспективи цих методів для діагностики патологій носового клапана.

Л. 6. Бібліогр.: 4 найм.

UDC 615.47

Diagnostic possibilities of electromyographic method for nasal valve function research/ O.G. Avrunin, T.V. Zhemchuzhkina, T.V. Nosova // *Bionics of Intelligence: Sci. Mag.* – 2010. – № 3 (74). – P. 99–104.

Possibilities of EMG-methods for functional diagnostic of nasal valve are described. The methods for EMG-signal analysis of nasal valve are proposed. Possibilities and perspectives of these methods for diagnostic of nasal valve pathologies are discussed.

Fig. 6. Ref.: 4 items.

УДК 615.47



ПРИМЕНЕНИЕ МЕТОДА ВЕЙВЛЕТ-АНАЛИЗА ДЛЯ ОПРЕДЕЛЕНИЯ АНОМАЛЬНЫХ УЧАСТКОВ НА ЭЛЕКТРОЭНЦЕФАЛОГРАММЕ ЧЕЛОВЕКА

К.Г. Половенко

ХНУРЭ, г. Харьков, Украина, bykh@kharkov.kture.ua

Проанализировано применение метода вейвлет-анализа для определения аномальных участков на ЭЭГ человека. Для детальной расшифровки ЭЭГ использовался масштабный анализ с базисной функцией Добеши на основе вейвлет-преобразований. Показаны преимущества вейвлет-обработки для определения характерных особенностей ЭЭГ при сравнении со спектральным анализом.

ЭЛЕКТРОЭНЦЕФАЛОГРАММА, РИТМЫ, СПАЙКИ, ЭПИЛЕПСИЯ, ДИСКРЕТНОЕ ВЕЙВЛЕТ ПРЕОБРАЗОВАНИЕ, ПРЕОБРАЗОВАНИЕ ФУРЬЕ

Введение

В настоящее время электроэнцефалограмма (ЭЭГ) широко используется для оценки функционального состояния головного мозга. Для установления степени и характера общемозговых нарушений, характеризующих реактивность коры головного мозга на внешние раздражения, определяются вызванные потенциалы головного мозга. Обработка информации вызванных потенциалов способствует достоверной постановке диагноза. К сожалению, современные методы автоматизированной обработки ЭЭГ не позволяют выявлять аномалии [1, 2]. Поэтому поиск новых методов автоматизированной обработки ЭЭГ является актуальным. Одним из таких методов является метод вейвлет-анализа [7, 8].

Целью работы является исследование возможности применения вейвлет-анализа для определения качественных и количественных особенностей электроэнцефалограммы в норме и патологии.

1. Материалы и методы исследования

В работе используются методы дискретного вейвлет-анализа [8, 9, 11] и масштабного анализа на основе вейвлет-преобразований [3].

Обработка сигналов производилась в пакете SYSTAT AutoSignal 1.6, который обеспечивает фильтрацию, обработку и анализ сложных сигналов с помощью интерактивных графических инструментов и детальных отчетов.

Для детальной расшифровки ЭЭГ использовался масштабный анализ с базисной функцией Добеши на основе вейвлет-преобразований, суть которого состоит в разложении сигнала по базисной функции Добеши, образованной сдвигами и масштабными копиями вейвлетной функции. Данный метод позволяет детально рассмотреть отдельные локальные участки сигналов, что дает возможность повысить информативность и точность результатов. Эта особенность использовалась для расширения функциональных возможностей программного обеспечения, разработанного автором, в составе системы вейвлет-преобразований для

частотно-временных характеристик паттернов ЭЭГ сигналов [3, 5].

2. Анализ результатов работы

Проводилась обработка ЭЭГ с помощью методов дискретного вейвлетного и масштабного анализов. Эти методы применялись для качественной и количественной оценки параметров сигналов на аномальных участках ЭЭГ с помощью пакета SYSTAT AutoSignal 1.6. На рис. 1 и 2 представлены результаты обработки ЭЭГ-сигналов с помощью построений Фурье и дискретных вейвлет спектров, полученных с отведения *Poz*. На рис. 1 представлены: исходный сигнал ЭЭГ здорового пациента (условно считаем ЭЭГ без патологий), Фурье и вейвлет спектры. При построении вейвлет-спектра (рис. 1б) наблюдаются незначительные изменения в сигнале ЭЭГ человека, что говорит о случайных вызванных потенциалах в мозге. На рис. 2 представлен сигнал ЭЭГ человека, больного эпилепсией, где имеют место явно выраженные всплески (устойчивые острые пики) (рис. 2а), которые характеризуются также ускорением ритмов ЭЭГ, чередованием быстрых и медленных колебаний, увеличением амплитуды колебаний

Из данного рисунка видно, что построение Фурье-спектра является мало информативным при определении как наличия, так и типа аномалии сигналов. Частотный Фурье-спектр отображает лишь глобальные сведения о частотах исследуемого сигнала и не дает представления о локальных свойствах сигнала при его временных изменениях. Наличие множественных всплесков с окраской от темно-синего до ярко красного в вейвлет-спектре (рис. 2б) свидетельствует о явных патологиях. Минимальные уровни спектра соответствуют низкочастотным составляющим сигнала (окрашены в темно-синий цвет), а максимальные – высокочастотным (окрашены в темно-красный). Эти изменения цветовой палитры объясняются тем, что у данного пациента наблюдаются серьезные патологические изменения в низкочастотной области: именно эпилепсия характеризуется наличием

низкочастотных составляющих спектра сигнала. Таким образом, можно сделать вывод, что дискретный вейвлет-анализ пригоден для выявления такой патологии, как эпилепсия, и имеет явные преимущества перед традиционным спектральным анализом, обеспечивая, в частности, корректное описание нестационарных процессов и сохраняя более полную информацию об особенностях ЭЭГ.

Известно, что вейвлет-анализ представляет собой разложение сигнала по специальному базису, а базисные функции получаются сдвигом и масштабированием (сжатием и растяжением) одной функции порождающего вейвлета [5]. Каждому значению масштаба вейвлет-анализа соответствует бесконечное количество сдвинутых друг относительно друга локализованных в пространстве функций.

Масштабный анализ определяет общий подход для поиска различных вейвлет-базисов. При вычислении вейвлет-преобразований для дискретных сигналов открываются широкие возможности его практической реализации.

Автором предложено использовать масштабный анализ с помощью вейвлет-преобразований с базисной функции Добеши. Был разработан алгоритм и программное средство «Масштабный анализ по Добеши», где осуществляется разложение сигнала с определенным сдвигом: чем меньше сдвиг, тем больше площадь преобразования и больше информации для детальной обработки сигнала.

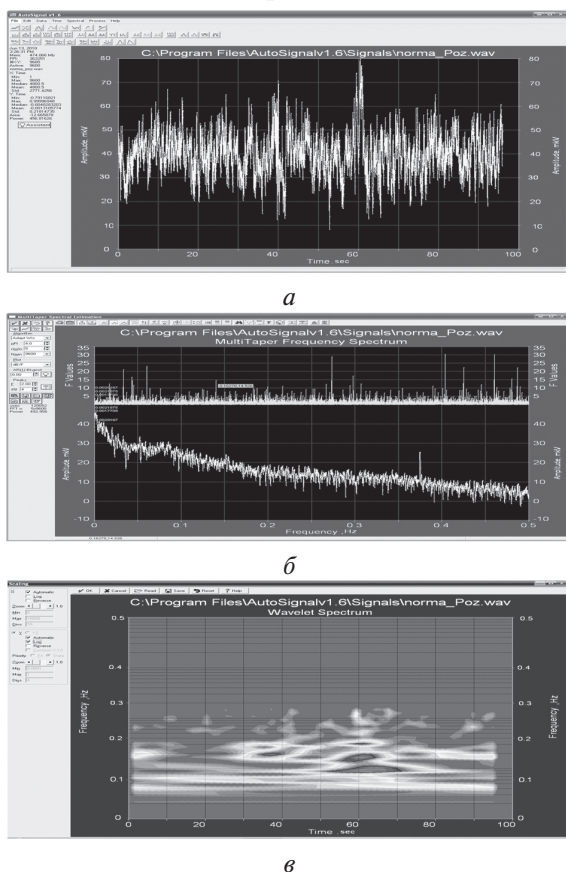


Рис. 1. Исходный сигнал пациента здорового пациента (а) с отведения Poz, его Фурье-спектр (б) и вейвлет-спектр (в)

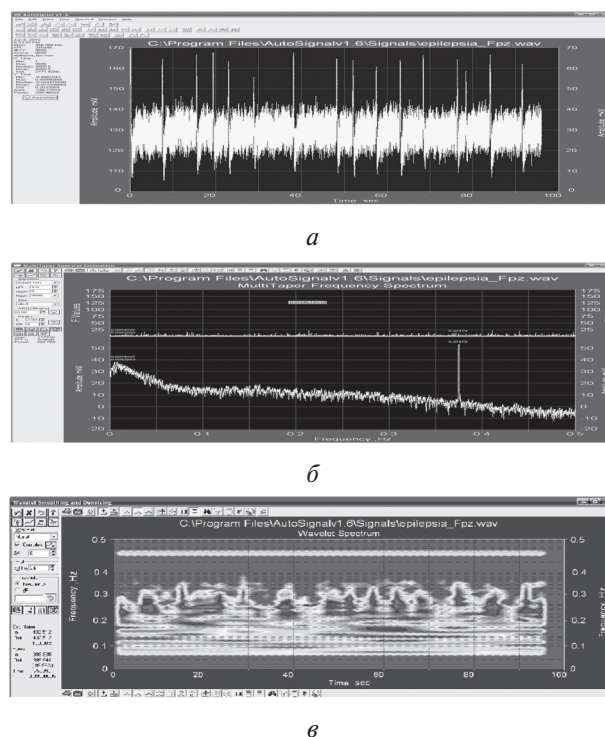


Рис. 2. Исходный сигнал пациента больного эпилепсией (а) с отведения Poz, его Фурье-спектр (б) и вейвлет-спектр (в)

Данный программный продукт позволяет аппроксимировать исходный сигнал, провести его детализацию, вычислить коэффициенты аппроксимации и детализации. Выбирая эти коэффициенты, удастся детально проанализировать исходный сигнал.

С помощью данного программного средства для двух разных людей, имеющих разные диагнозы, были выбраны коэффициенты детализации и аппроксимации сигнала на определенных уровнях, при этом анализировался уровень изменения деконпозиции и порядка функции Добеши. На рис. 3-6 показаны результаты обработки ЭЭГ-сигналов с помощью разработанного программного средства «Масштабный анализ по Добеши».

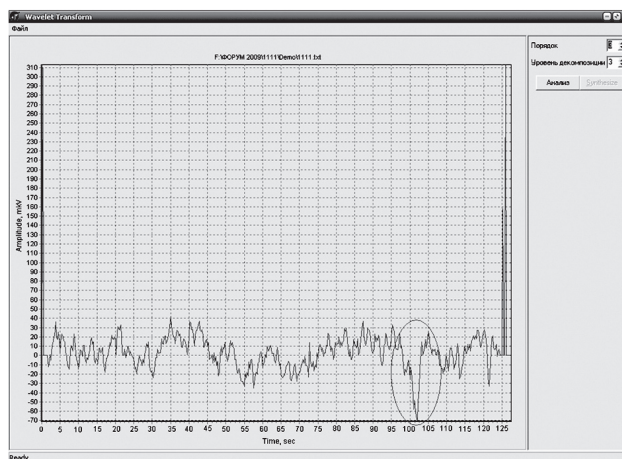


Рис. 3. Исходный сигнал здорового человека, снятого с отведения T4_1

На рис. 3, 4 овалами обведены характерные участки нормальной ЭЭГ и показано их детальное представление (рис. 4) при разложении сигнала, имеющийся выброс имеет амплитуду в пределах нормы ЭЭГ-сигнала. На рис. 5, 6 представлен анализ ЭЭГ с помощью разработанного программного средства при разных порядках уровня декомпозиции и функции Добеши. Здесь овалами обведены аномальные участки, характеризующие патологическую активность коры головного мозга человека. Выбросы на рис. 5, 6 имеют амплитуду, намного превышающую норму ЭЭГ-сигнала. При увеличении порядка базисной функции и уровня декомпозиции можно подробно рассмотреть изменение сигнала во времени и выделить характерные особенности аномальных участков.

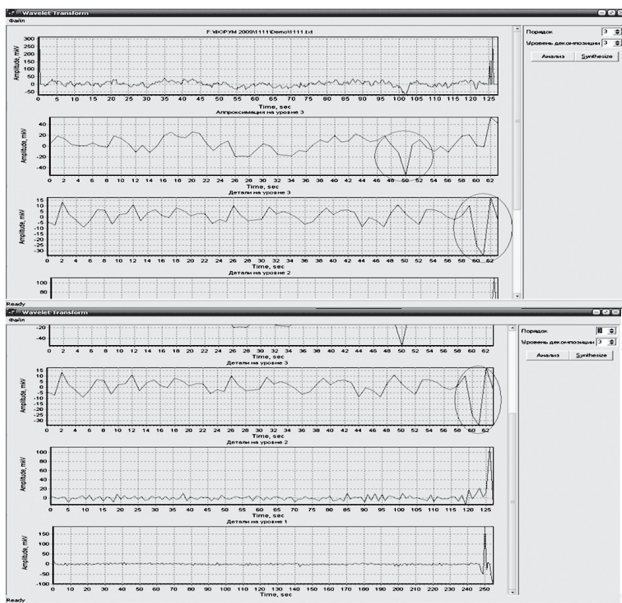


Рис. 4. Разложение сигнала масштабным анализом нормальной ЭЭГ. Порядок Добеши 3, уровень декомпозиции 3

На рис. 6 видно, что чем больше порядок базисной функции Добеши и уровень декомпозиции, тем подробнее можно рассмотреть сигнал для выявления его характерных участков на всех уровнях аппроксимации и детализации, что повышает степень определения той или иной аномалии.

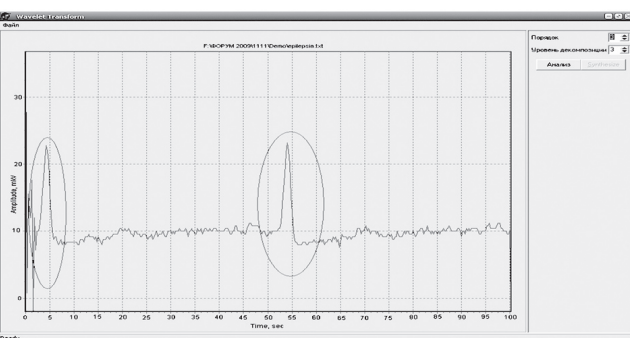


Рис. 5. Электроэнцефалограмма пациента, больного эпилепсией, снятой с отведения Т4_1

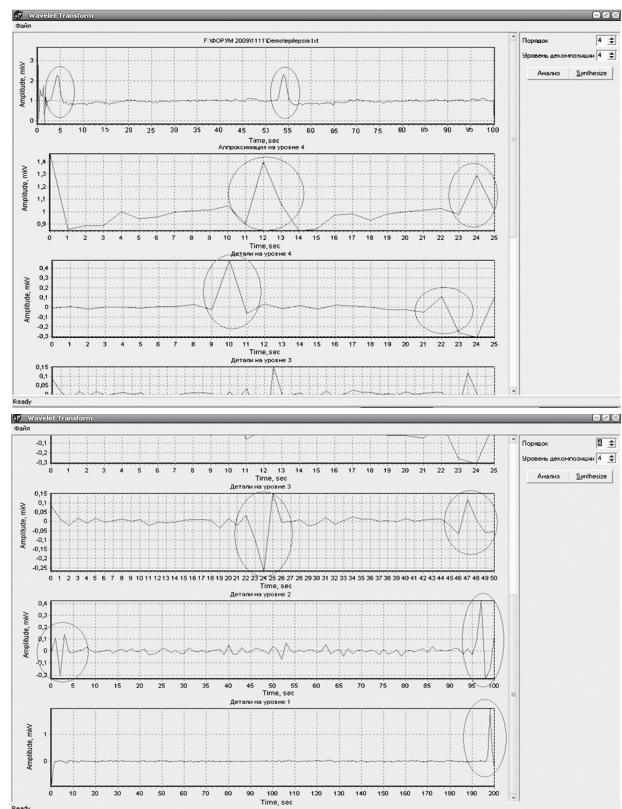


Рис. 6. Разложение сигнала масштабным анализом эпилептической ЭЭГ. Порядок Добеши 4, уровень декомпозиции 4

Аналогичные результаты были получены при исследовании ЭЭГ пациентов с 5 патологиями: черепно-мозговая травма, опухоль головного мозга, шизофрения, болезнь Паркинсона, инсульт.

Выводы

Получены качественно новые результаты, показывающие возможность ранжирования патологий внутри диагностической группы, что в дальнейшем может использоваться для определения степеней сложности заболеваний головного мозга.

Вейвлетный метод-анализ показал надежные результаты даже при наличии только одного образца для каждого типа аномалий.

Вейвлет-метод обработки чувствителен к любым преобразованиям исходной формы сигнала и позволяет представить их в удобной форме. Вейвлет-преобразования эффективны как для выявления кратковременных аномалий сигнала (спайков), так и устойчивых изменений его параметров, а также пригоден для установления факта появления патологии и определения ее характера.

Предложен метод обработки и оценки электроэнцефалографических кривых, основанный на масштабном анализе с базисной функцией Добеши. Установлено, что с увеличением уровня декомпозиции (3-6) осуществляется более подробная детализация исходного сигнала, что позволяет более точно диагностировать патологию.

Список литературы: 1. *Сахаров, В.Л.* Методы математической обработки электроэнцефалограмм [Текст]: Учебное пособие / В.Л. Сахаров, А.С. Андреев // Таганрог: "Антон", 2000. — 44 с. 2. Клиническая электроэнцефалография [Текст] / Под ред. В.С. Русинова // М.: "Медицина", 1973. — 340 с. 3. *Рангайян, Р.М.* Анализ биомедицинских сигналов. Практический подход [Текст]: пер. с англ. / под ред. А.П. Немирко // М.: ФИЗМАТЛИТ, 2007. — 440 с. 4. *Айфичер, Эммануил С.* Цифровая обработка сигналов: практический подход [Текст]: 2-е издание / Айфичер Эммануил С., Джервис Барри У.: пер. с англ. — М.: Издательский дом "Вильямс", 2004. — 992 с. 5. *Добеши И.* Десять лекций по вейвлетам [Текст] / И. Добеши. — Ижевск: НИЦ "Регулярная и хаотическая динамика, 2001. — 464 с. 6. *Голуб, В.А.* Автоматическое распознавание острых волн в электроэнцефалографии [Текст] / В.А. Голуб, О.В. Мачульская, Н.П. Сереженко // Сб. докладов междунар. науч.-практич. конф. «Кибернетика и технологии XXI» века. — Воронеж, 2000. — С. 119-123. 7. *Голуб, В.А.* Автоматическое распознавание паттернов в электроэнцефалографии [Текст] // В.А. Голуб, О.В. Мачульская, Н.П. Сереженко // Сб. докладов междунар. науч.-практич. конф. «Кибернетика и технологии XXI» века. — Воронеж, 2001. — С. 458-462. 8. *Голуб, В.А.* Выявление паттернов ЭЭГ с помощью вейвлет-преобразования [Текст] // В.А. Голуб, И.Н. Козлова, Н.П. Сереженко // Вестник ВГУ, сер. Системный анализ и информационные технологии. — 2007. — № 2. — С. 61-64. 9. *Голуб, В.А.* Выявление паттернов ЭЭГ с помощью вейвлет-преобразования [Текст] / В.А. Голуб, И.Н. Козлова, Н.П. Сереженко // Информатика: проблемы, методология, технологии: Материалы 9-й междунар. науч.-метод. конф., Воронеж, 12-13 февраля 2009 г., 2009.

Т. 1. — С. 197-203. 10. *Блейхут, Р.* Быстрые алгоритмы цифровой обработки сигналов [Текст] / Р. Блейхут. — М.: Мир, 1999. — 448 с. 11. *Есауленко, И.Э.* Распознавание паттернов электроэнцефалограмм с использованием вейвлет-анализа [Текст] / И.Э. Есауленко, Н.П. Сереженко, С.Н. Семенов // Системный анализ и управление в биомедицинских системах. — 2009. — Т. 8, № 4. — С. 855-857.

Поступила в редколлегию 3.09.2010.

УДК 615.47

Застосування методу вейвлет-аналізу для визначення аномальних ділянок на електроенцефалограмі людини / К.Г. Половенко // Біоніка інтелекту: наук.-техн. журнал. — 2010. — № 3 (74). — С. 105–108.

В статті розглядаються перспективні напрями обробки електроенцефалограм за допомогою вейвлет-аналізу з метою виявлення аномальних ділянок. Показані переваги вейвлетного аналізу для визначення характерних особливостей електроенцефалограм при порівнянні зі спектральним методом.

Л. 11. Бібліогр.: 11 найм.

UDK 615.47

Applications of method of veyvlet-analysis for determination of anomalous areas on the electroencephalogram of human / K.G. Polovenko // Bionics of Intelligence: Sci. Mag. — 2010. — № 3 (74). — P. 105–108.

In the article perspective directions of treatment of electroencephalograms are examined by a veyvlet-analysis with the purpose of exposure of anomalous areas. Advantages of veyvletnogo analysis are rotined for determination of characteristic features of electroencephalograms when compared to a spectral method.

Fig. 11. Ref.: 11 items.

УДК 004.6



СЛАБОСТРУКТУРИРОВАННЫЕ ТЕКСТОВЫЕ ДОКУМЕНТЫ КАК ИСТОЧНИКИ ДАННЫХ

В.А. Губин

ХНУРЭ, г. Харьков, Украина, gubinv@rambler.ru

Исследованы возможности и предпосылки автоматизированной идентификации данных в электронных текстовых документах, размещенных в корпоративных сетях и в сети Internet. Выделен класс слабоструктурированных текстовых документов, представлены их основные признаки и свойства как источник данных. Приведены различные сценарии взаимного расположения в таких документах текстовых фрагментов, соответствующих атрибутам и значениям данных.

СЛАБОСТРУКТУРИРОВАННЫЙ ДОКУМЕНТ, ДАННЫЕ, АТРИБУТ ДАННЫХ, ЗНАЧЕНИЕ ДАННЫХ

Введение

Объем контента, содержащегося в сети Internet и в корпоративных сетях, растет с каждым годом в геометрической прогрессии. Большую его часть составляет мультимедиа: видео-, аудио-файлы, файлы с изображениями различных форматов. Но и объемы электронных текстовых документов в абсолютном измерении также увеличиваются все более возрастающими темпами. В связи с этим все более и более актуальной становится задача эффективного поиска и доступа к информации, содержащейся в этих документах.

Изначально предполагалось преимущественно предполагается и в настоящий момент, что с документами, опубликованными в сети Internet, в первую очередь будет знакомиться человек. Этой цели подчинено и внутреннее представление документов и то, как они отображаются в окне просмотра того или иного Internet-браузера. Ориентированность на восприятие человеком той информации, которая содержится в подавляющем числе опубликованных в сети документов, мешает доступу к ней и обработке ее тем или иным программным обеспечением. В данной работе предпринимается попытка выделить класс документов, для которых эта проблема может быть преодолена.

1. Цели и задачи исследования

Целью данной работы является исследование возможностей и предпосылок автоматизированной идентификации данных в электронных текстовых документах, размещенных в корпоративных сетях и в сети Internet. Для этого из всего множества документов необходимо выделить класс текстовых документов, содержимое которых предназначено для непосредственного ознакомления человеком, но, в силу особенностей их внутренней структуры, имеющих предпосылки для автоматизированной идентификации содержащихся в них данных. Целью также является исследование основных признаков и свойств такого рода документов как источников данных.

2. Исследование особенностей слабоструктурированных текстовых документов

Главное предназначение электронных текстовых документов – быть источником информации. Причем в зависимости от формата текстовых документов информация представлена в виде данных, ориентированных, в первую очередь, на автоматизированную обработку, либо в виде, изначально предполагающем, что основным агентом, получающим к ним доступ, будет человек.

К первому типу документов можно отнести документы в формате xml, rdf и подобных им. Эти форматы ориентированы на описание данных. Их содержимое аналогично содержимому таблиц базы данных. Поэтому при автоматизированной обработке таких документов не возникает проблем с идентификацией и доступом к данным, содержащимся в них. Причем сами эти документы, как правило, также получены как итог работы того или иного программного обеспечения.

Ко второму типу документов можно отнести документы в форматах html, doc, rtf и подобных им. В этом случае изначально предполагается, что с содержимым таких документов будет знакомиться человек. Разумеется, речь не идет о непосредственном доступе человека к содержимому. Человек знакомится с экранным представлением таких документов в окне просмотра браузера Internet Explorer, например, либо в редакторе текстов Microsoft Word, либо в любой другой, аналогичной, среде.

На рис. 1 изображен фрагмент электронного текстового документа в формате html, содержащий информацию о мобильном телефоне.

Совершенно очевидно, что представленный выше документ содержит некоторые сведения. Они относятся к тем или иным характеристикам мобильного телефона. Просматривая этот документ, человек может легко извлечь информацию, например, о параметрах дисплея, об объеме встроенной памяти, о разрешении фотокамеры телефона и т.д.

Но если ставится задача автоматизированного извлечения данных из такого рода текстов, то, в общем случае, есть риск столкнуться с существ-

венными трудностями, связанными с необходимостью понимания смысла естественно-языковых источников.

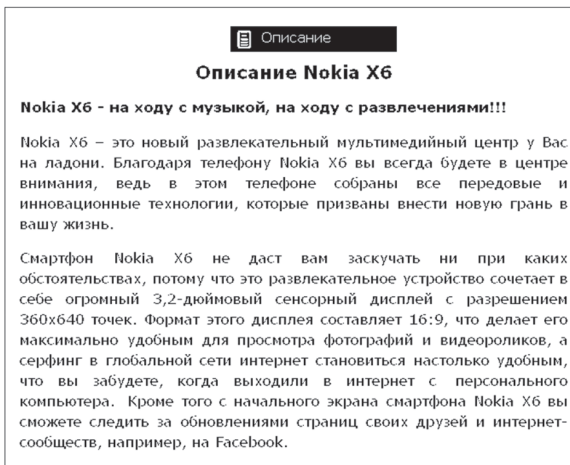


Рис. 1. Фрагмент html-документа с описанием мобильного телефона

Если нажать на этой же странице на кнопку «Характеристики», то получится несколько иное текстовое представление тех же сведений о мобильном телефоне (рис. 2).

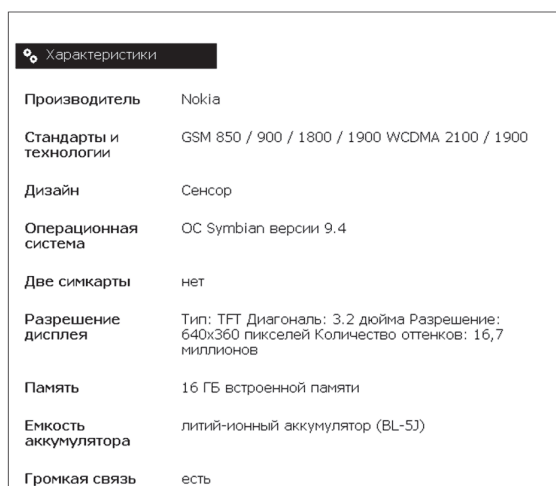


Рис. 2. Фрагмент html-документа с характеристиками мобильного телефона

Это представление уже более напоминает то, что имеет место в таблицах базы данных или в xml-документах. Здесь имеется больше предпосылок для автоматизированного извлечения данных. Это объясняется тем, что, в отличие от документа, представленного на рис. 1, данный документ содержит обособленные форматированием текстовые фрагменты. Причем некоторые из них объективно относятся к значениям данных («Nokia», «GSM 850 / 900 / 1800 / 1900 WCDMA 2100 / 1900» «Сенсор»). Другие же обособленные текстовые фрагменты («Производитель», «Стандарты и технологии», «Дизайн», «Операционная система») напоминают имена полей в таблице базы данных и, по сути, являются атрибутами данных, помогающими правильно интерпретировать содержимое

тех обособленных текстовых фрагментов, которые выше были отнесены к значениям данных.

Таким образом, представленный на рис. 2 документ объективно содержит данные, представленные парами «атрибут-значение». Но с формальной точки зрения содержимое документа – набор равнозначных фрагментов текста. Следовательно, если рассмотреть фрагмент внутреннего представления соответствующего html-документа (рис. 3), то можно увидеть, что нет формальных признаков, указывающих на то, чем является тот или иной обособленный фрагмент текстового документа: значением данных либо атрибутом данных. Из этого следует, что данные в таких документах не имеют строгого описания. Для того чтобы эти данные получили строгое описание, необходимо идентифицировать каждый такой обособленный фрагмент текстового документа либо как атрибут данного, либо как значение данного.



Рис. 3. Фрагмент внутренней разметки документа с характеристиками мобильного телефона

Таким образом, из всего множества текстовых документов можно выделить класс текстовых документов, называемых слабоструктурированными. Можно выделить следующие их признаки и свойства как источников данных:

1. Документ содержит внутреннюю разметку.
2. Содержимое документа разбито внутренним форматированием на обособленные текстовые фрагменты.
3. Каждый фрагмент объективно представляет собой либо значение данных, либо атрибут данных.
4. Во внутренней разметке документов нет формальных признаков, указывающих на то, что есть значение данных, а что есть атрибут данных.

Для такого рода документов будет наиболее эффективным их представление в виде совокупности объектов [1] с последующей возможной идентификацией текстовых фрагментов как атрибутов и значений данных.

Примеры типов слабоструктурированных текстовых документов:

1. Анкеты.

2. Страховые формы.
3. Налоговые декларации.
4. Счета.
5. Транспортные накладные.
6. Контракты.
7. Технические параметры изделия.
8. Прайс-листы.
9. Типовые договоры.
10. Карточки учета.
11. Результаты спортивных матчей.

Разумеется, этот перечень далеко не исчерпывает все возможные типы слабоструктурированных текстовых документов. Причем каждый тип может иметь свои подтипы.

Если бы данные в текстовых документах всегда были представлены в таком виде: таблица, состоящая из двух столбцов, в первой ячейке каждой строки таблицы – атрибут данных, а во второй – значение данных, то проблем с автоматизированной классификацией того, что есть что, не было бы. Но сложность автоматизированной классификации обособленных фрагментов текстового документа заключается и в том, что топологически эти фрагменты могут располагаться относительно друг друга по-разному, а не только так, как показано на рис. 2. Пример фрагмента реального документа, подтверждающего данную мысль, представлен на рис. 4.

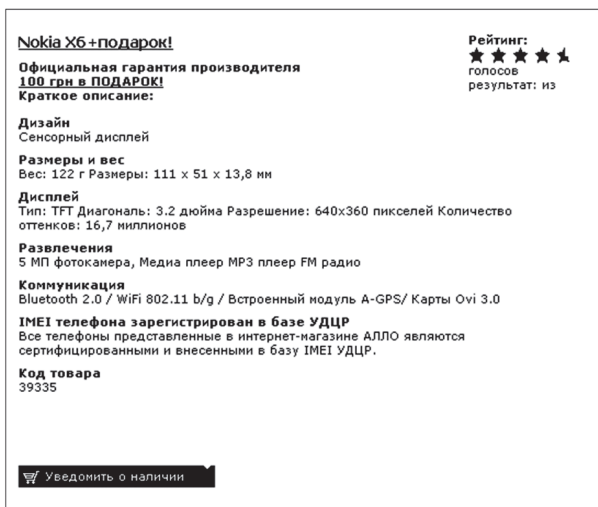


Рис. 4. Альтернативное представление характеристик мобильного телефона

Можно выделить основные варианты взаимного расположения значения и атрибута данных в слабоструктурированных текстовых документах:

1. Атрибут данных располагается в первом столбце таблицы, а значения данных – в ячейках соответствующих строк.
2. Атрибут располагается в первой строке таблицы, а значения данных – в ячейках соответствующего столбца.
3. Атрибут и значение данных располагаются в одном и том же абзаце.
4. Атрибут и значение данных располагаются в смежных абзацах.

5. Атрибут является элементом списка, а значения данных являются элементами вложенного списка.

6. Атрибут присутствует в документе, но не имеет явного соответствующего значения данных.

7. Значение данных присутствует в документе, но не имеет явного соответствующего атрибута.

8. Возможны комбинации вышеприведенных сценариев и, возможно, другие сценарии.

Возможный подход к определению того, чем является тот или иной обособленный фрагмент текстового документа, определяется тем, что в совокупности документов одного и того же типа фрагменты, соответствующие атрибуту данных, будут оставаться без изменений, а фрагменты, соответствующие значению данных, будут изменяться от документа к документу. Таким образом, при попытке извлечения данных из слабоструктурированных текстовых документов необходимо оценить только степень устойчивости к изменению каждого фрагмента текстового документа без необходимости проникновения в семантику текста.

Выводы

Исследования, представленные в данной работе, показывают, что из всего спектра электронных текстовых документов можно выделить подкласс документов, имеющих потенциал для автоматизированной идентификации данных. Приведены примеры и выделены особенности слабоструктурированных текстовых документов как источников данных.

Намечен возможный подход к определению того, чем является тот или иной обособленный фрагмент слабоструктурированного текстового документа: атрибутом данных или значением данных.

Научной новизной работы является следующее: из множества текстовых документов выделен класс слабоструктурированных текстовых документов и представлены их основные признаки и свойства как источник данных.

Список литературы: 1. Губин, В.А. Объектное представление электронных текстовых документов [Текст] / В.А. Губин, А.Н. Гвоздинский // Радиоэлектроника и информатика. – 2007. – № 1 (36). – С. 61-63.

Поступила в редколлегию 8.09.2010.

УДК 004.6

Слабоструктуровані текстові документи як джерела даних / В.О. Губін // Біоніка інтелекту: наук.-техн. журнал. – 2010. – № 3 (74). – С. 109–111.

У статті виділено клас слабоструктурованих текстових документів. Представлені їх основні ознаки і властивості як джерел даних.

Л. 4. Бібліогр.: 1 найм.

UDK 004.6

Semi-Structured Text Documents as Data Sources / V.A. Gubin // Bionics of Intelligence: Sci. Mag. – 2010. – № 3 (74). – P. 109–111.

In article the class of semi-structured text documents is isolated. Their main features and properties as data sources are represented.

Fig. 5. Ref.: 1 items.

УДК 004.047:681.3.01



ПОДДЕРЖКА ЦЕЛОСТНОСТИ БАЗ ДАННЫХ ПРИ ДИНАМИЧЕСКИХ СТРУКТУРНЫХ ИЗМЕНЕНИЯХ

С.С. Таянский

ХНУРЭ, г. Харьков, Украина, tanyansky_ss@yahoo.com

Рассмотрен основанный на логике подход к описанию и поддержке ограничений целостности в базах данных при их модификациях. Сформулированы условия, при которых модификация базы данных может быть выполнена корректно. Определены границы модификаций БД, гарантирующие целостность и согласованность неоднородных структур данных.

БАЗА ДАННЫХ, ЦЕЛОСТНОСТЬ, ИНФОРМАЦИОННЫЙ ОБЪЕКТ, СЕМАНТИКА ДАННЫХ, МОДИФИКАЦИЯ БАЗЫ ДАННЫХ, ЛОГИЧЕСКОЕ ПРАВИЛО

Введение

Одним из основополагающих понятий в технологии баз данных (БД) является понятие целостности. В общем случае это понятие, прежде всего, связано с тем, что БД отражает в информационном виде некоторый объект или совокупность взаимосвязанных объектов реального мира. Под целостностью будем понимать соответствие информационной модели предметной области (ПрО), хранимой в БД, объектам реального мира и их взаимосвязям в каждый момент времени. Любое изменение в ПрО, значимое для построенной модели, должно отражаться в БД, и при этом должна сохраняться однозначная интерпретация информационной модели в терминах этой ПрО.

Поддержка целостности в ее классическом понимании включает понятие структурной целостности, которая трактуется как то, что система управления БД может допускать работу только с однородными структурами данных, например, только типа “реляционное отношение” (отсутствие дубликатов кортежей, обязательное наличие первичного ключа, отсутствие понятия упорядоченности кортежей и др.) [1].

При работе с распределенными или интегрируемыми информационными системами (ИС) на практике приходится сталкиваться с неоднородными структурами данных. Для обработки неоднородных данных наиболее трудно решаемой задачей является поддержка целостности. С одной стороны, это связано с несоответствиями между представлениями или форматами данных, поступающих из разных источников. С другой стороны – при взаимодействии множества источников данных часто возникает проблема нарушения семантической целостности. При различной семантике данных в поддерживаемых БД необходимо найти способы не только для разрешения семантических конфликтов, но и для поддержки согласованности информации БД.

Исследования в этой области ведутся с момента практического использования БД в распределенных и крупномасштабных системах. Оригиналь-

ные подходы были рассмотрены в [2], а также при построении испытательной распределенной базы данных [3]. Подходы к управлению распределенными ресурсами представлены в работах Sheth A. P., Larson J.A. [4], Garcia-Solaso M., Saltor F., Castellanos M. [5], в которых рассматриваются задачи, возникающие при достижении локальной автономности. Среди переведенных источников можно выделить работу группы авторов [6], которая отличается широтой и глубиной охвата материала по вопросам проектирования и использования современных систем БД.

Основным выводом из рассмотренного материала является то, что проблема поддержки целостности в неоднородных информационных системах является актуальной и при этом требует умения разумно использовать сочетание технологических и архитектурных решений.

Таким образом, основной проблемой при решении задач поддержки целостности неоднородных БД является обобщенное представление объектов ПрО. В статье используется определение информационного объекта ПрО на основе требований ИС и рассмотрен подход к обеспечению их семантической однозначности.

Целью работы является представление средств описания ПрО и формальное определение семантики данных, используя введенные свойства БД с динамически изменяемой структурой. Для обеспечения универсальности подхода будем рассматривать основанный на логике подход к описанию и поддержке ограничений целостности в БД при их модификациях.

1. Определение семантики базы данных

Используя разработанные средства представления ПрО как набор информационных объектов O , логические правила существования запишем как выражение (1).

$$o \dashv o_1, \dots, o_n \text{ или } l \dashv l_{1, \dots, n} \quad (1)$$

В дальнейшем множество заданных информационных объектов $o_1, \dots, o_n \in O$ будем называть свободными, а множество правила вида (1) будем

обозначать $L = \{l_1, l_2, \dots, l_n\}$ или $L = \{l_{1,\dots,n}\}$ и называть L - правилами.

Содержательно правило (1) заключается в том, что если все объекты $l_{1,\dots,n}$ входят в O , то и объект l также должен входить в O . Поэтому поддержка актуального состояния БД может привести к дополнительным действиям в следующем смысле: если объекты $l_{1,\dots,n}$ добавляются к O , то к нему нужно добавить и объект l . Такие дополнительные действия необходимы для обеспечения целостности БД.

Ограничением целостности будем называть множество правил $I \subseteq L$, выполнение которых переводит БД из одного целостного состояния в другое целостное состояние. Отличие ограничений I от L - правил, описывающих свойства ПрО, заключается в следующем:

1. Детерминанты (левая часть) правил могут иметь информационные объекты с отрицанием “ \neg ”.

2. Ограничения могут привести к дополнительным манипулированиям объектами при модификациях БД (вставка, удаления).

3. Как следствие пунктов 1) и 2), два ограничения могут оказаться несовместными в том смысле, что их поддержка требует выполнения конфликтующих действий (например, одно ограничение требует вставки некоторого информационного объекта o , а другое – удаления того же объекта).

4. Ограничения I имеют приоритет перед обычными L -правилами, то есть некоторый объект можно вывести с помощью L -правила, только если его отрицание не выводится с помощью некоторого ограничения [8].

Учитывая тот факт, что теория, на основании которой построены принципы целостности БД, является непротиворечивой [9], введем определения для множества объектов O и множества правил из I .

Определение 1. Множество O совместно, если оно не содержит никакого объекта одновременно с его отрицанием.

Определение 2. Множество ограничений I совместно, если, начав с любого объекта l , нельзя породить какой-либо объект l' и его отрицание.

Основная концепция предлагаемого подхода состоит в том, что и БД и множество ограничений целостности совместны.

Декларативно понятие семантики ПрО определим как множество S заданных информационных объектов $\{o_j\}$ (в частности, достаточно базиса множества информационных объектов B) и порожденных с помощью L -правил информационных объектов. Исходя из этого, можно сделать вывод о том, что семантики S может оказаться несовместной, даже если O и L совместны.

Например, если $O = \{l_a, \neg l_b\}$, а $L = \{l_b \rightarrow l_a\}$, то O и L совместны, а семантика $S = \{l_a, l_b, \neg l_b\}$ несовместна.

Под модификацией БД будем понимать добавление или удаление информационного объекта, при выполнении которых множество O остается совместным. На множества объектов, являющиеся кандидатами в семантики модифицируемых БД, наложим условия, аналогичные требованиям из работы [10]. Покажем, что при их выполнении всегда можно найти, по крайней мере, одну совместную семантику S' для модифицируемой БД.

В дальнейшем будем рассматривать модификации с детерминированной семантикой, то есть основанные на таких ограничениях целостности, для которых всегда имеется ровно одна возможная семантика S' модифицированной БД.

В этом контексте для обеспечения поддержки целостности необходимо решить следующие задачи:

1. Определить свойства детерминированных множеств ограничений.

2. Свести детерминированные множества ограничений к эквивалентным множествам простых ограничений, то есть ограничений вида $l \rightarrow l_1$.

Отметим еще одно существенное отличие рассматриваемого подхода к модификациям от других подходов, например описанных в [10]. Оно состоит в определении требования так называемого “минимального изменения”. Требование заключается в следующем: чтобы некоторое множество объектов S' можно было рассматривать как возможную семантику модифицированной БД, оно должно представлять минимальное изменение семантики S исходной БД.

Так, в работе [10] разница между S и S' измеряется через их симметрическую разность $S \setminus S' = (S - S') \cup (S' - S)$. Однако, как будет показано, минимальность симметрической разности не обязательно означает сохранение максимально возможного множества информационных объектов исходного множества O . Другими словами, минимальность симметрической разности не гарантирует выполнения аксиом изменения сформулированных в [11].

Введем другую меру изменения, из минимальности которой следует минимальность симметрической разности, и при этом в результирующем множестве O будет сохраняться максимально возможное количество объектов. Такое определение минимального изменения было подсказано подходом, использованном в работе [12].

Для применения методов поиска наименьших достаточных изменений далее рассмотрим основные определения и обозначения, связанные с БД и ограничениями целостности; определим семантику модификаций и установим существование, по крайней мере, одной семантики для модифицируемой БД; определим характеристики детерминированных множеств ограничений и покажем, что

каждое детерминированное множество ограничений можно свести к эквивалентному множеству простых ограничений вида $I \sqcup I_1$.

2. Свойства семантики базы данных

Определим БД как множество, состоящее из констант, переменных и информационных объектов ПрО. Информационный объект I будем называть позитивным, а его отрицание $\neg I$ – негативным, при этом свободным информационным объектом будем называть позитивный или негативный предикат, не содержащий констант и\или переменных.

В контексте введенных определений БД определим как множество информационных объектов O вместе с множеством правил L и ограничениями I , то есть БД можно представить тройкой элементов вида (2):

$$DB = \langle O, L, (I \subseteq L) \rangle. \quad (2)$$

Отметим, что O может изменяться при изменениях требований ПрО, в то время как I может оставаться без изменения. Это позволит исключить этап работы с функциональными L^f и структурными L^r правилами, и рассматривать O как конечный набор информационных объектов рассматриваемой ПрО [8].

Пусть O – множество свободных информационных объектов, L – множество правил на O , I – некоторое фиксированное множество ограничений, содержащее все правила из L . Через $O^I(L)$ обозначим множество всех экземпляров правил из L , в которых все составные объекты заменены информационными объектами из I . Свяжем с I семантический оператор η_I , который определим следующим образом (3):

$$\eta^I(O) = \{I \mid I \sqcup I_{1, \dots, n} \in O^I(L) \text{ и} \\ I \sqcup I_{1, \dots, n} \in I \cup I\}. \quad (3)$$

Для простоты представления при очевидности множества ограничений будем использовать символ η вместо η^I . Так как η^I – монотонный оператор по отношению включения множеств (если $I_1 \subseteq I_2$, то $\eta^{I_1} \subseteq \eta^{I_2}$), то последовательность (4) имеет предел.

$$\eta_0(O) = O \dots \eta_n(O) = \eta(\eta_{n-1}(O)), \forall n > 0. \quad (4)$$

Обозначим соответствующий предел как η^O и назовем наименьшей неподвижной точкой η относительно O , то есть $\eta^O(O) = O$. Для одного информационного объекта I предел будем обозначать как η^I .

Так как оператор η монотонный, можно доказать, что неподвижная точка η^O также монотонна по аргументу O (относительно включения множеств).

Утверждение 1. Если O и O' – два множества свободных информационных объектов такие, что

$O \subseteq O'$, то $\eta^O \subseteq \eta^{O'}$, то есть η^O монотонна по своему аргументу.

Доказательство. Проведем доказательство по индукции и покажем, что $\forall k > 0 \mid \eta_k(O) \subseteq \eta_k(O')$. Из $\eta_0(O) = O$ и $\eta_0(O') = O'$ следует $\eta_0(O) \subseteq \eta_0(O')$. Предположив, что $\eta_k(O) \subseteq \eta_k(O')$, для некоторого k применим тот факт, что η – монотонный оператор, и получим, что $\eta(\eta_k(O)) \subseteq \eta(\eta_k(O'))$ или $\eta_{k+1}(O) \subseteq \eta_{k+1}(O')$. Таким образом, $\forall k > 0 \mid \eta_k(O) \subseteq \eta_k(O')$. Пусть k_1 и k_2 – число итераций, необходимых для достижения неподвижных точек, то есть $\eta_{k_1}(O) = \eta^O$, а $\eta_{k_2}(O') = \eta^{O'}$. Определив $k = \max(k_1, k_2)$ из условия $\eta_k(O) \subseteq \eta_k(O')$, получим, что $\eta^O \subseteq \eta^{O'}$.

Доказательство закончено.

Рассмотрим утверждение, определяющее множество свободных информационных объектов O , замкнутое относительно правил I , то есть такое множество O , что $\eta(O) \subseteq O$.

Лемма 1. Для выполнения условия $\eta(O) \subseteq O$ необходимо и достаточно, чтобы $\eta^O = O$.

Доказательство. Для доказательства необходимости отметим, что из $\eta(O) \subseteq O$ и $O \subseteq \eta(O)$ следует $\eta(O) = O$. Тогда, поскольку $\eta_0(O) \subseteq O$ и $\eta_1(O) = \eta(O) = O$, верно, что $\eta_0(O) = O$.

Для доказательства достаточности воспользуемся тем, что $\eta(\eta^O) = \eta^O$ и $\eta(O) = O$, получим $\eta_0(O) = O$.

Доказательство закончено.

В общем случае η^O может оказаться несовместным множеством, причем эта несовместность может быть вызвана либо ограничениями I , либо информационными объектами O . Например, если $O = \{I_a\}$ и $I = \{I_a \sqcup I_b, \neg I_b \sqcup I_a\}$, то получаем несовместное множество $\eta^O = \{I_a, I_b, \neg I_b\}$.

Воспользовавшись введенными обозначениями, уточним некоторые свойства семантики БД.

Определение 3. Семантика множества информационных объектов S определяется как наименьшая неподвижная точка η относительно S в обозначении η^S . Базу данных будем называть совместной, если ее семантика совместна.

Определение 4. Множество ограничений I является совместным, если для всякого информационного объекта $I \in O$ существует множество информационных объектов O' такое, что $\eta^{O' \cup \{I\}}$ является совместным множеством.

Из определений 3 и 4 следует, что I совместно, если для всякого информационного объекта I существует совместная БД, содержащая I . Также имеет место утверждение, характеризующее совместные ограничения.

Утверждение 2. Множество ограничений I является совместным тогда и только тогда, когда для любого информационного объекта I множество η_I является также совместным.

Доказательство. Прямая часть доказательства очевидна, так как при $O = \emptyset$ множество $\eta^{O \cup \{l\}}$ совместно для любого информационного объекта l , откуда следует, что I также совместно.

Для доказательства обратной части предположим, что I совместно. Тогда опираясь на определение совместности множества ограничений (см. определение 2), должно существовать множество свободных информационных объектов O такое, что $\eta^{O \cup \{l\}}$ совместно. Так как $\{l\} \subseteq O \cup \{l\}$, то из утверждения о монотонности оператора η^O (см. утверждение 1) следует, что $\eta_l \subseteq \eta^{O \cup \{l\}}$ и так как $\eta^{O \cup \{l\}}$ совместно, то и η_l также совместно.

Доказательство закончено.

Например, рассмотрим множество $I = \{l_a \downarrow l_{b,c}; \neg l_a \downarrow l_{b,c}\}$. Множество I является совместным, так как для каждого информационного объекта l множество $\eta^l = \{l\}$ совместно. При этом I препятствует тому, что l_a и l_b одновременно входят в O .

В дальнейшем будем рассматривать только совместные множества ограничений, то есть будет предполагаться, что ограничения совместны, даже если об этом явно не будет говориться. Кроме того, будем считать, что ограничения не содержат переменных, то есть БД содержит только свободные информационные объекты, а I является множеством правил (ограничений), в которые входят также только свободные информационные объекты.

3. Меры изменения модифицируемой базы данных

Пусть O – совместное множество свободных информационных объектов некоторой БД и пусть I – ограничения с семантикой S (то есть $S = \eta^O$). Под модификацией O будем понимать операции добавления или удаления информационного объекта l , при которых сохраняется совместность БД. Добавление l означает, что l должен присутствовать в семантике модифицированной БД, а удаление l означает, что l не должен входить в семантику модифицированной БД. Будем обозначать через $ins(l)$ – добавление l , а через $del(l)$ – удаление l .

Для того чтобы модификация была допустима, семантика S' модифицированной БД должна удовлетворять следующим условиям:

1. Множество S' должно быть совместно.
2. Множество S' должно включать l при $ins(l)$, и не должно включать l в случае $del(l)$.
3. Множество S' должно быть дедуктивно замкнуто относительно I .
4. Содержимое множества S' определяется набором информационных объектов, которые были в S (и добавленных объектов l в случае $ins(l)$).
5. Семантика S' должна быть получена с помощью минимального изменения семантики S исходной БД.

Рассматривая множество допустимых семантик модифицируемой БД необходимо отметить, что в

общем случае может оказаться 0, 1 или более множеств S' , удовлетворяющих указанным условиям. В дальнейшем будем рассматривать множества ограничений I , которые допускают только одно множество S' для любой операции $ins(l)$ или $del(l)$.

Первые три условия определяют точные требования для допущения модификации БД, остальные нуждаются в дополнительных комментариях. Уточним некоторые свойства. Четвертое условие говорит, что для того чтобы информационный объект был включен в S' , он должен быть выводим (с использованием правил I) из заданного множества O . Здесь рассматривается O как множество информационных объектов, которые сформировались в БД после модификации.

Например, рассмотрим $O = \{l_a, l_b\}$ и $I = \{l_c \downarrow l_a; l_d \downarrow \neg l_{c,b}\}$. Добавим $\neg l_c$, тогда $S' = \{\neg l_c, l_b, l_d\}$. Действительно, каждый объект из S' можно вывести из $(S' \cap O) \cup \{\neg l_c\} = \{\neg l_c, l_b\}$, где l_b – информационный объект, оставшийся от исходного O , а $\neg l_c$ – добавленный объект.

Для выполнения пятого условия переход от S к S' можно “измерять”, используя симметрическую разность $S \setminus S'$. То есть, $S - S'$ представляет собой множество объектов, удаленных из S во время модификации, а $S' - S$ – это множество объектов, которые добавляются к S . Таким образом, $S \setminus S'$ описывает общее изменение при модификации. Отсюда следует, что для выполнения пятого условия достаточно, чтобы $S \setminus S'$ было минимальным (по отношению к включению множеств). Такой подход нахождения минимальной разности можно найти у ряда авторов [10, 13].

В дальнейшем будем применять меру изменения более точную, чем симметрическая разность, в том смысле, что минимальность изменения влечет минимальность симметрической разности. Подобную меру можно найти в работе [12]. Среди множеств S' , удовлетворяющих условиям 1-4 и минимизирующим $S \setminus S'$, используемая мера выбирает такое, которое минимизирует $S - S'$. То есть такое, при котором удаляется минимально возможное множество информационных объектов из S (или такое, которое сохраняет максимально возможное множество объектов из S).

Использование такой меры позволит оставаться в рамках аксиом изменения, предложенных в [11], одна из которых содержательно утверждает, что при восстановлении совместности БД во время модификации допускается удаление минимально возможного количества информации.

Например, пусть для некоторой БД $O = \{l_b, l_c\}$, $I = \{l_d \downarrow l_{a,c}; l_e \downarrow l_{a,c}; l_k \downarrow l_{a,c}\}$ и $S = \eta^O = \{l_b, l_c\}$, при этом необходимо добавить объект l_c . В этом случае для модифицируемой БД возможны две семантики $S_1 = \{l_a, l_b, l_c, l_d, l_e, l_k\}$ и $S_2 = \{l_a, l_b\}$. Содержательно семантика S_1 получена добавлением объекта l_a к

исходному множеству, а S_2 – путем добавления l_a и удалением l_c из S . В данном примере обе семантики удовлетворяют условиям 1-4, и обе минимизируют симметрическую разность. Но в соответствии с введенным условием минимизации $S_1 - S_2$ предпочтительнее семантика S_1 , так как она сохраняет больше информационных объектов из S .

Определим меру изменения как пару $(S_1 - S_2, S_2 - S_1)$ в обозначении $S_1 \circ S_2$ и зададим частичный порядок \prec на парах множеств информационных объектов как выражение вида (5):

$$(l'_{1...n}, l'_{1...m}) \prec (l''_{1...n}, l''_{1...m}) \Leftrightarrow ((l'_{1...n} \subseteq l''_{1...n}) \text{ или } (l'_{1...n} = l''_{1...n} \text{ и } l'_{1...m} \subseteq l''_{1...m})). \quad (5)$$

Такое определение частичного порядка отдает приоритет минимальности $S_1 - S_2$. При этом справедливо следующее утверждение.

Утверждение 3. Пусть $\mathbf{S} = \{S_1, S_2, \dots, S_n\}$ – семейство множеств возможных семантик, где зафиксирована некоторая семантика S_p и пусть $S_1 \subseteq \mathbf{S}$ – семантика, для которой $S_p \circ S_1$ минимально по отношению частичного порядка “ \prec ” среди остальных $S_p \circ S_i$, где $S_i \subseteq \mathbf{S}$. Тогда $S_p - S_1$ минимально по включению множеств среди всех $S_p - S_i$ и $S_p \setminus S_1$ минимально по включению множеств среди всех $S_p \setminus S_i$.

Доказательство. Докажем первую часть утверждения (минимальность $S_p - S_1$) от противного. Предположим, что $S_p - S_1$ не минимально, тогда существует $S_i \subseteq \mathbf{S}$ такое, что $S_p - S_i \subset S_p - S_1$ таких, что $S_p \circ S_i \prec S_p \circ S_1$. Из того, что $S_p \circ S_1$ минимально по отношению частичного порядка, следует, что $S_p \circ S_i = S_p \circ S_1$, что влечет равенство $S_p - S_i = S_p - S_1$, что противоречит предположению $S_p - S_i \subset S_p - S_1$. Таким образом, можно утверждать, что $S_p - S_1$ минимально.

Для доказательства второй части утверждения предположим, что $S_p \setminus S_i \subseteq S_p \setminus S_1$, тогда имеет место включение $S_p - S_i \subseteq S_p - S_1$ и $S_i - S_p \subseteq S_1 - S_p$. Проверим условие $S_p \circ S_i \prec S_p \circ S_1$. Так как $S_p - S_i \subseteq S_p - S_1$, то возможно два случая: или $S_p - S_i \subset S_p - S_1$, что влечет $S_p \circ S_i \prec S_p \circ S_1$, или $S_p - S_i = S_p - S_1$. Для второго случая используем включения $S_p - S_i \subseteq S_p - S_1$ и $S_i - S_p \subseteq S_1 - S_p$ и также получаем $S_p \circ S_i \prec S_p \circ S_1$. Воспользовавшись минимальностью $S_p \circ S_1$ по условию “ \prec ” сделаем вывод, что $S_p \circ S_i = S_p \circ S_1$. Другими словами, $S_p - S_i = S_p - S_1$ и $S_i - S_p = S_1 - S_p$, то есть $(S_p - S_i) \cup (S_i - S_p) = (S_p - S_1) \cup (S_1 - S_p)$ или $S_p \setminus S_i = S_p \setminus S_1$. Таким образом, $S_p \setminus S_1$ является минимальным по включению множеств.

Доказательство окончено.

В соответствии с заданными свойствами семейства \mathbf{S} (см. утверждение 3) можно утверждать, что из минимальности $S_p \circ S_i$ по отношению частичного порядка следует минимальность $S_p - S_i$ и $S_p \setminus S_i$ по отношению включения множеств (из утверждения о минимальности по включению).

4. Эквивалентность семантик модифицируемой базы данных

Для БД с заданным множеством информационных объектов O и ограничениями целостности I выделим два множества O' и O'' , которые определяются следующими условиями:

1. Множество O' удовлетворяет правилу $l \dashv l_{1...n}$, если $l_{1...n} \in O'$.

2. Множество O' является локальной границей C^I (относительно правил I), если O' совместно и удовлетворяет всем правилам из I .

3. Множество O' порождает информационный объект l в O в обозначении $O' \stackrel{O}{\Rightarrow} l$ (или $O' \Rightarrow l$, если O подразумевается), если $l \in \eta^{O'}$.

4. Множество O' порождает множество информационных объектов O'' в O в обозначении $O' \stackrel{O}{\Rightarrow} O''$ (или $O' \Rightarrow O''$, если O подразумевается), если $O' \stackrel{O}{\Rightarrow} l$ для всех $l \in O''$.

Определение 5. Пусть l – информационный объект и $DB = \langle O, I \rangle$ – совместная БД с семантикой S . Множество информационных объектов S^{ins} будем называть семантикой с добавленным элементом, а S^{del} будем называть семантикой с удаленным элементом для l и O .

Для семантик S^{ins} и S^{del} необходимо, чтобы они были:

- 1) совместны;
- 2) $S \circ S^{ins}$ ($S \circ S^{del}$) были минимальны по отношению “ \prec ”;
- 3) S^{ins} (S^{del}) должны удовлетворять всем ограничениям I .

Также для S^{ins} необходимо выполнение условий:

- 4') $l \in S^{ins}$;
- 5') $(S^{ins} \cap O) \cup \{l\} \Rightarrow S^{ins}$, а для S^{del} выполнение условий:
 - 4'') $l \notin S^{del}$;
 - 5'') $(S^{del} \cap O) \Rightarrow S^{del}$.

Условия 1, 3 и 4' подтверждают, что семантика S^{ins} (S^{del}) является локальной границей C^{ins} (C^{del}), содержащей l . Требование 5' выражает условие 1-4 (см. выше), то есть обоснованность элементов модифицированной семантики, а условие 5'' выражает минимальность изменений БД.

Покажем, что для любого информационного объекта l существует, по крайней мере, одна S^{ins} или S^{del} семантика. Или, другими словами, любое добавление или удаление информационного объекта в совместную БД можно выполнить так, что результат модификации также порождает совместную БД.

Лемма 2. Пусть O – совместное множество свободных информационных объектов некоторой БД с семантикой S , пусть также l информационный

объект из O . Предположим, что две семантики S_1^{ins} (S_1^{del}) и S_2^{ins} (S_2^{del}) удовлетворяют условиям 1, 2, 4' и 5'. Положим $\varpi_1 = S_1^{ins} \cap O$ ($\varpi_1 = S_1^{del} \cap O$) и $\varpi_2 = S_2^{ins} \cap O$ ($\varpi_2 = S_2^{del} \cap O$). Тогда имеют место тождества:

1. $S^\circ S_1^{ins} \prec S^\circ S_2^{ins}$
 $(S^\circ S_1^{del} \prec S^\circ S_2^{del}) \Leftrightarrow \varpi_1 \supseteq \varpi_2$;
2. $S^\circ S_1^{ins} = S^\circ S_2^{ins}$
 $(S^\circ S_1^{del} = S^\circ S_2^{del}) \Leftrightarrow \varpi_1 = \varpi_2$;
3. $S_1^{ins} = \eta^{\varpi_1 \cup l}$; $S_1^{del} = \eta^{\varpi_1}$.

Доказательство. Докажем тождества только для случая добавления информационного объекта, доказательство случая удаления проводится аналогично.

Проведем доказательства для каждого условия отдельно. Докажем третье условие. Отметим, что если $\varpi_1 \cup l \subseteq S_1^{ins}$, то $\eta^{\varpi_1 \cup l} \subseteq \eta^{S_1^{ins}}$. Так как S_1^{ins} замкнуто относительно l , из леммы 4.1 вытекает $\eta^{S_1^{ins}} = S_1^{ins}$, и это дает возможность утверждать, что $\eta^{\varpi_1 \cup l} \subseteq S_1^{ins}$. Поскольку S_1^{ins} удовлетворяет $(S^{ins} \cap O) \cup \{l\} \Rightarrow S^{ins}$ (см. определение 5), то $S_1^{ins} \subseteq \eta^{\varpi_1 \cup l}$. Таким образом, $S_1^{ins} = \eta^{\varpi_1 \cup l}$.

Для первого условия предположим, что $S^\circ S_1^{ins} \prec S^\circ S_2^{ins}$, тогда очевидно, что $S - S_1^{ins} \subseteq S - S_2^{ins}$, а это, в свою очередь, эквивалентно выражению $S \cap S_2^{ins} \subseteq S \cap S_1^{ins}$. Из рассмотренных соотношений вытекает следующее выражение включения $O \cap S \cap S_2^{ins} \subseteq O \cap S \cap S_1^{ins}$. Так как $O \subseteq S$, то есть $O \cap S = O$, то справедлива сокращенная запись предыдущего выражения включения: $O \cap S_2^{ins} \subseteq O \cap S_1^{ins}$. Таким образом, справедливо $\varpi_2 \subseteq \varpi_1$.

Докажем обратное утверждение. Пусть $\varpi_2 \subseteq \varpi_1$, имеем $\varpi_2 \cup \{l\} \subseteq \varpi_1 \cup \{l\}$. По условию 3 данной леммы (доказанного выше), а также, используя свойство неподвижной точки (η монотонна на множестве свободных информационных объектов) получаем $S_2^{ins} \subseteq S_1^{ins}$. Этот факт дает основание утверждать, что $S \cap S_2^{ins} \subseteq S \cap S_1^{ins}$, что эквивалентно $S - S_1^{ins} \subseteq S - S_2^{ins}$. Для подтверждения истинности последнего выражения включения рассмотрим два случая: 1) подмножество $S - S_1^{ins} \subseteq S - S_2^{ins}$ и 2) равенство $S - S_1^{ins} = S - S_2^{ins}$. Из 1-ого случая сразу получаем, что $S^\circ S_1^{ins} \prec S^\circ S_2^{ins}$. Из выражения 2-ого случая получаем, что $S \cap S_1^{ins} = S \cap S_2^{ins}$, из чего, в силу свойства $O \cap S = O$, вытекает равенство $O \cap S_1^{ins} = O \cap S_2^{ins}$. Этот факт дает возможность говорить о равенстве $\varpi_1 = \varpi_2$. Используя условие 3 данной леммы можно отметить, что $S_1^{ins} = S_2^{ins}$, из чего следует, что $S^\circ S_1^{ins} = S^\circ S_2^{ins}$. Таким образом, справедливо выражение $S^\circ S_1^{ins} \prec S^\circ S_2^{ins}$.

Для доказательства второго условия предположим, что $S^\circ S_1^{ins} = S^\circ S_2^{ins}$, тогда имеем $S - S_1^{ins} = S - S_2^{ins}$ или, соответственно, $S \cap S_1^{ins} = S \cap S_2^{ins}$ и $S_1^{ins} - S = S_2^{ins} - S$. Из того, что $S_1^{ins} = (S \cap S_1^{ins}) \cup$

$(S_1^{ins} - S)$ и $S_2^{ins} = (S \cap S_2^{ins}) \cup (S_2^{ins} - S)$, получаем $S_1^{ins} = S_2^{ins}$. Этот факт дает основание утверждать, что $O \cap S_1^{ins} = O \cap S_2^{ins}$, то есть $\varpi_1 = \varpi_2$. Докажем обратное, пусть $\varpi_1 = \varpi_2$, тогда используя условие 3 данной леммы, получим, что $S_1^{ins} = S_2^{ins}$, откуда следует условие $S^\circ S_1^{ins} = S^\circ S_2^{ins}$.

Доказательство закончено.

Введем эквивалентную характеристику для семантик S^{ins} и S^{del} , используя информационные объекты семантики модифицированной БД. Эту характеристику будем рассматривать как дополнительное требование к определению 5.

Утверждение 4. Пусть O – совместное множество свободных информационных объектов некоторой БД с семантикой S , и l информационный объект из O . Пусть также S^{ins} и S^{del} семантики, удовлетворяющие условиям 1, 2, 4' и 5', а также $\varpi = S^{ins} \cap O$ ($\varpi = S^{del} \cap O$). Тогда мера $S^\circ S^{ins}$ ($S^\circ S^{del}$) минимальна по отношению \prec , если $\varpi = S^{ins} \cap O$ ($\varpi = S^{del} \cap O$) максимально по отношению включения множеств.

Доказательство. Введем $\varpi' = (S^{ins})' \cap O$ ($\varpi' = (S^{del})' \cap O$), где $(S^{ins})'$ и $(S^{del})'$ – семантики, удовлетворяющие условиям 1, 2, 4' и 5', и предположим, что $\varpi \subseteq \varpi'$. В соответствии с пунктом 1 леммы 4.2 получаем $S^\circ (S^{ins})' \prec S^\circ S^{ins}$ ($S^\circ (S^{del})' \prec S^\circ S^{del}$), а так как $S^\circ S^{ins}$ ($S^\circ S^{del}$) минимально по отношению \prec , то можно утверждать, что $S^\circ (S^{ins})' = S^\circ S^{ins}$ ($S^\circ (S^{del})' = S^\circ S^{del}$). Используя пункт 2 леммы 4.2, имеем равенство $\varpi = \varpi_1$. То есть ϖ максимально по отношению включения множеств.

Докажем обратное. Предположим, что ϖ максимально и пусть $(S^{ins})'$ и $(S^{del})'$ семантики, удовлетворяющие условиям 1, 2, 4' и 5', для которых выполняются условия $S^\circ (S^{ins})' \prec S^\circ S^{ins}$ и $S^\circ (S^{del})' \prec S^\circ S^{del}$. Положим $\varpi' = (S^{ins})' \cap O$ ($\varpi' = (S^{del})' \cap O$), тогда в соответствии с пунктом 2 леммы 4.2 получаем, что $S^\circ (S^{ins})' = S^\circ S^{ins}$ ($S^\circ (S^{del})' = S^\circ S^{del}$), откуда вытекает, что $S^\circ (S^{ins})'$ ($S^\circ (S^{del})'$) минимально относительно порядка \prec .

Доказательство закончено.

Далее покажем, что множества возможных семантик можно объединить в некоторые непустые семейства $\mathbf{S}^{ins} = \{S_1^{ins}, S_2^{ins}, \dots, S_n^{ins}\}$ и $\mathbf{S}^{del} = \{S_1^{del}, S_2^{del}, \dots, S_n^{del}\}$, то есть для всякого информационного объекта существует, по крайней мере, одна S^{ins} и S^{del} семантика. Рассмотрим утверждение, которое определяет принципы вычисления соответствующих семейств.

Утверждение 5. Пусть DB – некоторая совместная БД, соответствующая виду (2) и l – свободный информационный объект из DB . Тогда справедливы выражения (А) и (Б):

$$\mathbf{S}^{ins} = \{ \eta^{\varpi \cup \{l\}} \mid \max(\varpi) \text{ при } \varpi \subseteq O \text{ и } \eta^{\varpi \cup \{l\}} \text{ – совместно} \} \quad (\text{А})$$

$$\mathbf{S}^{del} = \{ \eta^{\varpi} \mid \max(\varpi) \text{ при } \varpi \subseteq O \text{ и } l \notin \eta^{\varpi} \}. \quad (\text{Б})$$

Доказательство. Доказательство проведем для первого выражение, для второго – доказательство аналогично.

Переобозначим выражение утверждения следующим образом: $\Lambda_1 = \mathbf{S}^{ins}$ и $\Lambda_2 = \{\eta^{\varpi \cup \{l\}} \mid \max(\varpi) \text{ при } \varpi \subseteq O \text{ и } \eta^{\varpi \cup \{l\}} \text{ – совместно}\}$. Покажем равенство $\Lambda_1 = \Lambda_2$ путем доказательства обоюдного вхождения $\Lambda_2 \subseteq \Lambda_1$ и $\Lambda_1 \subseteq \Lambda_2$.

Докажем первое вхождение. Пусть $S^{ins} = \eta^{\varpi \cup \{l\}}$ для некоторого $\max_X(\varpi)$, где $X \subseteq O$. Факт того, что $\varpi \subseteq S^{ins} \cap O$ совместно – очевиден, покажем, что выполняется равенство $\varpi = S^{ins} \cap O$. Докажем этот факт от противного. Пусть $\varpi \neq S^{ins} \cap O$. Предположим, что некоторый $x \in (S^{ins} \cap O) - \varpi$ и $\varpi' = \varpi \cup x$. Так как $\varpi' \subseteq O$ и $\varpi \subset \varpi'$, то ϖ не является максимальным, что противоречит предположению (см. вначале предположение для $\max_X(\varpi)$) и, следовательно, $\varpi = S^{ins} \cap O$.

Отметим, что S^{ins} удовлетворяет условиям 1, 2, 4' и 5'. Для подтверждения истинности условия принадлежности $S^{ins} \subseteq \Lambda_1$ достаточно подтвердить выполнение условия утверждения 4.4 (о максимальнойности $\varpi = S^{ins} \cap O$ относительно включения множеств). Пусть $(S^{ins})'$ – семантика, удовлетворяющая условиям 1, 2, 4' и 5' такая, что $\varpi' = (S^{ins})' \cap O$ и $\varpi \subseteq \varpi'$. Используя пункт 3 леммы 2, можно сделать вывод, что $(S^{ins})'$ совместное множество и $(S^{ins})' = \eta^{\varpi' \cup \{l\}}$. Так как $\varpi \subseteq \varpi' \subseteq O$, то ϖ' максимально по отношению к $\varpi \subseteq O$ и $\eta^{\varpi' \cup \{l\}}$ – совместно, следовательно $\varpi = \varpi'$.

Из вышесказанного можно сделать вывод, что условие о максимальнойности $\varpi = S^{ins} \cap O$ относительно включения множеств (см. утверждения 4) выполняется и $S^{ins} \subseteq \Lambda_1$, откуда следует, что $\Lambda_2 \subseteq \Lambda_1$.

Для доказательства второго вхождения положим, что $S^{ins} \subseteq \Lambda_1$ и пусть $\varpi = S^{ins} \cap O$. В соответствии с пунктом 3 леммы 4.2 имеем $S^{ins} = \eta^{\varpi \cup \{l\}}$, откуда следует, что $\eta^{\varpi \cup \{l\}}$ – совместное множество. Необходимо показать, что ϖ является максимальным подмножеством O , то есть существует $\max_X(\varpi)$, где $X \subseteq O$.

Предположим, что ϖ не является максимальным, тогда существует $\varpi' \subseteq O$ такое, что $\varpi \subset \varpi'$, причем $\eta^{\varpi' \cup \{l\}}$ – совместно. Так как $\varpi \subset \varpi'$, то равенство $\varpi = S^{ins} \cap O$ не удовлетворяет условию о максимальнойности $\varpi = S^{ins} \cap O$ относительно включения множеств (см. утверждение 4), то есть $S^{ins} \notin \Lambda_1$, что является противоречием и, следовательно, $\eta^{\varpi \cup \{l\}}$ удовлетворяет условиям 1, 2, 4' и 5'.

Таким образом, существует $\max_X(\varpi)$, где $X \subseteq O$, для которого $\eta^{\varpi \cup \{l\}}$ совместно. Другими словами, $S^{ins} \subseteq \Lambda_2$, из чего следует, что $\Lambda_1 \subseteq \Lambda_2$.

Доказательство закончено.

5. Корректность модификаций базы данных

Покажем, что для совместной БД допустима любая модификация. Другими словами – существует, по крайней мере, одна S^{ins} и S^{del} семантика и БД с этой семантикой.

Утверждение 6. Для какой-либо совместной БД, соответствующей виду (4.2), и для любого добавления или удаления информационного объекта l существует хотя бы одна S^{ins} и S^{del} семантика.

Доказательство. Проведем доказательство для каждого действия отдельно. Для доказательства существования S^{ins} определим множество $\zeta^{ins} = \{\varpi \mid \varpi \subseteq O \text{ и } \eta^{\varpi \cup \{l\}} \text{ – совместно}\}$. Множество ζ^{ins} непустое, так как множество ограничений I является совместным (см. утверждение 2). Кроме этого отметим, что $\zeta^{ins} \subseteq O$, и поэтому в ζ^{ins} существует максимальный элемент в обозначении ϖ_{\max}^{ins} . Используя свойство семейства \mathbf{S}^{ins} (см. утверждение 5(A)), получаем, что $S^{ins} = \eta^{\varpi_{\max}^{ins} \cup \{l\}}$.

Для доказательства существования S^{del} определим множество $\zeta^{del} = \{\varpi \mid \varpi \subseteq O \text{ и } l \notin \eta^{\varpi}\}$. Множество ζ^{ins} также непустое и $\zeta^{del} \subseteq O$. Максимальный элемент в ζ^{del} обозначим ϖ_{\max}^{del} . Используя свойство семейства \mathbf{S}^{del} (см. утверждение 5(B)), получаем, что $S^{del} = \eta^{\varpi_{\max}^{del}}$.

Доказательство закончено.

Пусть O_1 и O_2 – два множества свободных информационных объектов. Рассмотрим ситуацию, когда $O_1 \subseteq O_2$ при $I_1 = I_2$ и $S_1 \subseteq S_2$. Предположим, что необходимо добавить или удалить некоторый информационный объект в (из) O_1 и/или O_2 . Обозначим через S_1^{ins} семантику модифицированного множества O_1 . Тогда существует семантика S_2^{ins} модифицированного множества O_2 такая, что $S_1^{ins} \subseteq S_2^{ins}$. Этот факт говорит о монотонности семантики модификаций.

Теорема 1. Пусть O_1 и O_2 два совместных множества свободных информационных объектов с одинаковым множеством ограничений I , при этом $O_1 \subseteq O_2$, и пусть l – добавляемый (удаляемый) информационный объект. Тогда для любой семантики $S_1^{ins} \subseteq \mathbf{S}^{ins}$ ($S_1^{del} \subseteq \mathbf{S}^{del}$) существует $S_2^{ins} \subseteq \mathbf{S}^{ins}$ ($S_2^{del} \subseteq \mathbf{S}^{del}$), для которых $S_1^{ins} \subseteq S_2^{ins}$.

Доказательство. Проведем доказательство только для одного случая – добавления информационного объекта, при удалении объекта доказательство проводится аналогично.

Предположим, что $S_1^{ins} \subseteq \mathbf{S}^{ins}$, в соответствии с утверждением 5(A) можно сделать вывод, что существуют множество информационных объектов ϖ' максимальное среди всех подмножеств O_1 , для которого $\eta^{\varpi' \cup \{l\}}$ – совместно и $S_1^{ins} = \eta^{\varpi' \cup \{l\}} \subseteq \mathbf{S}^{ins}$. Тогда из того, что по условию $O_1 \subseteq O_2$, следует, что $\varpi' \subseteq O_2$ и $\eta^{\varpi' \cup \{l\}}$ также совместно.

С другой стороны, так как O_2 конечное множество, то существует некоторое ϖ'' такое, что $\varpi' \subseteq \varpi''$ и ϖ'' является максимальным среди всех подмножеств O_2 и $\eta^{\varpi'' \cup \{l\}}$ – совместно. Используя утверждение 5(A), получаем, что $S_2^{ins} = \eta^{\varpi'' \cup \{l\}} \subseteq \mathbf{S}^{ins}$. Так как $O_1 \subseteq O_2$, то также $\eta^{\varpi' \cup \{l\}} \subseteq \eta^{\varpi'' \cup \{l\}}$, что соответствует выполнению условия $S_1^{ins} \subseteq S_2^{ins}$.

Доказательство закончено.

Выводы

В статье исследованы и предложены методы добавления и удаления информационных объектов в БД, представленной множеством информационных объектов O и множеством правил I . Определены границы модификаций БД, гарантирующие целостность и согласованность неоднородных структур данных. Для обеспечения поддержки неоднородности свойства БД и ограничения целостности описываются логическими выражениями.

Таким образом, автором получены следующие теоретические результаты:

1. На основе введенных логических правил существования (L -правил) введен отдельный класс правил – ограничений целостности.

2. Введены понятия совместности БД и совместности ограничений и на их основе расширено понятие семантики БД. Исследовано поведение семантик при модификации БД.

3. Сформулирован и доказан ряд утверждений, характеризующих свойства информационных объектов и совместности множества ограничений целостности.

4. Определены условия допустимой модификации БД.

5. Сформулировано и доказано утверждение о существовании хотя бы одной семантики из семейства возможных семантик при добавлении и/или удалении информационного объекта.

Практическое использование полученных результатов заключается в осуществлении контроля модификаций структуры БД. Нарушения структуры таблиц может привести к несогласованности данных, что естественным образом отразится на целостности данных. Появление информационных объектов, противоположных по смыслу, на практике может иметь место при приведении различных моделей данных к единому виду, что также влияет на семантику данных.

Следует отметить, что дальнейшее исследование рассмотренного материала следует направить на определения минимальных требований к семантике БД, при которой допустимы изменения ее структуры, а также описания средств реализации запросов к БД с неоднородной структурой.

Список литературы: 1. *Codd, E.* A relational model of data for large shared data banks. [Text] / E. Codd // *SACM* 13. – 1970. – № 6. – p. 1958-1982. 2. *Карденас, А.Ф.* Управление неоднородными распределенными базами данных [Текст] / А.Ф. Карденас // *ТИИЭР*. – 1987. – т. 75, № 5. – с. 72-86. 3. *Дуайер, П.А.,* Ларсон Дж.А. Опыт работы с испытательной распределенной базой данных / П.А. Дуайер, Дж.А. Ларсон // *ТИИЭР*. – 1987. – т. 75, № 5. – с. 126-138.

4. *Sheth, A.P.* Federated database for managing distributed, heterogeneous, and autonomous databases [Text] / A.P. Sheth, J.A. Larson // *Computing Surveys*. – 1990. – № 22(3). – p. 183-236. 5. *Garcia-Solaso, M.* Semantic heterogeneity in multidatabase system [Text] / F. Saltor, M. Castellanos // *In Bukhres and Elmagarmid*. – 1996. – p. 129-195. 6. *Гарсиа-Молина, Г.* Системы баз данных. Полный курс [Текст] / Г. Гарсиа-Молина, Дж. Ульман, Дж. Уидом. Пер. с англ.: – М.: Издательский дом “Вильямс”. – 2003. – 1088 с.: ил. – 3000 экз. – ISBN 5-8459-0384-X. 7. *Танянский, С.* Характеристические свойства объектов информационных систем [Текст] / С. Танянский. // ”Штучний інтелект”: науковий журнал – 2007. – № 1. – с. 78-89. 8. *Лоран, Д.* Детерминированное поддержание ограничений целостности [Текст] / Д. Лоран, Н. Спиратос, Д. Стамат // *Программирование*. – 1998. – № 2. – С. 38-57. 9. *Мейер, Д.* Теория реляционных баз данных [Текст] / Д. Мейер. – М.: Мир – 1987. – 608 с.: ил. – 14000 экз. 10. *Martk, V.* Revision Programming, Database Updates and Integrity Constraints. International Conference on Database Theory [Text] / V.W. Martk, M. Truszczyrski // *M. ICDT. LNCS* 893. – 1995. – p. 368-382. 11. *Alchourron, C.* On the logic of theory change: Partial meet contradiction and revision functions [Text] / C. Alchourron, A. Gardenfors, D. Makinson // *Journal of Symbolic Logic*. – 1985. – v. 50. – p. 510-530. 12. *Dekhtyar, M.* On Conservative Enforced Updates [Text] / M. Dekhtyar, A. Dikovskiy, N. Spyrtatos // 4th International Conference, LPNMR'97, Eds. J.Dix, U.Furbach, A.Nerode, Dagstuhl Castle, Germany, July 28-31, 1997, Lect. Notes in AI (CS), N 1265, – 1997. – p. 244-257. 13. *Borytda, A.* Language features for flexible handling of exceptions in information systems [Text] / A. Borytda // *ACM Trans Database-Syst*. – 1985. – № 10. – p. 565-603.

Поступила в редколлегию 9.09.2010.

УДК 004.047:681.3.01

Підтримка цілісності баз даних при динамічних структурних змінах / С.С. Таняньський // *Біоніка інтелекту: наук.-техн. журнал*. – 2010. – № 3 (74). – С. 112–119.

Розглянуто заснований на логіці підхід до опису і підтримки обмежень цілісності в базах даних при їхніх модифікаціях. Сформульовано умови, при яких модифікація бази даних може бути виконана коректно. Визначено границі модифікацій БД, що гарантують цілісність і погодженість неоднорідних структур даних.

Бібліогр.: 13 найм.

UDK 004.047:681.3.01

Support of integrity of databases at dynamic structural changes / S. Tanyansky // *Bionics of Intelligence: Sci. Mag.* – 2010. – № 3 (74). – P. 112–119.

The approach based on the logician to the description and supports of restrictions of integrity in databases is considered at their updating. Conditions at which database updating can be executed correctly are formulated. Borders of updating of the DB, guaranteeing integrity and a coordination of non-uniform structures of the data are defined.

Ref.: 13 items.

УДК 519.681



ЛОГИКА СЕКВЕНЦИАЛЬНЫХ СХЕМ АЛГОРИТМОВ

С.П. Кашубин

Институт проблем машиностроения НАН Украины,
Харьков, Украина, s.kashubin@gmail.com

Описана новая теория, названная *логикой секвенциальных схем алгоритмов*, которая предназначена главным образом для программирования. В рамках теории исследуются вопросы доказательства правильности алгоритмов, их эквивалентные преобразования и распараллеливание.

АЛГОРИТМ, ПРЕДСТАВЛЕНИЕ ЗНАНИЙ, ДОКАЗАТЕЛЬСТВО ПРАВИЛЬНОСТИ АЛГОРИТМА, ЭКВИВАЛЕНТНЫЕ ПРЕОБРАЗОВАНИЯ, РАСПАРАЛЛЕЛИВАНИЕ, ТЕОРИЯ ПРОГРАММИРОВАНИЯ, СЕКВЕНЦИАЛЬНАЯ СХЕМА

Введение

Чтобы решить задачу, необходимо знать ее *алгоритм*, под которым понимается точная конечная система правил, определяющая совокупность действий над некоторыми объектами (исходными и промежуточными данными) для решения задачи. Иногда вместо слова «алгоритм» мы будем употреблять словосочетание «*схема алгоритма*».

Общее понятие алгоритма было известно еще в средние века, но его точное математическое определение было сформировано лишь в тридцатых годах XX столетия. Дело в том, что в начале XX века накопилось большое количество задач, которые не поддавались решению, несмотря на значительные усилия. Возникло подозрение, что для некоторых из этих задач не существует алгоритма вообще. В связи с такой ситуацией появилась необходимость в точном определении алгоритма. Осознавая эту ситуацию, а также опираясь на исследования в области оснований математики и математической логики, было выработано сразу несколько стандартных математических определений алгоритма. Первые стандартные варианты этого понятия были сформулированы практически одновременно и независимо друг от друга А.Тьюрингом (машина Тьюринга), А.Черчем (лямбда-исчисление) и Э.Постом (машина Поста). Чуть позже С.Клини ввел понятие рекурсивной функции, а А.А.Марков – нормальные алгоритмы. Как оказалось, все эти определения эквивалентны между собой. В настоящее время «алгоритм» является одним из основных и употребительных понятий математики и не только ее.

Процесс программирования обычно начинается с описания области знаний промежуточным языком, дающим возможность построения алгоритмов с последующей реализацией этих алгоритмов в программный код. В этом промежуточном языке понятие «алгоритм» более приспособлено для целей программирования, чем определения, рассмотренные выше. Такие алгоритмы будем называть прикладными алгоритмами, а область знаний, которая изучает их свойства – *прикладной теорией алгоритмов*.

В направлении развития прикладной теории алгоритмов сделаны важные шаги. В 1958 году А.А. Ляпунов [1] и Ю.А. Янов [2] разработали понятие логической схемы алгоритма, а годом позже Л.А. Калужнин ввел понятие граф-схемы алгоритма [3], затем И.Д. Заславский [4] продолжил работы в данном направлении. В.С. Королюк [5] и Е.Л. Ющенко [6] разработали понятие адресного алгоритма. Еще одна модель была предложена И.В. Вельбицким [7] и получила название Р-схемы. Для моделирования программного обеспечения и других систем в 1994 году Г. Бучем, Дж. Румбахом и А. Джекобсоном был создан UML – универсальный язык моделирования [8]. Основная парадигма для современной информатики, характерная для интеллектуальных систем, опирается на идею представления знаний. Сложность и многообразие структур знаний породили ряд способов их представления, таких как логическая модель, фреймовская и продукционная системы, семантические сети. Новая концепция Web также названа семантическими сетями, где знания должны быть описаны при помощи онтологий. Под онтологией (в информатике) понимается описание предметной области в виде, удобном для компьютерной обработки. Данная концепция была предложена и активно развивается комитетом W3C, создателем и руководителем которой является Тим Бернарс-Ли – автор сети World Wide Web. Базовыми языками для онтологий сейчас являются языки RDF, RDFS и OWL.

В 1983 году автором была опубликована работа [9], где описан язык секвенциальных схем (ЯСС). Этот язык использовался при создании и документировании систем программного обеспечения, например, система *konofa*, которая предназначена для приведения формул исчисления высказываний к конъюнктивно нормальной форме [10]. В результате исследований ЯСС был сделан вывод о том, что этот язык необходимо уточнить и доработать. В настоящей работе представлена новая уточненная и дополненная версия языка. Среди дополнений имеются средства логического вывода, что дает основание считать ЯСС логикой. Полное название этой логики – «*логика секвенциальных схем алгоритмов*», а краткое – «*логика S_2* » или просто « *S_2* ».

1. Структура области знаний

Мы будем придерживаться точки зрения, согласно которой знания ориентированы на возможность решения задач, а поскольку задачи решаются по алгоритмам, то и алгоритмы являются знаниями. Языки описания знаний бывают естественными и искусственными (формальными). На естественных языках мы разговариваем с детства, а языки искусственные включают языки программирования, языки моделирования, алгоритмические языки и т.д. Естественные языки могут содержать специфическую символику: математическую, химическую и т.д. Многие языки используются в различных формах, например, в виде устной речи, системы жестов, письменности и т.д., однако, любая форма языка сводится к текстам. Знания, как правило, хранятся в различных библиотеках и Интернете.

Для описания отдельного языка используется метаязык. *Метаязык* — это язык для исследования какого-либо другого языка, который в данном случае называется *объектным языком*. К метаязыкам в информатике относятся стандарты типа Dublin Core, XML, RDF и т.д. Для логики S_2 планируется создание своего собственного метаязыка.

Пусть A и B алгоритмы, где B манипулирует объектами, которые являются схемой алгоритма A или ее частями. Тогда A по отношению к B называется *алгоритмом-объектом*, а B по отношению к A — *метаалгоритмом*. Примерами метаалгоритмов являются программы ИКС-технологии [11], а по отношению к ним программное обеспечение, созданное этой технологией — алгоритмами-объектами.

Все алгоритмы выполняются последовательно или параллельно в среде *абстрактного компьютера (АК)*, который имеет для этих целей необходимое количество исполнителей и памяти. *Исполнителем* может быть как человек, так и некоторое автоматическое устройство. Память состоит из отдельных ячеек. *Ячейка памяти* — это место для записи и хранения информации. АК обладает *вычислительным ресурсом*, состоящим из набора алгоритмов, правильно решающих свои задачи.

1.1. Предметная область

Объектами *предметной области* является все то, что может быть как-то воспринято и на что может быть направлена наша мысль. Объекты используются как данные при решении задач. Предметная область логики S_2 структурирована на основе понятия *типа*, который представляет собой множество объектов. Любые данные, которыми манипулируют алгоритмы, относятся к тому или иному типу.

Для того чтобы объекты и типы можно было использовать в рассуждениях, им следует дать имена. Имена, обозначающие конкретные объекты, называются *константами*, а их совокупность определя-

ет тип. *Переменные* — это имена ячеек памяти, куда разрешено записывать константы одного типа. Другими словами, тип переменной — это область ее определения.

Тип вводится выражения вида:

$$\text{type } T: c_1, c_2, \dots, c_n,$$

где *type* — ключевое слово; T — имя типа; c_1, c_2, \dots, c_n — константы.

Переменная определяется следующим образом:

$$\text{variable } x_1, x_2, \dots, x_m: T,$$

где *variable* — ключевое слово; T — имя типа; x_1, x_2, \dots, x_m — переменные.

Обычно логика предикатов имеет одну предметную область, которая является областью определения для всех возможных переменных. Однако с точки зрения здравого смысла лучше рассматривать предметную область, состоящую из нескольких типов объектов, как это сделано в классической математике и в программировании. Кванторы $\forall x$ и $\exists x$ не обязаны теперь пробегать всю полную предметную область, а только тот тип объектов, какому соответствует переменная x .

1.2. Высказывания и предикаты

Тексты, описывающие некоторую область знаний, включают совокупность *высказываний*, под которым подразумевается смысл некоторого повествовательного предложения. Высказывание *верно (истинно)*, если то, о чем оно повествует, действительно имеет место; и высказывание *неверно (ложно)*, если то, о чем оно повествует, не соответствует действительности.

Простые высказывания объединяются в более сложные при помощи логических связок «или», «и», «нет», «если» и т.д. Самыми простыми высказываниями являются *атомные высказывания* — нерасчленимые высказывания для порождения других.

Высказывания, соединенные связкой «или», называются *дизъюнкцией*. Дизъюнкции бывают двухчленные, трехчленные, четырехчленные и т.д. Если один из членов дизъюнкции является истинным, то и вся дизъюнкция считается истинной, иначе она является ложной.

Конъюнкцией называются высказывания, соединенные связкой «и». По аналогии с дизъюнкцией конъюнкция тоже может быть многочленной. Если все члены конъюнкции являются истинными, то и вся конъюнкция считается истинной, иначе она является ложной.

Высказывание вида «не A » называется *отрицанием*, где A — высказывание, «не» — унарная логическая связка. Отрицание истинное, если A — ложное и отрицание ложное, если A — истинное.

Импликацией называется высказывание вида «если A , то B », где A, B — высказывания, «если ..., то ...» — логическая связка. Импликация является

ложной, если A есть истина, а B – ложь. Во всех остальных сочетаниях истинностных значениях A и B импликация принимает значение истина.

Эквивалентность (читается « A эквивалентно B ») имеет значение «истина» тогда, когда совпадают значения высказываний A и B . Во всех остальных случаях эквивалентность считается ложной.

С высказываниями тесно связаны *предикаты*, под которыми будем понимать шаблоны текста, состоящие из постоянной части и имен типов для подстановки других имен. Предикаты дают осмысленные высказывания при замене имен типов константами того же типа. Примеры предикатов: «число ЧИСЛА простое», « $2 + \text{ЧИСЛА} = 8$ », «ЧЕЛОВЕК любит ЧЕЛОВЕК», где тип ЧИСЛА – обозначает целые числа, тип ЧЕЛОВЕК – обозначает людей. Предикаты можно обозначать с помощью постоянной части текста, за которой сразу идет перечень имен типов в круглых скобках, например, для предыдущих примеров: «простое число(ЧИСЛА)», «два плюс нечто равно восемь(ЧИСЛА)», «любит(ЧЕЛОВЕК, ЧЕЛОВЕК)». Предикат истинен, если при замене имен типов на константы таких же типов получается истинное высказывание, и предикат ложен, если при аналогичных заменах получается ложное высказывание.

Предикат мы будем определять при помощи следующего выражения:

$$\text{predicate: } P(T_1, T_2, \dots, T_n),$$

где P – имя постоянной части предиката; T_1, T_2, \dots, T_n – имена типов; $n \geq 0$ – порядок (арность) предиката.

Бинарный предикат $P(T_1, T_2)$ можно изобразить в виде $(T_1 P T_2)$ или же вообще без скобок. 0 -арный предикат $P()$ обозначает некоторое высказывание, которое, как правило, записывается в виде P .

По определению произвольный предикат $P(T_1, T_2, \dots, T_n)$ разрешим на основе связанной с ним таблицы 1.

Таблица 1

T_1	T_2	...	T_n
a_{11}	a_{12}	...	a_{1n}
a_{21}	a_{22}	...	a_{2n}
\vdots	\vdots		\vdots
a_{k1}	a_{k2}	...	a_{kn}

В колонках таблицы записаны константы соответствующих типов, а в строках – константы, при которых предикат превращается в истинное высказывание, если заменить этими константами соответствующие типы в $P(T_1, T_2, \dots, T_n)$. Если типы заменить значениями, не являющимися строками таблицы, то получившееся высказывание будет ложным.

Предикаты изучаются в логике предикатов, которая представляет собой универсальный абстрактный язык описания знаний и является неотъемлемой частью фундамента современной математики. В этом смысле логика S_2 строится как специальная логика предикатов для описания, изучения и манипулирования алгоритмами.

2. Формулы

Алфавит языка S_2 содержит основные, специальные и вспомогательные символы. Основные символы – это буквы естественного языка, специальные символы – логические связки, вспомогательные – это разделители типа скобок и запятых. Из основных символов составляются слова, обозначающие константы, переменные, предикаты и другие понятия S_2 .

Пусть $P(T_1, T_2, \dots, T_n)$ – предикат, а v_1, v_2, \dots, v_n – константы и переменные, относящиеся соответственно к типам T_1, T_2, \dots, T_n , тогда выражение $P(v_1, v_2, \dots, v_n)$, полученное подстановкой v_1, v_2, \dots, v_n вместо T_1, T_2, \dots, T_n , называется *атомарной формулой* или *атомом*. Атомы служат «кирпичиками» для построения более сложных формул при помощи логических связок и скобок. Связками могут быть: \neg – «нет», $\&$ – «и», \vee – «или», \supset – «если ..., то ...», \sim – «эквивалентно».

Логические формулы, или просто формулы определим по индукции:

1. Атомарная формула есть формула.
2. Если A, B – формулы, то выражения $\neg A, (A \vee B), (A \& B), (A \supset B), (A \sim B)$

также являются формулами.

Совокупность различных переменных формулы называются ее *параметрами* или *аргументами*. Любая формула выражает некоторый предикат, зависящий от своих параметров.

Иногда для удобства восприятия большую и сложную формулу необходимо изобразить в более коротком виде. Сделать это можно при помощи следующего формального определения:

$$A(x_1, x_2, \dots, x_n) ::= \text{«определяемая формула»},$$

где A – имя определяемой формулы; x_1, x_2, \dots, x_n – параметры определяемой формулы, ::= – символ «равно по определению». Введенное обозначение формулы можно записывать без параметров и скобок, т.е. в виде A , если это не приведет к недоразумениям.

Разрешение формулы осуществляется по индукции:

1. Если формула является атомарной, то она разрешима по определению.
2. Если истинностные значения формул A и B известны, то истинность формул

$$\neg A, (A \vee B), (A \& B), (A \supset B), (A \sim B)$$

определяются по правилам для связок.

Формула соответствует *логической функции*, областью определения которой являются все возможные значения ее параметров, а область значений – истина или ложь. В матлогике логическая функция обычно задается таблицей истинности.

Пусть имеем *type* $T: c_1, c_2, \dots, c_k$ и *variable* $x: T$, тогда допустимы следующие два формальных определения:

$$\forall x A(x) ::= A(c_1) \& A(c_2) \& \dots \& A(c_k),$$

$$\exists x A(x) ::= A(c_1) \vee A(c_2) \vee \dots \vee A(c_k).$$

Символ \forall называется *квантором общности*, а символ \exists – *квантором существования*. Формулы $\forall x A(x)$ и $\exists x A(x)$ могут быть использованы для построения других формул, при этом вхождение переменной x в построенную формулу называется *связанным*.

Формула A называется *общезначимой* (символически $\vDash A$), тогда когда она является истинной для всех допустимых распределений значений ее параметров. При помощи выражения $\Rightarrow A$ мы будем обозначать тот факт, что формула A является истинной для текущих значений своих параметров.

Мы говорим, что формула B является следствием из формул A_1, A_2, \dots, A_m (это символически записывается в виде $A_1, A_2, \dots, A_m \models B$) тогда, когда выполняется следующее условие: если все A_1, A_2, \dots, A_m принимают значение «истина», то B также принимает значение «истина». Выражение $A_1, A_2, \dots, A_m \models B$ назовем *логической секвенцией*. Понятие секвенции в математической логике впервые ввел Г. Генцен [12] для формализации логического вывода.

Можно доказать, что для логики S_2 секвенция $A_1, A_2, \dots, A_m \models B$ разрешима. Другими словами – можно построить алгоритм, который определяет истинностное значение этой секвенции.

Пусть $\Gamma \models B_1, \Gamma \models B_2, \dots, \Gamma \models B_n$ – логические секвенции, тогда по отношению к ним можно применить такое определение как

$$\Gamma \models B_1, B_2, \dots, B_n ::= \Gamma \models B_1, \Gamma \models B_2, \dots, \Gamma \models B_n.$$

Для списка формул введем следующее определение:

$$\Delta(x_1, x_2, \dots, x_n) ::= A_1, A_2, \dots, A_k,$$

где $\Delta(x_1, x_2, \dots, x_n)$ – обозначение списка формул; A_1, A_2, \dots, A_k – список формул для $k \geq 1$; x_1, x_2, \dots, x_n – объединение параметров формул A_1, A_2, \dots, A_k . Параметры x_1, x_2, \dots, x_n называются *параметрами списка формул* Δ .

Если $\Rightarrow A_1, \Rightarrow A_2, \dots, \Rightarrow A_k$ – список истинных формул для $k \geq 1$, то тогда имеем определение:

$$\Rightarrow A_1, A_2, \dots, A_k ::= \Rightarrow A_1, \Rightarrow A_2, \dots, \Rightarrow A_k.$$

Пусть $\Delta(x_1, x_2, \dots, x_n)$ – список формул с параметрами x_1, x_2, \dots, x_n . Тогда будем считать, что выражение $\Rightarrow \Delta(x_1, x_2, \dots, x_n)$ определяет со-

стояние памяти для x_1, x_2, \dots, x_n . В S_2 списку $\Delta(x_1, x_2, \dots, x_n)$ соответствует таблица с колонками x_1, x_2, \dots, x_n и со строками, для которых справедливо $\Rightarrow \Delta(x_1, x_2, \dots, x_n)$.

3. Вычислительные секвенции

Наиболее общей моделью алгоритма (также как и другой системы) является модель черного ящика, который изображен на рис.1. В этом случае алгоритм представляется в виде прямоугольника, внутреннее устройство которого скрыто или неизвестно для внешнего наблюдателя. Алгоритм взаимодействует с внешней средой посредством своих входов и выходов так, что состояние выходов функционально зависит от состояний входов.



Рис. 1

Алгоритм кратко описывается своей *спецификацией*, под которой будем понимать совокупность *вычислительных секвенций*, представляющих собой выражения вида:

$$\Lambda(x_1, x_2, \dots, x_n) \xrightarrow{r} \Delta(y_1, y_2, \dots, y_m),$$

где r – имя спецификации алгоритма; список формул $\Lambda(x_1, x_2, \dots, x_n)$ называется *условием секвенции*; список формул $\Delta(y_1, y_2, \dots, y_m)$ называется *заключением секвенции*. Параметры условия x_1, x_2, \dots, x_n называются *входами секвенции (алгоритма)*, а параметры заключения y_1, y_2, \dots, y_m – *выходами секвенции (алгоритма)*. Условие и заключение можно рассматривать как конъюнкции их формул. *Условие секвенции выполняется*, если имеет место $\Rightarrow \Lambda(x_1, x_2, \dots, x_n)$; это влечет за собой дальнейшее выполнение алгоритма. Если же условие не выполняется, то алгоритм также не выполняется.

Имя спецификации состоит из двух частей: первая часть написана строчными буквами, вторая – начинается с заглавной буквы, а все остальные строчные. Первая часть имени совпадает с именем алгоритма, а вторая часть – выбирается произвольно. Иногда имя спецификации над стрелкой мы будем опускать, если это не приведет к недоразумениям. Введем определение:

$$\rightarrow (y_1, y_2, \dots, y_m) ::= r(x_1, x_2, \dots, x_n) ::=$$

$$\Lambda(x_1, x_2, \dots, x_n) \xrightarrow{r} \Delta(y_1, y_2, \dots, y_m),$$

где символ $::=$ называется *символом присвоения*. Такую секвенцию допускается изображать без стрелки, т.е. в виде выражения

$$(y_1, y_2, \dots, y_m) ::= r(x_1, x_2, \dots, x_n).$$

Вычислительную секвенцию, представленную в форме с символом присвоения, назовем *секвенцией присвоения*.

В логике S_2 имеются секвенция такая как

$$\Delta(x) \xrightarrow{\text{copy}} \Delta(y) \text{ или } \rightarrow y := \text{copy}(x),$$

которая копирует значение x в y . Для подобных секвенций будем использовать следующее определение:

$$\rightarrow y := x ::= \rightarrow y := \text{copy}(x).$$

Вычислительная секвенция с одним выходом называется *функциональной секвенцией*.

Пусть $\rightarrow y := f(x_1, x_2, \dots, x_n)$ — функциональная секвенция присвоения, тогда $f(x_1, x_2, \dots, x_n)$ назовем *функцией*. Считается, что 0-арная функция — это константа [13].

В классической логике предикатов используется понятие *терма*, который по смыслу соответствует математическому выражению. Терм в S_2 вводится по индукции следующим образом:

1. Переменные, константы и функции являются термами.

2. Если x — переменная, t и v — произвольные термы, $t|_v^x$ — результат замены x на v в t , то тогда $t|_v^x$ также является термом.

4. Секвенциальные схемы алгоритмов

Пара вычислительных секвенций называется *независимой*, если она обладает следующими свойствами:

1. Входы одной из них не содержатся среди выходов другой, однако, входы и выходы одной и той же секвенции могут совпадать.

2. Все выходы секвенций различаются между собой.

Совокупность вычислительных секвенций r_1, r_2, \dots, r_n является *независимыми секвенциями*, если любая пара из них независима.

Пусть r_1, r_2, \dots, r_n — независимые секвенции, тогда выражение

$$p: [r_1, r_2, \dots, r_n]$$

называется *блок-объединением*, где p — имя этого объединения, $n \geq 0$. Блок-объединение будет записываться в одну строку (по необходимости с переносом). Считаем, что блок-объединение определяет совокупность алгоритмов, выполняемых параллельно.

При выполнении блок-объединения его секвенции могут быть выполнены все, частично или даже ни одна из них. Те секвенции, которые выполняются до конца, назовем *активными секвенциями*. Блок-объединение, содержащий активные секвенции, называется *активным блок-объединением*. Состояние памяти AK может изменяться только под воздействием активных блок-объединений. Совокупность активных секвенций активного блок-объединения получила название *активной строки*.

4.1. Алгоритмические блоки

Полное описание алгоритма в логике S_2 представлено *секвенциальной схемой алгоритма*, которая состоит из отдельных частей, названных (*алгоритмическими*) *блоками*. Вся секвенциальная схема также есть блок, а те блоки, которые в него входят, называются *внутренними блоками* или *подблоками*. Для подблока тот блок, куда он входит, называется *внешним блоком*. В общем случае блок обозначается выражением вида

$$p: [T],$$

где p — имя блока, T — внутренние блоки. Квадратные скобки и имена отдельных блоков можно опускать, если это не приведет к недоразумениям. Блок может быть пустым.

В состав некоторых блоков кроме вычислительных секвенций войдут еще и *управляющие секвенции*, которые представляют собой такие выражения как

$$(\Delta \downarrow p) \text{ или } (\Delta \uparrow p),$$

где p — имя блока, Δ — список формул (он может быть пустым). Выражение $(\Delta \downarrow p)$ получило название *break-секвенции*, а $(\Delta \uparrow p)$ — *continue-секвенции*. Список формул Δ является *условием управляющей секвенции*, понятие которой совпадает с понятием условия вычислительной секвенции. Совокупность параметров списка формул Δ являются *входами управляющей секвенции*.

Алгоритмический блок (секвенциальную схему алгоритма) определим так:

1. Блок-объединение есть блок.

2. Если $[K], [M]$ — блоки, тогда выражение

$p: \left[\begin{array}{l} [K] \\ [M] \end{array} \right]$ также является блоком, и называется он *композицией блоков*.

3. Если после одного из подблоков блока p вставить управляющую секвенцию, то полученное выражение есть блок, и называется он *управляемым блоком*.

Из определения блока следует, что секвенциальная схема состоит из последовательности строк вычислительных и управляющих секвенций, структурированных квадратными скобками. Управляющая секвенция всегда образует одну строку.

Выполнение алгоритмического блока осуществляется в соответствии со следующим правилом:

1. Объединения секвенций — его выполнение рассмотрено выше.

2. Композиция выполняется последовательно сверху вниз.

3. Если при выполнении управляемого блока p очередь дошла до break-секвенции $(\Delta \downarrow p)$ и Δ есть истина, то тогда произойдет выход из блока p . Если Δ ложное, то тогда будет выполняться компонент секвенциальной схемы, следующий за $(\Delta \downarrow p)$.

4. Если при выполнении управляемого блока p очередь дошла до continue-секвенции $(\Delta \uparrow p)$ и Δ есть истина, то тогда блок p начнет выполняться с самого начала. Если Δ ложное, то тогда будет выполняться компонент секвенциальной схемы, следующий за $(\Delta \uparrow p)$.

Пример 1 (алгоритмический блок)

$$\left[\begin{array}{l} y_1 := x + y; \quad y_2 := z - t \quad y_3 := x - z \\ y_4 := y_3 + y; \quad y_5 := y_1 + y_2 \\ y_6 := y_4 * y_5 \end{array} \right].$$

Этот блок вычисляет выражение

$$(((x + y) + (z - t)) * ((x - z) + y))$$

и относится к типу *линейных блоков*, под которыми понимаются блоки без управляющих секвенций.

4.2. Информационный граф блока

Для анализа информационных взаимосвязей линейного блока можно использовать орграф со взвешенными дугами. *Орграф* или *ориентированный граф* – это упорядоченная пара $G = (V, D)$, где V – множество *вершин* или *узлов*, D – множество дуг. *Дуга* – это упорядоченная пара вершин графа (v, w) , где вершина v называется началом, а w – концом дуги. Можно сказать, что дуга $v \rightarrow w$ ведет от вершины v к вершине w . Граф является взвешенным, если множеству дуг поставлена в соответствие некоторая величина, называемая *весом* дуги. Тот факт, что дуга (v, w) имеет вес u , запишем с помощью упорядоченной тройки (u, v, w) . Таким образом, орграф с взвешенными дугами – это упорядоченная пара $G = (V, D)$, где V – множество *вершин* или *узлов*, D – множество взвешенных дуг (u, v, w) .

Интерпретация алгоритмического блока взвешенным орграфом осуществляется по следующему правилу:

1. Блок индексируется так, что между его вычислительными секвенциями и индексами устанавливается взаимно-однозначное соответствие.
2. Индексы секвенций – это вершины орграфа.
3. Пусть s и r – две секвенции, а u – выход секвенции s , который одновременно является входом секвенции r . Тогда тройка (u, i_s, i_r) – взвешенная дуга орграфа, где i_s, i_r – индексы секвенций s и r .

Граф, построенный по этому правилу, назовем *информационным графом блока*, а его пример показан на рис. 2.

Граф на рис. 2 имеет три уровня: $\{1, 2, 3\}$, $\{4, 5\}$, $\{6\}$, каждый из которых соответствует одному блок-объединению.

4.3. Производные блоки

Те блоки, которые были введены выше, а именно объединение, композиция и управляемый блок, назовем *основными блоками*, а блоки, которые мы сейчас определим, назовем *производными*. Производные блоки, как правило, моделируют извест-

ные конструкции программирования, такие как оператор цикла, условный оператор и другие.

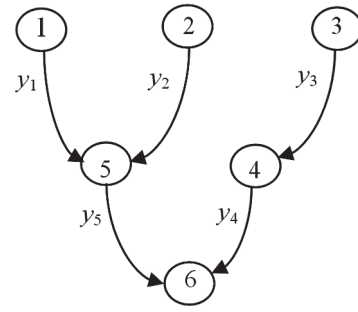


Рис. 2. Информационный граф для блока из примера 1

While-цикл с постусловием определяется следующим образом:

$$[T](A) ::= p: \left[\begin{array}{l} [T] \\ (A \uparrow p) \end{array} \right],$$

где A – формула под названием «*условие цикла*»; $[T]$ – блок под названием «*тело цикла*».

Пусть $[T](A)$ – *while-цикл с постусловием*, тогда *while-цикл с предусловием* определится так:

$$(A)[T] ::= r: \left[\begin{array}{l} (\neg A \downarrow r) \\ [T](A) \end{array} \right].$$

Цикл *for* соответствует такому блоку как

$$\left[\begin{array}{l} i := a \\ A(i) \left[\begin{array}{l} T \\ i := f(i) \end{array} \right] \end{array} \right],$$

где $i := a$ – инициализация цикла; $A(i)$ – условие цикла; $i := f(i)$ – шаг цикла; T – тело цикла.

Условный блок if-then вводится следующим образом:

$$r: [A \rightarrow [T]] ::= r: \left[\begin{array}{l} [] \\ (\neg A \downarrow r) \\ [T] \end{array} \right],$$

где $[T]$ – блок, $A \rightarrow [T]$ – условный блок *if-then*; A – формула, которая называется *условием блока*. Блок $[T]$ выполняется, когда условие A есть истина.

Оператор *if-then-else* моделируется при помощи следующей композиции двух условных блоков *if-then*:

$$p: \left[\begin{array}{l} A \rightarrow \left[\begin{array}{l} [K] \\ (\downarrow p) \end{array} \right] \\ \neg A \rightarrow [M] \end{array} \right].$$

Эту композицию можно сократить на основе такого определения:

$$A \rightarrow [K][M] ::= p: \left[\begin{array}{l} A \rightarrow \left[\begin{array}{l} [K] \\ (\downarrow p) \end{array} \right] \\ \neg A \rightarrow [M] \end{array} \right].$$

Возникает вопрос, какие алгоритмы можно представить секвенциальными схемами? Ответ на него дает теорема, приведенная ниже.

Теорема 1 (о полноте). При помощи секвенциальной схемы можно реализовать любой алгоритм.

Доказательство:

1. «Алгоритм любой сложности можно реализовать, используя только три конструкции: следования, выбора и повторения» (Теорема Дейкстры).
2. Конструкция следования соответствует композиции блоков S_2 , конструкции выбора и повторения определены как производные блоки (см. выше).
3. Утверждение теоремы (пп 1, 2).

5. О доказательстве в логике S_2

Обычно в программировании правильность программ проверяется путем тестирования, однако это не дает нужного эффекта – ошибки могут выявляться на протяжении всей жизни программы. Выдающийся ученый в области информатики Э. В. Дейкстра говорил, что программист должен использовать теорию, необходимую для подтверждения алгоритма. В этом смысле можно сказать, что такой теорией является логика S_2 , где делается попытка представить рассуждения относительно алгоритмов в виде формального доказательств.

Алгоритм является правильным, если его секвенциальная схема построена из правильных вычислительных секвенций, и существует доказательство спецификации этого алгоритма. Доказательство вычислительных секвенций строится на основе *схемных правил*, а спецификации – на основе *правил изменения памяти и логических правил*. Все виды правил являются фигурами заключения генценовского типа, приспособленные для логики S_2 . Последовательность построения секвенциальной схемы правильного алгоритма такова: вначале доказываются все секвенции схемы, затем из них строится сама схема, и наконец доказываются ее спецификации.

5.1. Доказательство секвенций схемы алгоритма

Рассмотрим некоторые из схемных правил.

Пусть $\Lambda \xrightarrow{r} \Delta$ – вычислительная секвенция, $\rightarrow x := t$ – функциональная секвенция присвоения, t – терм, $\Lambda|_t^x$ – результат замены переменной x на t в списке формул Λ . Тогда справедливо правило замены входа секвенции, сформулированное ниже.

Правило замены входа секвенции

$$\frac{\rightarrow x := t; \Lambda \xrightarrow{r} \Delta}{\Lambda|_t^x \xrightarrow{r} \Delta}.$$

На основе этого правила можно получить секвенцию, где в формулах условия содержатся произвольные термы.

Пусть $\Lambda \xrightarrow{r} \Delta$ – вычислительная секвенция, x – параметр списка формул Δ , v – переменная такого же типа, как и x , $\Delta|_v^x$ – результат замены x на v в Δ . Тогда имеет место правило замены выхода (см. ниже).

Правило замены выхода

$$\frac{\Lambda \xrightarrow{r} \Delta}{\Lambda \xrightarrow{r} \Delta|_v^x}.$$

В схемных правилах, приведенных ниже, используются следующие обозначения: Δ , Γ , Ψ – списки формул, A , B – одиночные формулы.

Правило перестановки в условии

$$\frac{\Delta, A, B, \Gamma \xrightarrow{r} \Psi}{\Delta, B, A, \Gamma \xrightarrow{r} \Psi}$$

Правило перестановки в заключении

$$\frac{\Delta \xrightarrow{r} \Psi, A, B, \Gamma}{\Delta \xrightarrow{r} \Psi, B, A, \Gamma}$$

Правило сокращения в условии

$$\frac{\Delta, A, A, \Gamma \xrightarrow{r} \Psi}{\Delta, A, \Gamma \xrightarrow{r} \Psi}$$

Правило сокращения в заключении

$$\frac{\Delta \xrightarrow{r} \Psi, A, A, \Gamma}{\Delta \xrightarrow{r} \Psi, A, \Gamma}$$

Правило введения конъюнкции в условии

$$\frac{\Delta, A, B, \Gamma \xrightarrow{r} \Psi}{\Delta, A \& B, \Gamma \xrightarrow{r} \Psi}$$

Правило введения конъюнкции в заключении

$$\frac{\Delta \xrightarrow{r} \Psi, A, B, \Gamma}{\Delta \xrightarrow{r} \Psi, A \& B, \Gamma}$$

Правило введения дизъюнкции в условии

$$\frac{A, \Delta \xrightarrow{r} \Psi; B, \Delta \xrightarrow{r} \Psi}{A \vee B, \Delta \xrightarrow{r} \Psi}$$

Правило удаления импликации в заключении

$$\frac{\Delta \xrightarrow{r} A; \Delta \xrightarrow{r} A \supset B}{\Delta \xrightarrow{r} B}$$

1-е правило утончения

$$\frac{\Gamma \xrightarrow{r} \Psi}{\Delta, \Gamma \xrightarrow{r} \Psi}$$

2-е правило утончения

$$\frac{\rightarrow (y_1, y_2, \dots, y_m) := r(t_1, t_2, \dots, t_n)}{\Delta \rightarrow (y_1, y_2, \dots, y_m) := r(t_1, t_2, \dots, t_n)}$$

где t_1, t_2, \dots, t_n – термы.

Выражение под чертой 2-го правила утончения также является секвенцией присвоения. Для этого вида секвенций может быть полезным следующее определение:

$$\Delta \rightarrow (y_1, y_2, \dots, y_n) := (t_1, t_2, \dots, t_n) ::=$$

$$\Delta \rightarrow y_1 := t_1, \Delta \rightarrow y_2 := t_2, \dots, \Delta \rightarrow y_n := t_n,$$

где t_1, t_2, \dots, t_n – термы.

Конечная последовательность вычислительных секвенций s_1, s_2, \dots, s_n называется *доказательством*

секвенции схемы алгоритма s_n , если эта последовательность удовлетворяет следующему условию: каждая секвенция $s_i, i \leq n$ либо является спецификацией алгоритма из вычислительного ресурса AK , либо получена по одному из схемных правил и некоторых секвенций $s_j, j < n$.

5.2. Доказательство спецификации алгоритма

Алгоритм изменяет свою память по шагам. На очередном шаге находится активный блок-объединение, который сразу выполняется, что приводит к изменению состояния памяти. Работу алгоритма со спецификацией $\Delta \rightarrow \Omega$ можно описать последовательностью состояний памяти

$$\begin{aligned} &\Rightarrow \Delta_1 \\ &\Rightarrow \Delta_2 \\ &\dots \\ &\Rightarrow \Delta_m, \end{aligned}$$

где $\Rightarrow \Delta_1$ – начальное состояния памяти (Δ_1 совпадает с Δ), $\Rightarrow \Delta_m$ – конечное состояние памяти. Алгоритм удовлетворяет спецификации, если имеет место $\Delta_m \models \Omega$. Алгоритм является правильным, если он удовлетворяет всем своим спецификациям.

Активный блок-объединение изменяет состояние памяти по правилу, приведенному ниже.

Правило изменения памяти

$$\frac{\Rightarrow \Delta; \Delta \vdash \Lambda_1, \Lambda_2, \dots, \Lambda_n; \Lambda_1 \rightarrow \Gamma_1; \Lambda_2 \rightarrow \Gamma_2; \dots; \Lambda_n \rightarrow \Gamma_n}{\Rightarrow \exists y_1 \exists y_2 \dots \exists y_m (\Delta), \Gamma(y_1, y_2, \dots, y_m)},$$

где $\Rightarrow \Delta$ – состояние памяти до выполнения блока, $\Lambda_1 \rightarrow \Gamma_1, \Lambda_2 \rightarrow \Gamma_2, \dots, \Lambda_n \rightarrow \Gamma_n$ – активные секвенции, $\Gamma(y_1, y_2, \dots, y_m)$ – объединение формул из заключенных активных секвенций, y_1, y_2, \dots, y_m – объединение выходов активных секвенций, $\exists y_1 \exists y_2 \dots \exists y_m (\Delta)$ – формулы, полученные из списка Δ присваиванием перед каждой его формулой кванторов $\exists y_1 \exists y_2 \dots \exists y_m$, выражение под чертой правила – это состояние памяти после выполнения блока. Лишние кванторы в списке формул $\exists y_1 \exists y_2 \dots \exists y_m (\Delta)$ можно будет сократить, и в результате останутся только кванторы перед формулами, содержащими связанные с ними переменные. Если список формул Δ не содержат переменных y_1, y_2, \dots, y_m , то тогда это правило запишется в виде правила изменения памяти.

Правило изменения памяти

$$\frac{\Rightarrow \Delta; \Delta \vdash \Lambda_1, \Lambda_2, \dots, \Lambda_n; \Lambda_1 \rightarrow \Gamma_1; \Lambda_2 \rightarrow \Gamma_2; \dots; \Lambda_n \rightarrow \Gamma_n}{\Rightarrow \Delta, \Gamma_1, \Gamma_2, \dots, \Gamma_n}.$$

Доказательство спецификации алгоритма $\Delta \rightarrow \Omega$ представляет собой конечную последовательность

$$\begin{aligned} &\Rightarrow \Delta_1 \\ &\Rightarrow \Delta_2 \\ &\dots \\ &\Rightarrow \Delta_m \\ &\Delta_m \models \Omega, \end{aligned}$$

где $\Rightarrow \Delta_1, \Rightarrow \Delta_2, \dots, \Rightarrow \Delta_m$ – состояния памяти, которые пробегает алгоритм в процессе его выполнения, $\Rightarrow \Delta_1$ – начальное состояния памяти, $\Rightarrow \Delta_m$ – конечное состояние памяти. Последовательность состояний памяти должна удовлетворять следующим условиям:

1. Δ_1 совпадает с Δ .
2. Каждое состояние памяти $\Rightarrow \Delta_i, 1 < i \leq m$ получено по правилу изменения памяти.
3. Логическая секвенция $\Delta_m \models \Omega$ доказывается при помощи логических правил классической математической логики.

6. Эквивалентность алгоритмов

Секвенция $\Lambda(x_1, x_2, \dots, x_n) \rightarrow \Delta(y_1, y_2, \dots, y_m)$ и секвенция $\Gamma(x_1, x_2, \dots, x_n) \rightarrow \Omega(y_1, y_2, \dots, y_m)$ называются эквивалентными, если для одинаковых значений входов x_1, x_2, \dots, x_n их выходам y_1, y_2, \dots, y_m всегда будут присваиваться одинаковые значения. Формально эквивалентность секвенций изобразится при помощи такого выражения как

$$\begin{aligned} \Lambda(x_1, x_2, \dots, x_n) \rightarrow \Delta(y_1, y_2, \dots, y_m) \approx \\ \Gamma(x_1, x_2, \dots, x_n) \rightarrow \Omega(y_1, y_2, \dots, y_m), \end{aligned}$$

где \approx – символ эквивалентности вычислительных секвенций (спецификаций).

Из определения понятия эквивалентности секвенций вытекает следующая теорема.

Теорема 2. Если для двух вычислительных секвенций справедливо $\Lambda \rightarrow \Delta \approx \Gamma \rightarrow \Omega$, то тогда $\Lambda \models \Gamma$ и $\Gamma \models \Lambda$.

Следствие. Если для двух вычислительных секвенций справедливо $\Lambda \rightarrow \Delta \approx \Gamma \rightarrow \Omega$, то тогда $\Lambda \rightarrow \Omega$ и $\Gamma \rightarrow \Delta$.

Два алгоритма эквивалентны, если между множествами их спецификаций существует такое взаимно-однозначное соответствие, где каждая пара – эквивалентные секвенции. Эквивалентность алгоритмов обозначается выражением $P \approx R$, где P, R – алгоритмы, \approx – символ эквивалентности алгоритмов, который совпадает с символом эквивалентности спецификаций.

Приведем без доказательства несколько теорем, которые могут быть использованы для эквивалентных преобразований секвенциальных схем.

Теорема 3. Пусть T, R – блок-объединения такие, что R отличается от T лишь порядком расположения секвенций, тогда $T \approx R$.

Теорема 4. Пусть TA – секвенциальная схема, в которой выделено вхождение формулы A , а TB – секвенциальная схема, полученная из TA заменой A на формулу B . Тогда, если $A \sim B$, то $TA \approx TB$.

Теорема 5. Пусть TA – секвенциальная схема, в которой выделено вхождение блока A , а TB – секвенциальная схема, полученная из TA заменой A на блок B . Тогда, если $A \approx B$, то $TA \approx TB$.

Теорема 6. Пусть $[\Psi_S]$ – блок-объединение, в котором выделена последовательность вычислительных секвенций S . Тогда $[\Psi_S] \approx \begin{bmatrix} [\Psi] \\ [S] \end{bmatrix}$ и $[\Psi_S] \approx \begin{bmatrix} [S] \\ [\Psi] \end{bmatrix}$, где $[\Psi]$ – блок-объединение, в котором удалена последовательность секвенций S , $[S]$ – блок-объединение, состоящий из последовательности секвенций S .

Следствие. Пусть $[S_1, S_2, \dots, S_n]$ – блок-объединение, который состоит из вычислительных сек-

венций S_1, S_2, \dots, S_n , тогда $[S_1, S_2, \dots, S_n] \approx \begin{bmatrix} S_1 \\ S_2 \\ \dots \\ S_n \end{bmatrix}$.

Теорему 6 и следствие из нее можно использовать для распараллеливания алгоритмов.

Выводы

Новая логика S_2 является специальной логикой предикатов для изучения и манипулирования прикладными алгоритмами. Ее можно использовать для проектирования знаниеориентированных интеллектуальных систем программного обеспечения. В логике создается механизм формального доказательства правильности алгоритмов и их эквивалентности. Алгоритмы можно распараллеливать. Схемы алгоритмов удобно редактировать в стандартных редакторах формул, таких как MathType.

Список литературы: 1. *Ляпунов, А.А.* О некоторых общих вопросах кибернетики [Текст] / А.А. Ляпунов // Сб. «Проблемы кибернетики». – М.: Физматгиз, 1958. – Вып. 1. – С. 5-22. 2. *Янов, Ю.И.* О логических схемах алгоритмов [Текст] / Ю.И. Янов // Сб. «Проблемы кибернетики». – М.: Физматгиз, 1958. – Вып. 1. – С. 46-74. 3. *Калужнин, Л.А.* Об алгоритмизации математических задач // Сб. «Проблемы кибернетики». – М.: Физматгиз, 1959. – Вып. 2. – С. 51-58. 4. *Заславский, И.Д.* Граф-схемы с памятью [Текст] / И.Д. Заславский // Сб. «Труды математического института им. В.А. Стеклова». – М.-Л.: изд-во «Наука», 1964. – Вып. LXXII. – С. 99-192. 5. *Королюк, В.С.* О по-

нятии адресного алгоритма [Текст] / В.С. Королюк // Сб. «Проблемы кибернетики». – М.: Физматгиз, 1960. – Вып. 4. – С. 95-110. 6. *Ющенко, Е.Л.* Адресное программирование [Текст] / Е.Л. Ющенко. – К.: Изд-во «Техническая литература», 1963. – 288 с. 7. *Вельбицкий, И.В.* Технология программирования [Текст] / И.В. Вельбицкий. – К.: Техника, 1984. – 279 с. 8. *Рамбо, Дж.* UML: специальный справочник [Текст] / Дж. Рамбо, А. Якобсон, Г. Буч. – СПб.: Питер, 2002. – 656 с. 9. *Кашубин, С.П.* Язык секвенциальных схем алгоритмов [Текст] / С.П. Кашубин; Ин-т пробл. Машиностроения АН УССР. – Харьков, 1983. – 20 с. – Деп. В ВИНТИ 05.07.83, №3963. 10. *Кашубин, С.П.* Приведение формул исчисления высказываний к конъюнктивно нормальной форме [Текст] / С.П. Кашубин; Ин-т пробл. Машиностроения АН УССР. – Харьков, 1986. – 30 с. – Деп. В ВИНТИ 18.06.86, №4483-В86. 11. *Кашубин, С.П.* ИКС-технология (создание интегрированных компьютерных систем) [Текст] / С.П. Кашубин // Автоматизированные системы управления и приборы автоматки. 2006. Вып. 135. С. 47-50. 12. Математическая теория логического вывода: Сб. переводов. – М.: изд-во «Наука», 1967. – 352 с. 13. *Шенфильд, Дж.* Математическая логика: Пер. с англ. – М.: Изд-во «Наука», 1975. – 528 с. 14. *Дейкстра, Э.* Дисциплина программирования: Пер. с англ. – М.: Мир, 1978. – 280 с.

Поступила в редколлегию 10.09.2010.

УДК 519.681

Логіка секвенціальних схем алгоритмів / С.П. Кашубін // Біоніка інтелекту: наук.-техн. журнал. – 2010. – № 3 (74). – С. 120–128.

Запропоновано нову теорію під назвою «Логіка секвенційних схем алгоритмів», яка призначена головним чином для програмування. В статті досліджуються питання доказу правильності алгоритмів, їх еквівалентні перетворення та розпаралелювання.

Табл. 1. Іл. 2. Бібліогр.: 14 найм.

UDC 519.681

The logic of sequential schemes of algorithms / S.P. Kashubin // Bionics of Intelligence: Sci. Mag. – 2010. – № 3 (74). – P. 120–128.

A new theory called “The logic of sequential schemes of algorithms”, designed mainly for programming, is proposed. The article investigates the proof of the correctness of algorithms, their equivalent transformations and parallelization.

Tabl. 1. Fig. 2. Ref.: 14 item.

УДК 681.518:004.93.1'



ОПТИМІЗАЦІЯ ПАРАМЕТРІВ ПЛАНУ НАВЧАННЯ СИСТЕМИ РОЗПІЗНАВАННЯ МАГНІТОКАРДІОГРАМ

А.С. Довбиш¹, С.С. Мартиненко²

¹СумДУ, м. Суми, Україна, kras@id.sumdu.edu.ua

²СумДУ, м. Суми, Україна, smart@unesco.sumdu.edu.ua

У рамках інформаційно-екстремальної інтелектуальної технології розглядається оптимізація кроку зміни кута кола зчитування яскравості магнітокардіограми у полярній системі координат у процесі навчання системи розпізнавання. Ефективність запропонованого алгоритму досліджено на прикладі розпізнавання магнітокардіограм, одержаних за довільних початкових умов.

СИСТЕМА РОЗПІЗНАВАННЯ, МАГНІТОКАРДІОГРАМА, ІНФОРМАЦІЙНО-ЕКСТРЕМАЛЬНА ІНТЕЛЛЕКТУАЛЬНА ТЕХНОЛОГІЯ, ПОЛЯРНІ КООРДИНАТИ

Вступ

Одним із сучасних перспективних методів дослідження стану серцево-судинної системи є магнітокардіографія, яка має вищу діагностичну інформативність у порівнянні з електрокардіографічними зображеннями [1, 2]. Але розпізнавання магнітокардіограм вимагає від лікаря-кардіолога високого професійного рівня та тривалого практичного досвіду. Тому актуальною є задача автоматизації розпізнавання магнітокардіограм. На жаль, відомі результати, наведені, наприклад, у працях [3, 4], не дозволяють одержати необхідну для практичного застосування достовірність розпізнавання. Основними причинами такого стану є:

- відсутність ефективних алгоритмів розпізнавання, що обумовлено довільними початковими умовами формування зображень;
- застосування неоптимальних в інформаційно-розрахунковій частині параметрів стиснення та оброблення магнітокардіограм, які суттєво впливають на функціональну ефективність системи розпізнавання.

Одним із перспективних напрямів аналізу і синтезу систем розпізнавання, що навчаються, є використання ідей і методів інформаційно-екстремальної інтелектуальної технології (ІЕІ-технологія), що ґрунтується на максимізації інформаційної спроможності системи шляхом введення в процесі її навчання додаткових інформаційних обмежень [5-8]. При цьому важливим етапом розпізнавання магнітокардіограм, від якого залежить достовірність розпізнавання, у рамках ІЕІ-технології є формування багатовимірної апріорної у загальному випадку нечіткої навчальної матриці шляхом стиснення та оброблення відеоінформації. Функціональні параметри системи розпізнавання, що формують навчальну матрицю, відносяться до групи параметрів плану навчання, оскільки вони впливають як на структуру алгоритму навчання, так і на функціональну ефективність навчання системи [9, 10]

Одним із підходів до формування навчальної матриці є використання полярної системи коор-

динат [11, 12], яка забезпечує інваріантність алгоритму розпізнавання до геометричних деформацій образу.

У статті запропоновано інформаційно-екстремальний алгоритм оптимізації кроку квантування кута кола зчитування яскравості в пікселях рецепторного поля магнітокардіограм.

1. Постановка задачі

Нехай дано магнітокардіограми, які характеризують різні функціональні стани серцево-судинної системи пацієнта і утворюють алфавіт класів розпізнавання $X_m^o, m = \overline{1, M}$. Відомий вектор-кортеж параметрів функціонування $g_m = \langle x_m, \delta, \Delta_\phi \rangle$, де x_m – еталонний вектор-реалізація, який визначає геометричний центр контейнера класу X_m^o , побудованого в радіальному базисі простору ознак розпізнавання; δ – параметр поля контрольних допусків на ознаки розпізнавання; Δ_ϕ – крок квантування кута кола зчитування яскравості у відповідних пікселях рецепторного поля магнітокардіограми. При цьому у рамках ІЕІ-технології задано такі обмеження $x_m \in [0; d(x_m \oplus x_c) - 1]$, де x_c – еталонний вектор-реалізація класу X_c^o , який є найближчим до класу $X_m^o, \delta \in [0; \delta_H/2]$, де δ – нормоване (експлуатаційне) поле допусків на ознаки розпізнавання

На етапі навчання системи розпізнавання необхідно знайти оптимальні значення координат вектора параметрів функціонування g_m , які забезпечують максимальне значення інформаційного критерію функціональної ефективності (КФЕ) навчання системи розпізнавання

$$\bar{E}_{\max} = \frac{1}{M} \sum_{m=1}^M \max_G E_m, \quad (1)$$

де E_m – КФЕ навчання системи розпізнавати реалізацію класу X_m^o ; G – область допустимих значень параметрів функціонування.

При цьому з метою побудови на етапі навчання безпомилкових вирішальних правил за оптималь-

ним параметром Δ_ϕ^* сформувати вхідну навчальну матрицю $\|y_{m,i}^{(j)}\|, i=1, \overline{N}, j=1, \overline{n}$, де N, n – кількість ознак розпізнавання в словнику і реалізацій класу X_m^o відповідно.

На етапі екзамени треба прийняти рішення про належність реалізацій, що розпізнаються, до одного із класів заданого алфавіту.

2. Алгоритм оптимізації кроку квантування Δ_ϕ

Як параметр плану навчання системи розпізнавання будемо розглядати крок квантування RGB-складових зображень за кутом кола зчитування яскравості в пікселях рецепторного поля квантування. При цьому оброблення зображень в полярних координатах здійснюється за формулою

$$\Theta_j = \frac{1}{N_j} \sum_{i=1}^{N_j} \theta_i, \quad (2)$$

де Θ_j – усереднене значення яскравості пікселів, що входять у коло зчитування j -го радіусу, $j = \overline{0, R}$; θ_i – значення яскравості кольорової складової в i -му пікселі; N_j – загальна кількість пікселів в j -му колі зчитування; R – радіус кола зчитування.

Координати пікселя зчитування одержують за формулами $x = r \cos(\phi + \Delta_\phi)$ і $y = r \sin(\phi + \Delta_\phi)$, де ϕ – кут кола зчитування. Аналіз формули (2) показує, що як параметр плану навчання можна обрати крок квантування кола зчитування.

На рис. 1 показано розподіл пікселів на рецепторному полі магнітокардіограм, значення яких зчитувалося у випадках, коли зміна кута зчитування дорівнювала $\Delta_\phi = 1^\circ$ (0,017 рад), $\Delta_\phi = 5^\circ$ (0,087 рад.), $\Delta_\phi = 10^\circ$ (0,174 рад.) і $\Delta_\phi = 15^\circ$ (0,261) відповідно.

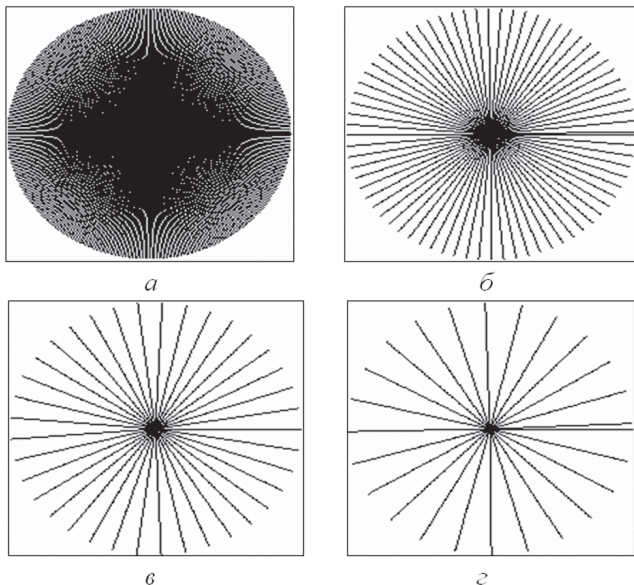


Рис. 1. Розподіл пікселів зображень при обробленні в полярних координатах при різних кроках кутів зчитування:

$a - \Delta_\phi = 1^\circ$, $b - \Delta_\phi = 5^\circ$, $v - \Delta_\phi = 10^\circ$, $z - \Delta_\phi = 15^\circ$

Аналіз рис. 1 показує, що щільність пікселів, з яких зчитується інформація, є нерівномірною. Найбільшою вона є в центрі зображення при $\Delta_\phi = 1^\circ$. Таким чином, відкидання розташованих на периферії рецепторного поля пікселів може зменшити достовірність розпізнавання. З іншого боку, пікселі, які зчитуються при мінімальному кроці квантування кута кола зчитування, можуть давати неінформативні або заважаючі ознаки, що так само знижує достовірність розпізнавання. Тобто маємо задачу параметричної оптимізації.

Після формування вхідної навчальної матриці для заданого алфавіту класів розпізнавання $\{X_m^o | m = \overline{1, M}\}$ запускається інформаційно-екстремальний алгоритм навчання системи розпізнавання з оптимізацією параметра плану навчання Δ_ϕ за такою ітераційною структурованою процедурою

$$\Delta_\phi^* = \arg \min_{G_\phi} \max_{G_\delta} \max_{G_E} \{E\}, \quad (3)$$

де G_ϕ, G_δ, G_E – області допустимих значень крока квантування Δ_ϕ , параметра поля контрольних допусків δ і робоча область визначення функції КФЕ (1).

Ентропійний (за Шенноном) нормований КФЕ навчання системи розпізнавання, наприклад, реалізацій класу X_m^o , може мати вигляд [5]:

$$E_m^{(k)} = 1 + \frac{1}{2} \left(\frac{\alpha_m^{(k)}(d)}{\alpha_m^{(k)}(d) + D_{2,m}^{(k)}(d)} \log_2 \frac{\alpha_m^{(k)}(d)}{\alpha_m^{(k)}(d) + D_{2,m}^{(k)}(d)} + \frac{\beta_m^{(k)}(d)}{D_{1,m}^{(k)}(d) + \beta_m^{(k)}(d)} \log_2 \frac{\beta_m^{(k)}(d)}{D_{1,m}^{(k)}(d) + \beta_m^{(k)}(d)} + \frac{D_{1,m}^{(k)}(d)}{D_{1,m}^{(k)}(d) + \beta_m^{(k)}(d)} \log_2 \frac{D_{1,m}^{(k)}(d)}{D_{1,m}^{(k)}(d) + \beta_m^{(k)}(d)} + \frac{D_{2,m}^{(k)}(d)}{\alpha_m^{(k)}(d) + D_{2,m}^{(k)}(d)} \log_2 \frac{D_{2,m}^{(k)}(d)}{\alpha_m^{(k)}(d) + D_{2,m}^{(k)}(d)} \right), \quad (4)$$

де $\alpha_m^{(k)}(d)$ – помилка першого роду прийняття рішення на k -му кроці навчання; $\beta_m^{(k)}(d)$ – помилка другого роду; $D_{1,m}^{(k)}(d)$ – перша достовірність; $D_{2,m}^{(k)}(d)$ – друга достовірність; d – дистанційна міра, яка визначає радіуси гіперсферичних контейнерів, побудованих в радіальному базисі простору ознак розпізнавання.

3. Приклад розпізнавання магнітокардіограм

Як вхідні зображення розглядалися магнітокардіограми трьох класів (рис. 2), які характеризували нормальний стан серцево-судинної системи пацієнта, ішемічну хворобу та шуми в серці.

Оброблення магнітокардіограм цих класів здійснювалося в полярних координатах за формулою (2), що забезпечує інваріантність алгоритму розпізнавання до зсуву, повороту та зміни масштабу зображень. У результаті було одержано спектри

яскравості для трьох RGB-складових кольорових зображень магнітокардіограм. На рис. 3 показано спектри яскравості RGB-складових, одержані при обробленні магнітокардіограми класу X_2^o , що характеризує ішемічну хворобу серця.

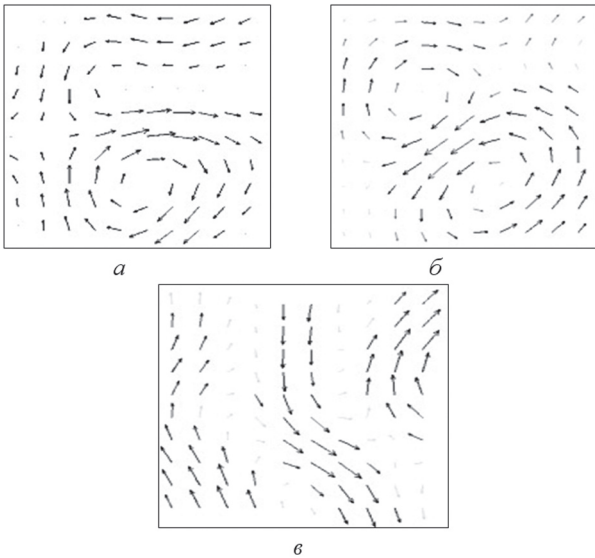


Рис. 2. Зображення магнітокардіограм: *a* — нормальний стан (клас X_1^o); *б* — ішемія (клас X_2^o); *в* — шуми в серці (клас X_3^o)

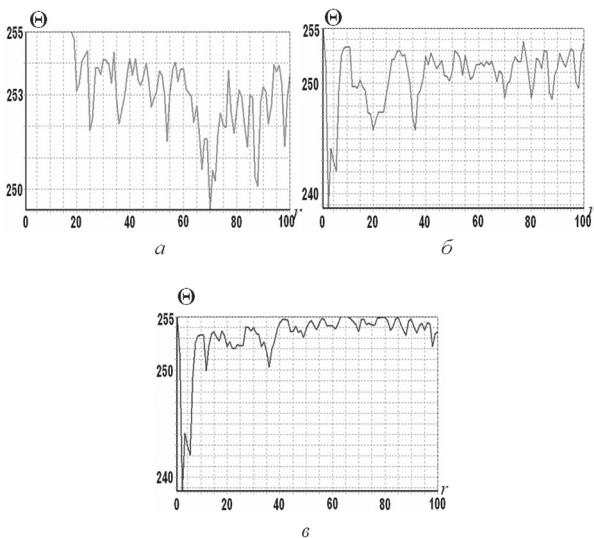


Рис. 3. Спектри яскравості для класу X_2^o : *a* — червона складова; *б* — зелена складова; *в* — синя складова

Дискретні спектри RGB-складових магнітокардіограм формують структурований вектор-реалізацію відповідного класу. Після формування навчальної матриці $\|y_{m,i}^{(j)}\|$, рядками якої є вектори-реалізації класу X_m^o , а стовпчиками — навчальна вибірка випадкових значень *i*-ої ознаки, було реалізовано алгоритм навчання (3) для значень кроку квантування кута зчитування, наведених на рис. 1. На рис. 4 подано динаміку зміни значення усеред-

неного критерію функціональної ефективності від параметру плану навчання — кроку квантування кута зчитування.

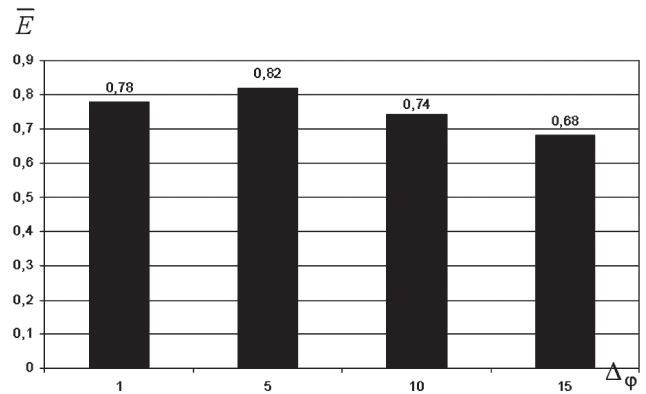


Рис. 4. Графік зміни усередненого КФЕ від значення кроку квантування кута зчитування

Аналіз рис. 4 показує, що усереднений критерій функціональної ефективності досягає свого максимуму при значенні кроку квантування $\Delta\varphi = 5^\circ$. При цьому усереднене значення КФЕ дорівнює $\bar{E}^* = 0,82$. На рис. 5 показано графіки залежності КФЕ від радіусу контейнера для кожного з класів при оптимальному параметрі плану навчання $\Delta\varphi = 5^\circ$.

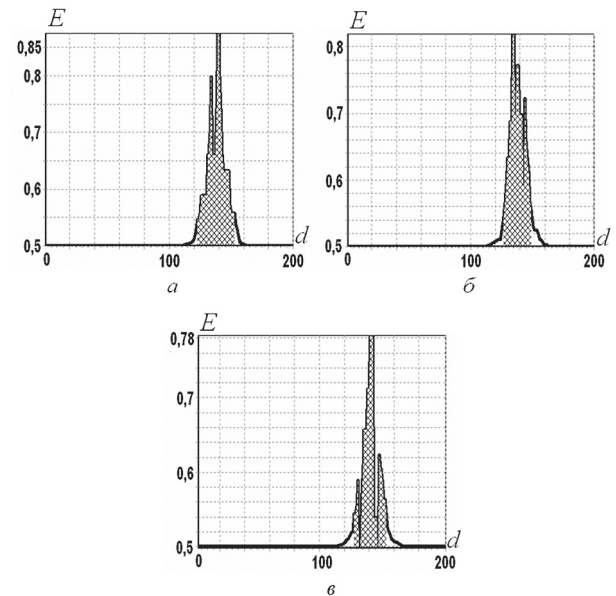


Рис. 5. Графік залежності КФЕ від радіусів контейнерів класів: *a*) клас X_1^o ; *б*) клас X_2^o ; *в*) клас X_3^o

У цьому випадку оптимальними геометричними параметрами контейнерів класів розпізнавання при визначених в процесі оптимізації еталонних векторах-реалізаціях відповідних класів є: для класу X_1^o оптимальний радіус контейнера дорівнює $d^* = 143$ ($E_1^* = 0,86$); для класу X_2^o — $d^* = 126$ ($E_2^* = 0,82$) і для класу X_3^o — $d^* = 132$ ($E_3^* = 0,78$).

Після побудови на етапі навчання оптимальних вирішальних правил було реалізовано алгоритм

екзамену, який полягав у обчисленні функцій належності реалізації образу, що розпізнається, до класів із заданого алфавіту [6]. Віднесення реалізації, що розпізнається, до відповідного класу здійснювалося за максимальним значенням функції належності. Априорно було відомо, що реалізація, що розпізнавалася на екзамені, належала класу X_1^o (рис. 6) і відрізнялася від аналогічної реалізації, показаної на рис. 2.

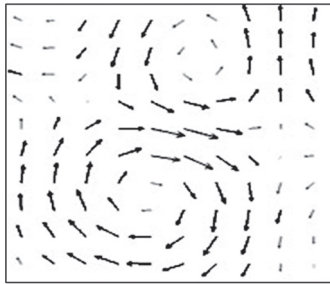


Рис. 6. Зображення класу X_1^o

У результаті було встановлено, що зображення, проказане на рис. 6, дійсно належить до класу X_1^o , оскільки функція належності для цього класу є додальною і більшою, ніж для інших класів

$$\bar{\mu}_1^* = \max_{\{m=1,2,3\}} \bar{\mu}_m = \max \left\{ \begin{matrix} 39,22; & -41,57; & -22,90 \\ x_1^o & x_2^o & x_3^o \end{matrix} \right\}.$$

Таким чином, оптимізація параметрів плану навчання у рамках ІЕІ-технології дозволяє суттєво підвищити достовірність розпізнавання зображень.

Висновки

1. Розроблено та досліджено алгоритм оптимізації кроку квантування кута кола зчитування яскравості у пікселях рецепторного поля, що дозволило підвищити асимптотичну достовірність розпізнавання магнітокардіограм.

2. Оброблення вхідних зображень у полярній системі координат на етапі формування вхідної вибірки дозволяє забезпечити інваріантність алгоритму розпізнавання до зсуву, повороту та зміни масштабу зображень.

Список літератури: 1. В. Hailer, P. Van Leeuwen, S. Lange, I. Chaikovsky, D. Gronemeyer. Magnetocardiography in hypertensive and coronary artery disease. International Congress Series 2007; 1300 : 488-451. 2. Embrechts M., Szymanski B., Sternickel K., Naenna T., Bragaspathi R. Use of machine learning for classification of magnetocardiograms. Proc. IEEE Conference on System, Man and Cybernetics; Washington DC, October 2003. p. 1400-5. 3. Fenici R., Brisinda D., Meloni AM. Effects of Filtering on Computer-Aided Analysis for Detection of Chronic Ischemic Heart Disease with Unshielded Rest Magnetocardiographic Mapping. Neurology and Clinical Neurophysiology (November 30, 2004) 2004:7. 4. Tantimongcolwat T., Naenna T., Isarankura-Na-

Ayudhya C., Embrechts M.J., Prachayasittikul V. Identification of ischemic heart disease via machine learning analysis on magnetocardiograms, Computers in Biology and Medicine, 38 (7), 2008, pp. 817-825. 5. Краснопоясовський, А.С. Інформаційний синтез інтелектуальних систем керування, що навчаються: Підхід, що ґрунтується на методі функціонально-статистичних випробувань [Текст] / А.С. Краснопоясовський. – Суми: Видавництво СумДУ. – 2003. – 257 с. 6. Довбиш, А.С. Основи проектування інтелектуальних систем: Навчальний посібник [Текст] / А.С. Довбиш. – Суми: Видавництво Сум ДУ, 2009. – 171 с. 7. Довбиш, А.С. Інформаційно-екстремальний метод розпізнавання електрограм [Текст] / А.С. Довбиш, С.С. Мартиненко // Вісник СумДУ. Технічні науки. – 2009. – №2. – С. 85-92. 8. Довбиш, А.С. Інтелектуальна система діагностування онкопатологій [Текст] / А.С. Довбиш, О.П. Чекалов, С.С. Мартиненко // Радіоелектронні і комп'ютерні системи. – Харків «ХАІ» 2009. – №3. – С. 78-83. 9. Дмитриев, А. К. Основы теории построения и контроля сложных систем [Текст] / А.К. Дмитриев, П.А. Мальцев. – Л.: Энергоатомиздат. Ленингр. отд-ние, 1988. – 192 с. 10. Краснопоясовський, А.С. Оптимізація кроку дискретизації реалізації образу в інтелектуальних системах [Текст] / А.С. Краснопоясовський, С.О. Заговора, М.В. Козинець // Управляющие системы и машины. – 2003. – № 1. – С. 42-47. 11. Анисимов, Б.В. Распознавание и цифровая обработка изображений [Текст] / Б.В. Анисимов, В.Д. Курганов, В.К. Злобин. – М.: Высшая школа, 1983. – 256 с. 12. Довбиш, А.С. Інтелектуальна система підтримки прийняття рішень для діагностування онкопатологій [Текст] / А.С. Довбиш, В.О. Жаловага // Вісник СумДУ. Серія: Технічні науки. – 2009. – № 3. – С. 5-12.

Надійшла до редколегії 10.09.2010.

УДК 681.518:004.93.1'

Оптимизация параметров плана обучения системы распознавания магнитокардиограмм / А.С. Довбыш, С.С. Мартиненко // Бионика интеллекта: науч.-техн. журнал. – 2010. – № 3 (74). – С. 129–132.

Рассматривается в рамках информационно-экстремальной интеллектуальной технологии оптимизация шага изменения угла окружности считывания яркости магнитокардиограммы в полярной системе координат в процессе обучения системы распознавания. Эффективность предложенного алгоритма исследована на примере распознавания магнитокардиограмм, полученных при произвольных начальных условиях.

Ил. 6. Библиогр.: 6 назв.

UDC 681.518:004.93.1'

Optimization of parameters of learning plan of magnetocardiograms recognition system / A.S. Dovbysh, S.S. Martynenko // Bionics of Intelligence: Sci. Mag. – 2010. – № 3 (74). – P. 129–132.

Article presents results of optimization of parameter of angle changing of brightness scanning circle of magnetocardiogram in polar coordinates in process of recognition system learning in bounds of information-extreme intellectual technology. Efficiency of algorithm is researched on example of magnetocardiograms recognition, gained in random starting conditions.

Fig. 6. Ref.: 6 items.

УДК 519.62



СЕРВІС-ОРІЄНТОВАНА МОДЕЛЬ ПЕРСОНАЛІЗОВАНОГО НАВЧАННЯ З ВИКОРИСТАННЯМ АГЕНТНИХ ТЕХНОЛОГІЙ

Н.С. Лесна¹, І.М. Келеберда², С.М. Шрестха³, Ю.А Кирпусь⁴

^{1–3} ХНУРЕ, м. Харків, Україна;

⁴ ХНУРЕ, м. Харків, Україна, kirpus_yuriy@mail.ru

Описано та проаналізовано побудову сервіс-орієнтованої моделі персоналізованого навчання з використанням агентних технологій. Основною функцією моделі є обробка та доставка потрібної користувачеві інформації.

ПРОГРАМНИЙ АГЕНТ, МОДЕЛЬ, ВЕБ-СЕРВІС, МУЛЬТИАГЕНТНІ СИСТЕМИ, JADE, XML

Вступ

Величезний обсяг інформації, який існує в Інтернеті на сьогоднішній день, і необхідність його обробки призводять до розвитку технологій та систем пошуку інформації.

Також зростає і актуальність проблеми випередження запиту користувача шляхом пропонування йому потенційно цікавої інформації. Цю проблему вирішують системи персоналізації на базі програмних агентів, які пропонують користувачеві потенційно цікаву для нього інформацію, без явного запиту з його боку.

В даній статті наведено дослідження процесу персоналізованого пошуку навчальних матеріалів в мережі Інтернет та створення мультиагентної моделі персоналізованого навчання шляхом надання учневі персоналізованих навчальних матеріалів.

Схожі задачі були частково розкриті в роботах [1]–[3]. Але дані роботи мало уваги приділяють саме процесу персонального пошуку та відображенню навчальних матеріалів.

1. Дослідження джерел інформації

В процесі дослідження було виявлено найефективніші джерела і сервіси для пошуку навчальної інформації, а саме:

- RSS-канали, які містять найактуальнішу інформацію;
- сервіс Google Scholar, який надає пошук з наукових джерел (статті, публікації та ін.);
- сервіс rip1.com, який дозволяє здійснювати пошук людей та їх контакти.

1.1. RSS – канали

RSS — сімейство XML-форматів, призначених для опису сторінок новин, анонсів статей, змін в блогах і т.п. Інформація з різних джерел, представлена у форматі RSS, може бути зібрана, оброблена і представлена користувачеві в зручному для нього вигляді спеціальними програмами-агрегаторами.

RSS-канал — це безпосередньо файл у форматі RSS. Цей файл розташовується на сервері. У цьому файлі зберігаються останні новини (звичайно близько 10–15 штук). Він оновлюється в міру появи новин. Як і у будь-якого документа, розташованого на сервері, у RSS- файла є адреса [4].

1.2. Сервіс — rip1.com

Rip1.com — один з наймолодших проектів, призначений виключно для пошуку інформації про людей. Даний сервіс, пов'язаний з так званою концепцією «Невидимого Інтернету» [5].

Механізм rip1.com, за твердженням його творців, здобуває інформацію саме з «тіньової» частини Інтернету. Скануючи он-лайн бази даних і анкети різних соціальних мереж, форумів, персональних сторінок і т. д., сервіс збирає інформацію про користувачів. Тому на його стартовій сторінці не одна форма для введення запиту, а цілих п'ять — для прізвища, імені, міста, штату і країни. Недолік, як у багатьох, — відсутність підтримки кирилиці.

1.3. Сервіс для пошуку з наукових джерел — Google Scholar

Google Scholar — сервіс для пошуку з наукових джерел статей, книжок, дисертацій, опублікованих різними науковими організаціями і професійними спільнотами.

Google Scholar являє собою вільний доступ до пошуку Web Engine, який індексує повний текст наукової літератури через масив видавничих форматів і дисципліни. Випущений в листопаді 2004 року індекс Google Scholar включає в себе найбільше наукове видавництво найбільш рецензованих он-лайн журналів Європи і Америки. Сервіс схожий на функцію вільно доступних Scirus від Elsevier, CiteSeerX і getCITED, та на Scopus Elsevier's і Thomson ISI 'S веб-науки.

Google Scholar дозволяє користувачам здійснювати пошук по цифровим або фізичним копіям статей: он-лайнних або в базах даних електронних бібліотек. Google Scholar порівняно швидкий та легкий у використанні. «Наукові» результати пошуку з'являються за допомогою посилання на повні тексти статей журналу, технічні доповіді, дисертації, книги та інші документи, в тому числі обрані веб-сторінки, які вважаються науковими.

Через свої функції Google Scholar забезпечує доступ до описів статей, які посилаються на статтю, яку користувач розглядає на даний момент. Саме ця функція, зокрема, передбачається індексацією цитування раніше доступних тільки в Scopus і Web of Knowledge [6].

Google Scholar відтворює індексування сканованих повнотекстових журналів, зміст яких надають як комерційні видавці, так і видавці відкритих джерел. Процес сканування набуває динамічного розвитку. На сьогоднішній день цей процес розповсюджено на спеціалізовані бібліографічні бази даних, такі як Open WorldCat OCLC і Національну бібліотеку медицини PubMed.

2. Агентна платформа — JADE

JADE написана на мові програмування Java з використанням таких розвинутих можливостей, як Java RMI, Java CORBA IDL, Java Serialization і Java Reflection API.

Обрана агентна платформа спрощує розробку мультиагентних систем завдяки використанню FIPA-специфікацій і за допомогою низки інструментів (tools), які підтримують фази виправлення помилок (debugging) і розгортання (deployment) системи. Ця агентна платформа може поширюватися серед комп'ютерів з різними операційними системами, і її можна конфігурувати через віддалений GUI-інтерфейс. Процес конфігурування цієї платформи досить гнучкий: її можна змінити навіть під час виконання програм, для цього необхідно просто перемістити агентів з однієї машини на іншу. Єдиною вимогою цієї системи є установка на машині Java Run Time 1.2. Комунікаційна архітектура пропонує гнучкий та ефективний процес обміну повідомленнями, де JADE створює чергу і управляє потоком ACL-повідомлень, які є приватними для кожного агента. Агенти здатні звертатися до черги за допомогою комбінації декількох режимів своєї роботи: блокування, голосування, перерви в роботі і зіставлення з еталоном (що стосується методів пошуку). На даний момент в системі використовується Java RMI, event-notification, і POP, але легко можна додати й інші протоколи. Також передбачена можливість інтеграції SMTP, HTTP і WAP. Більшість комунікаційних протоколів, які вже визначені міжнародним співтовариством розробників агентних середовищ, доступні і можуть ілюструватися на конкретних прикладах після визначення поведінки системи та її основних станів. SL і онтологія управління агентами також імплементовані разом з підтримкою визначених користувачем контентних мов, а також онтології, які можуть бути імплементовані і зареєстровані агентами, та використані системою. З метою істотного розширення працездатності JADE передбачена можливість інтеграції з JESS і Java-оболонкою ECLIPS.[7]

3. Технологія веб-сервіс

Веб-служба, веб-сервіс (англ. web service) — набір послуг у вигляді прикладних програм, ідентифікованих URI, загальнодоступні інтерфейси та

зв'язування (binding) яких визначаються засобами XML. Веб-служби забезпечують пряму взаємодію через Інтернет з іншими агентами програмного забезпечення, використовуючи повідомлення, основані на форматі XML. Опис цієї програмної системи може бути знайдено іншими програмними системами, які можуть взаємодіяти з нею відповідно до цього опису за допомогою повідомлень, основаних на XML, і передаватися за допомогою Інтернет-протоколів. Веб-служба є одиницею модульності при використанні сервісно-орієнтованої архітектури програми.

Основу технології веб-сервісу складають:

- найпростіші комунікаційні Інтернет-протоколи HTTP та SMTP;
- протоколи SOAP для керування повідомленнями в універсальному XML-форматі;
- мова WSDL описання інтерфейсу взаємодії компонентів розподіленої системи[8].

Базова архітектура включає наступні технології Web-служб:

- обмін повідомленнями,
- забезпечення віддаленого виклику процедур,
- опис послуг веб-служб,
- реєстрація (публікація) та пошук (виявлення) описів послуг.

Базова архітектура веб-служб визначає взаємодію між агентами програмного забезпечення як обмін повідомленнями між запитувачами сервісу (requesters) і постачальниками послуг (providers — провайдерами).

Requesters — агенти програмного забезпечення, які запитують виконання сервісу. Провайдери — агенти програмного забезпечення, які забезпечують виконання сервісу. Кожен з агентів може бути одночасно як запитувачем, так і провайдером. Providers відповідальні за реалізацію сервісу і публікацію його опису, requesters — повинні мати спосіб пошуку опису послуг. Таким чином, базова архітектура веб-служб проявляється у виконанні трьох ролей: провайдера сервісу, агентства реєстрації та пошуку сервісу і клієнта, який запитує сервіс. Взаємодія включає операції декларації (публікації), пошуку та зв'язування (bind) сервісу. Ці ролі і операції є для веб-служб артефактами (artifacts): модулями програмного забезпечення та їх описами. У типовому сценарії веб-служб провайдер сервісу містить програмні модулі з мережевим доступом, що реалізують сервіс. Також провайдер сервісу формує опис сервісу та його декларацію (публікацію), що доступна клієнту-споживачеві або агентству реєстрації (service discovery agency). Клієнт-споживач виконує операції пошуку для отримання опису сервісу локально або через реєстраційне агентство і, використовуючи цей опис, виконує процедуру зв'язування з провайдером сервісу, ініціює взаємодію з модулями, які надають цей сервіс. [8].

Як зазначалося вище, основною одиницею обміну в технології веб-служб є повідомлення, що мають структуру XML-документа. Описи сервісу використовують також XML-нотацію, що містить всі необхідні деталі для взаємодії в мережі, включаючи опис форматів повідомлення, транспортних протоколів та його локалізацію. Інакше кажучи, формування і транспортування XML-повідомлень підкоряються правилам і специфікаціям SOAP-протоколу; провайдер публікує WSDL-файл, що містить опис семантики та синтаксису повідомлення, кінцеву адресу сервісу (endpoint), що дає можливість клієнту-споживачеві правильно згенерувати SOAP-повідомлення і відправити його у правильному напрямку. Сервіс-провайдер отримує це повідомлення, обробляє запит і відсилає результат клієнтові також у вигляді SOAP-повідомлення. Технологія, типова для такого типу взаємодії веб-служб, включає SOAP, WSDL і HTTP [2].

4. Розробка моделі системи

Розроблена модель основана на раніше розглянутих технологіях і методах пошуку та доставки інформації. Дана модель відображає взаємодію програмних агентів для пошуку та класифікації інформації.

На рис. 1 наведена загальна схема розробленої моделі.

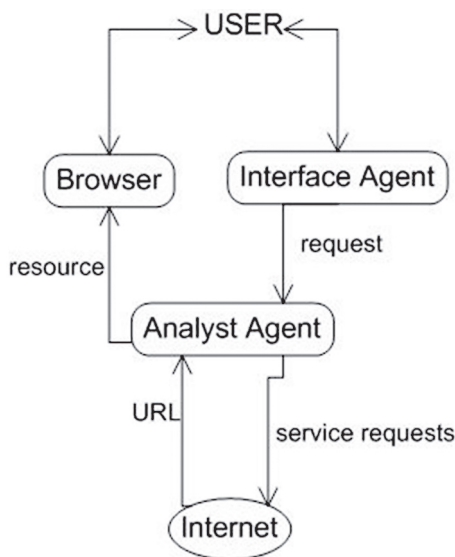


Рис. 1. Загальний вигляд розробленої моделі

В ролі «USER» виступає звичайний користувач, який безпосередньо взаємодіє через графічний інтерфейс з програмним агентом «Interface Agent». Користувач посилає запит на потрібну йому інформацію, запит обробляється даним агентом і передається наступному агенту «Analyst Agent», який в свою чергу аналізує запит, поданий вже попереднім агентом, та посилає свій запит до відповідного сервісу розміщеному в Internet. У відповідь «Analyst Agent» отримує дані, які потім аналізує та виводить у браузер користувача.

Розроблена модель припускає наступну класифікацію вихідних запитів від користувача (рис. 2):

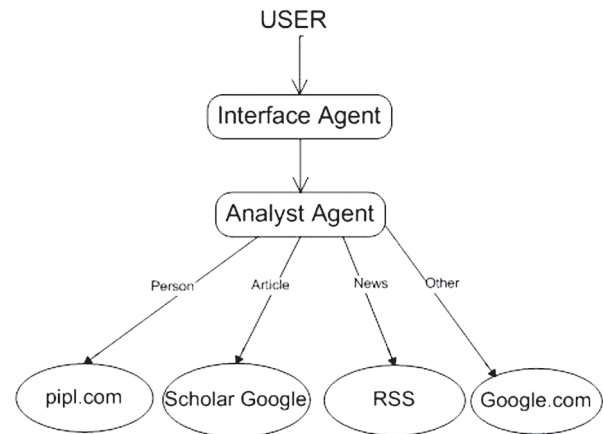


Рис. 2. Класифікація запитів користувача

– «Person» — персона. Запит направлений на пошук інформації про конкретну особу (автор статті, науковий діяч). Пошук здійснюється шляхом подання запиту на сервіс «pipl.com»;

– «Article» — стаття. Запит на пошук наукових видань (статей, конференцій, тезисів і т.і.). Пошук інформації здійснюється за допомогою сервісу Google Scholar;

– «News» — новини. Пошук актуальної інформації за ключовими словами запиту. Здійснюється шляхом сканування сервісів, які надають RSS-канали.

– «Other» — вся інша інформація. Запит на будь-яку інформацію, яка не підлягає класифікації, описаній вище. Пошук відбувається поданням звичайного запиту пошуковому сервісу «google.com».

Модель персоналізації класифікує всі ресурси, які проходять через неї, на 3 типи (рис. 3):

1) «Document» — документи. Вид інформації в вигляді файлів (.doc, .pdf, .rtf, ppt, .rar, .zip та ін.).

2) «Multimedia» — мультимедіа. Тип мультимедійної інформації (відео, аудіо).

3) «State» — сторінки. Звичайні посилання на веб-сторінки (статті, блоги, енциклопедії, форуми та ін.).

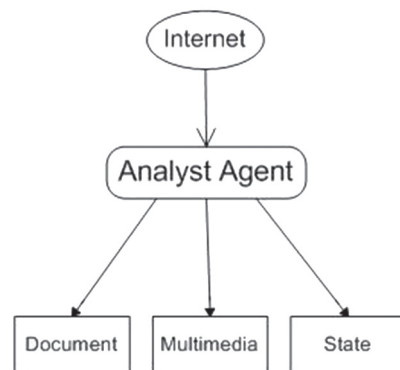


Рис. 3. Класифікація ресурсів

На рис. 4 зображена одна з основних частин розробленої моделі — модель агента «Inreface Agent». Вхідною інформацією до даної моделі є запит від користувача. Запит користувача потрапляє до модуля «Determination of the type request» — модуль обробки запиту, де визначається його тип і формується новий запит(вихідна інформація), який далі передається до агента «Analyst Agent».

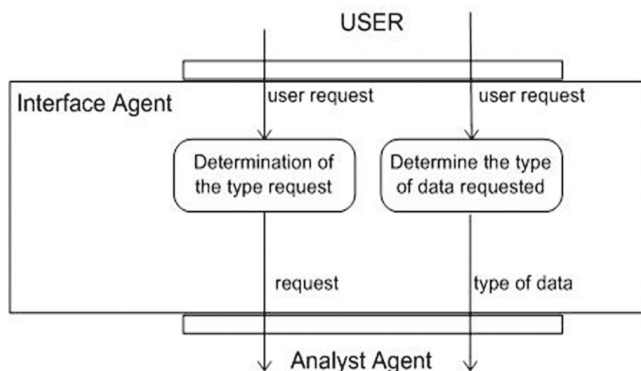


Рис. 4. Модель Інтерфейс-агента

Також в запиті користувача може міститися додаткова інформація щодо типу запитуваного ресурсу. Якщо дана інформація присутня, вона потрапляє в модуль «Determine the type of data request» — модуль обробки запиту користувача. «Inreface Agent» визначає наявність в запиті вказаного типу даних, привласнює йому відповідний тип та формує новий запит (вихідна інформація), який передає до агента «Analyst Agent». Якщо запит не містить явно вказаного типу, то модуль нічого не приймає і не передає.

Остання, основна, частина розробленої моделі — це модель агента «Analyst Agent». На рис. 5 наведена узагальнена схема моделі даного агента. Вхідною інформацією верхнього рівня є запит, переданий від агента «Inreface Agent». Якщо переданий запит містить тільки тип, привласнений попереднім агентом, то запит направляється до модуля «Select Service», де відбувається аналіз запиту, вибір відповідного джерела інформації(веб-сервісу) та формування запиту до конкретного сервісу. Якщо ж запит містить додаткову інформацію про тип запрошуваного ресурсу, то крім модуля «Select Service» запит направляється до модуля «Processing of data», де чекає до моменту приходу відповіді з даними від сервісу.

Після успішного виконання модуля «Select Service» на виході маємо запит, який направляється до конкретного сервісу. Відповідь від сервісу є вхідною інформацією нижнього рівня до моделі агента. Дана інформація потрапляє для подальшої обробки до модуля «Processing of data» — модуль обробки відповіді від сервісу та виводу інформації на браузер користувача.

В модулі «Processing of data» вхідні дані піддаються перетворенню, а саме фільтрації для конкретних типів даних, якщо ті були вказані в вхідному запиті від агента «Inreface Agent». Далі модуль виводить вихідну інформацію в браузер користувача.

Якщо у початковому запиті користувача не було вибрано додаткових параметрів, пов'язаних з типом ресурсів, «user» може переформулювати запит уже після виводу інформації в браузер шляхом вибору типу даних в графічному інтерфейсі. Таким чином відбудеться запит до модуля «Processing of data» агента «Inreface Agent», який переформулює запит та відсіює непотрібну інформацію.

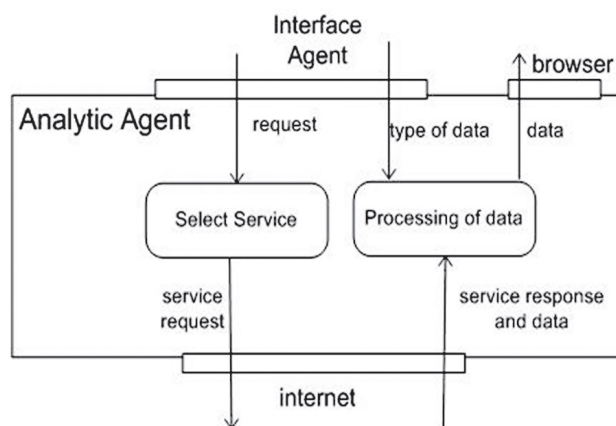


Рис. 5. Модель Агента-аналітика

Висновки

В даній статті була описана та проаналізована розроблена сервіс-орієнтована модель персоналізованого навчання з використанням агентних технологій, основною функцією якої є обробка та доставка потрібної користувачеві інформації.

В рамках створення веб-сервісу персоналізованого навчання було досліджено існуючі технології та засоби пошуку і доставки інформації.

Виявлено ефективність використання агентної платформи JADE. Розроблена модель включає моделі програмних агентів, які реалізують пошук та доставку інформації кінцевому користувачеві, та має засоби, необхідні для персоналізації інформації.

Список літератури: 1. Зайцева, Л. В. Модели и методы адаптации в системах компьютерного обучения [Текст] / Л. В. Зайцева. — СПб.: С-ПИТМО, 2003. — 1200 с. 2. Особенности разработки мультиагентных систем на основе платформы jade в рамках парадигмы semantic web [Текст] : материалы междунар. науч.-практ. конф., 8-10 июня 2009 г. Минск / редкол. : Е. Н. Владимирская (отв. ред.). — Минск : Темп, 2009. — 112 с. 3. Контроль и управление в сложных системах [Текст] : материалы IX междунар. науч.-практ. конф. 2 апреля 2009 г. Винница / редкол. : В. В. Грабко (отв. ред.). — Винница : Темп, 2009. — 58 с. 4. Что такое RSS? [Электронный ресурс] — Режим доступа: <http://www.webmascon.com/topics/technologies/9a.asp>—08.06.2005 г. — Загл. с экрана. 5. Pipl.com ищет людей

в скрытом Интернете [Электронный ресурс] – Режим доступа: <http://itua.info/news/internet/9274.html> – 01.01.2007 г. – Загл. с экрана. 6. О поисковой системе Google Scholar [Электронный ресурс] – Режим доступа: <http://ousnano.sbras.ru/sitepage.php?PageID=447> – 05.01.2010 г. – Загл. с экрана. 7. Образование, наука, производство: использование JADE для разработки компьютерных систем поддержки дистанционного обучения агентного типа [Текст] : тез. докл. науч.-практ. конф. (авг. 2005) / отв. ред. Н. Н. Глибовец. – Киев : ИСН, 2005. – 432 с. 8. *Галактионов, В. В.* Web – services – сервис-ориентированная технология для распределенных объектных вычислительных систем. Основные концепции, протоколы и спецификации [Текст] / В. В. Галактионов В. В. – М.: ОИЯИ, 2003. – 100 с.

Надійшла до редколегії 13.09.2010.

УДК 519.62

Сервис-ориентированная модель персонализированного обучения с использованием агентных технологий / Н.С. Лесная, И.Н. Келеберда, С.Н. Шрестха, Ю.А. Кирпусь

// Бионика интеллекта: науч.-техн. журнал. – 2010. – № 3 (74). – С. 133-137.

В статье рассматривается сервисно-агентный подход для разработки модели персонализированного обучения с основной функцией обработки и доставки необходимой информации пользователю. Приводится описание разработанной модели на основе агентной платформы JADE.

Ил. 5. Библиогр.: 7 назв.

UDK 519.7

Service-oriented model of personalized learning using agent technologies / N.S. Lesnaya, I.N. Keleberda, S.N. Shrestkha, Y.A. Kirpus' // Bionics of Intelligence: Sci. Mag. – 2010. – № 3 (74). – P. 133–137.

The article deals with a service-agent-based approach to develop a model of personalized training with the main function of processing and delivery of necessary information to the user. The description of the model-based agent platform JADE.

Fig. 5. Ref.: 7 items.

УДК



ІНФОРМАЦІЙНА ТЕХНОЛОГІЯ ПІДТРИМКИ КОНСУЛЬТУВАННЯ СТУДЕНТІВ ДИСТАНЦІЙНОЇ ФОРМИ НАВЧАННЯ

Н.С. Лєсна¹, А.В. Найденко², В.В. Сокол³

¹ ХНУРЕ, м. Харків, Україна, lmd@kture.kharkov.ua;

² ХНУРЕ, м. Харків, Україна, anaydenko@hotmail.com

³ ХНУРЕ, м. Харків, Україна, vitalysokol@kture.kharkov.ua

У рамках роботи було досліджено методи покращення консультування студентів дистанційної форми навчання, розглянуто етапи та моделі пошуку інформації, розглянуто та порівняно методи імовірнісного латентно-семантичного аналізу, добування ключових понять із тексту, класифікації та кластеризації текстових документів, автоматичного анотування текстів. В результаті було спроектовано та розроблено інформаційну технологію підтримки консультування студентів дистанційної форми навчання у вигляді веб-застосування з використанням технології ASP.NET. Інформаційна технологія оптимізує процес консультування студентів, зменшуючи об'єм роботи викладача та підвищуючи швидкість отримання студентом відповіді на поставлене запитання, що реалізовано за допомогою імовірнісного латентно-семантичного аналізу, який виявляє приховані фактори та дає змогу добути контекстно-залежні значення слів за допомогою статистичної обробки великих наборів текстових даних.

ДИСТАНЦІЙНЕ НАВЧАННЯ, ІНТЕЛЕКТУАЛЬНИЙ АНАЛІЗ ДАНИХ, ІМОВІРНІСНИЙ ЛАТЕНТНО-СЕМАНТИЧНИЙ АНАЛІЗ, TEXT MINING, TFIDF, NLP

Вступ

Сьогодні дистанційне навчання через Інтернет усе частіше розглядається не просто як зручна форма підвищення кваліфікації, а як альтернатива традиційній освіті, що дозволяє студентові одержати глибокі знання.

Термін «дистанційне навчання» означає таку організацію навчального процесу, при якій викладач розробляє навчальну програму, головним чином основу на самостійному навчанні студента. Основною відмінною рисою дистанційного навчання від інших типів освіти є широке використання телекомунікаційних технологій у навчальному процесі.

Особливу роль у дистанційному навчанні відіграє інтерактивна взаємодія викладача та студентів. У вузькому розумінні слова (стосовно роботи користувача програмного забезпечення взагалі) інтерактивна взаємодія – це діалог користувача з програмою, тобто обмін командами, запитаннями й відповідями. *При дистанційному навчанні засобом здійснення інтерактивної взаємодії є розроблена інформаційна технологія підтримки консультування студентів дистанційної форми навчання. Ця технологія виконує інтелектуальний пошук відповіді перед відправкою питання викладачеві, що дозволяє зменшити об'єм роботи викладача, заощадити час студента завдяки автоматичному швидкому отриманню відповіді на поставлене запитання.*

Постановка задачі. В рамках роботи було поставлено наступні задачі:

- ознайомитися з дистанційною формою навчання з використанням дистанційних освітніх технологій, форм, методів, засобів навчання й освітніх масивів мережі Інтернет;

- дослідити способи автоматизації процесу консультування студентів дистанційної форми навчання;

- дослідити методи інтелектуального пошуку у глобальних і локальних обчислювальних мережах і базах даних;

- ознайомитися із засобами інтелектуального пошуку текстової інформації;

- розглянути та вивчити окремі етапи використання Text Mining – глибинного аналізу текстової інформації;

- порівняти методи інтелектуального аналізу тексту для вирішення поставлених задач;

- спроектувати інформаційну технологію підтримки консультування студентів дистанційної форми навчання з використанням інтелектуального аналізу тексту;

- розробити програмне рішення інформаційної технології підтримки консультування студентів дистанційної форми навчання з використанням інтелектуального пошуку інформації.

Актуальність. Робота є досить актуальною, оскільки за допомогою інформаційної технології підтримки консультування студентів дистанційної форми навчання:

- множить педагогічні можливості викладачів навчального закладу;

- робить процес навчання більш предметним;

- створює додаткову мотивацію у студентів до вивчення матеріалів;

- заощаджує час викладача і студента за рахунок автоматичного інтелектуального пошуку відповіді на запитання студента до того, як воно буде залишене на розгляд викладача.

1. Автоматизація процесу консультування студентів дистанційної форми навчання

Інтерактивна взаємодія викладача та студентів відіграє особливу роль у дистанційному навчанні. Чим більше існує можливостей управляти про-

грамою, чим активніше користувач бере участь у діалозі, тим вище інтерактивність. У широкому розумінні інтерактивна взаємодія припускає діалог будь-яких суб'єктів один з одним з використанням доступних їм засобів і методів. При цьому передбачається активна участь у діалозі обох сторін – обмін запитаннями та відповідями, керування ходом діалогу, контроль над виконанням ухвалених рішень тощо. При дистанційному навчанні суб'єктами в інтерактивній взаємодії будуть виступати викладачі та студенти, а засобом здійснення подібної взаємодії – розроблена інформаційна технологія підтримки консультування студентів дистанційної форми навчання. Ця технологія виконує інтелектуальний пошук відповіді перед відправленням запитання викладачеві, що дозволяє зменшити об'єм роботи викладача, заощадити час студента завдяки автоматичному швидкому отриманню відповіді на поставлене запитання.

Існує широкий спектр пропонуваніх рішень і перспективних напрямків досліджень в області інформаційного пошуку, починаючи від побудови глобальних розподілених інформаційних структур і пошукових систем, закінчуючи елементарними на перший погляд питаннями аналізу документів. Усі вони, безумовно, важливі та корисні при розв'язку своїх специфічних задач. Проте, саме від методів аналізу багато в чому залежить ефективність існуючих пошукових систем, тому що вони є основою будь-якої пошукової системи та багато в чому визначають можливості й обмеження цих систем.

Крім цього існує ще один важливий фактор, що визначає ефективність будь-якого інформаційного пошуку – це людський фактор. Найчастіше у більшості досліджень в області інформаційного пошуку цей фактор або ігнорується, або його значення багато в чому недооцінюється. Але саме люди в остаточному підсумку користуються розробленими інформаційно-пошуковими системами. Облік людського фактора, специфіки його роботи, переваг і очікувань є перспективним і багатообіцяючим напрямком досліджень.

Представлені на сьогоднішній день у більшості популярних пошукових систем способи організації повнотекстового пошуку та методи аналізу документів не враховують у достатній мірі саме людський фактор. Тобто, не враховується той факт, що багато в чому пошук визначається слабко формалізованими та нечіткими умовами, у значній мірі залежними від досвіду та переваг самої людини. Далеко не завжди користувач інформаційно-пошукової системи може чітко й однозначно сформулювати саме той набір ключових слів, який і приведе його до бажаного результату. Мова йде про варіант пошуку на основі формування інформаційних запитів, що складаються із набору ключових слів. Цей варіант пошуку найпоширеніший і методологічно пророблений на сьогоднішній день.

2. Методи інтелектуального пошуку інформації

В рамках даної роботи було досліджено наступні методи інтелектуального аналізу текстових даних з метою пошуку релевантної інформації:

а) метод добування ключових понять із тексту. Первинною метою добування ключових понять (feature extraction) є ідентифікація фактів і відносин у тексті. У більшості випадків такими поняттями є іменники власні та загальні: імена й прізвища людей, назви організацій тощо. Алгоритми добування понять можуть використовувати словники, щоб ідентифікувати деякі терміни та лінгвістичні шаблони для визначення інших;

б) метод класифікації текстових документів. Класифікація (classification) – стандартна задача з області Data Mining. Її метою є визначення для кожного документа однієї або декількох заздалегідь заданих категорій, до яких цей документ відноситься. Особливістю задачі класифікації є припущення, що множина документів, що класифікуються, не містить “сміття”, тобто кожний з документів відповідає якій-небудь заданій категорії. Окремим випадком задачі класифікації є задача визначення тематики документа;

в) метод кластеризації текстових документів. Метою кластеризації (clustering) документів є автоматичне виявлення груп схожих документів серед заданої фіксованої множини. Відзначимо, що групи формуються тільки на основі попарної схожості описів документів, і ніякі характеристики цих груп не задаються заздалегідь;

г) метод автоматичного анотування текстів. Автоматичне анотування (summarization) дозволяє скоротити текст, зберігаючи його зміст. Розв'язок цієї задачі звичайно регулюється користувачем за допомогою визначення кількості добутих речень, або відсотком тексту, що добувається, стосовно усього тексту. Результат містить у собі найбільш значимі речення в тексті;

д) метод імовірнісного латентно-семантичного аналізу. Імовірнісний латентно-семантичний аналіз базується на статистичному аналізі даних, дозволяє зробити мовнезалежний семантичний пошук інформації з можливістю виявлення схованих факторів (тем), присутніх в інформаційному масиві, з урахуванням семантичної близькості слів.

Найбільш підходящим методом, який би дозволив максимально ефективно вирішити нашу задачу інтелектуального пошуку текстової інформації, є метод імовірнісного латентно-семантичного аналізу.

3. Семантичний інформаційний пошук

Окремою задачею інформаційного пошуку є латентно-семантичний аналіз тексту. Латентно-семантичний аналіз (LSA – Latent Semantic Analysis)

– це теорія й метод для добування контекстно-залежних значень слів за допомогою статистичної обробки великих наборів текстових даних [1]. Даний метод аналізу використовується не тільки в області пошуку інформації [2], але й у задачах фільтрації й класифікації [3].

Основна ідея латентно-семантичного аналізу полягає в тому, що сукупність усіх контекстів, у яких зустрічається й не зустрічається дане слово, задає множину взаємних обмежень, які дозволяють визначити подібність змістовних слів між собою.

Вихідною інформацією для LSA є матриця термів на документи, яка описує використовуваний для навчання системи набір даних. Елементи цієї матриці містять частоти використання кожного терму в кожному документі.

Один з найпоширеніших варіантів LSA заснований на використанні розкладання вихідної матриці за сингулярними значенням (SVD – Singular-Value Decomposition). Використовуючи SVD, велика вихідна матриця розкладається в множини із k ортогональних матриць, лінійна комбінація яких є достатнім наближенням вихідної матриці.

Згідно з теоремою про сингулярне розкладання, будь-яка речовинна прямокутна матриця X може бути розкладена в добуток трьох матриць:

$$X = U \Sigma V^T,$$

де матриці U і V – ортогональні, а Σ – діагональна матриця, значення на діагоналі якої називаються сингулярними значеннями матриці X .

Особливість такого розкладання в тому, що якщо в Σ залишити тільки k найбільших сингулярних значень, а в матрицях U і V тільки відповідні до цих значень стовпці, то добуток матриць, що вийшли, U_{lsa} , Σ_{lsa} і V_{lsa} буде найкращим наближенням вихідної матриці X матрицею рангу k .

$$M = \hat{X} = U_{lsa} \Sigma_{lsa} V_{lsa}.$$

Ідея такого розкладання та суть латентно-семантичного аналізу полягає в тому, що якщо в якості X використовувалася матриця термів на документ, то матриця \hat{X} , що містить тільки k перших лінійно незалежних компонентів X , відображає основну структуру асоціативних залежностей, присутніх у вихідній матриці, і в той же час не містить шуму.

Таким чином, кожний терм і документ представляються за допомогою векторів у загальному просторі розмірності k (так званім просторі гіпотез). Близькість між будь-якою комбінацією термів або документів може бути легко обчислена за допомогою скалярного добутку векторів.

Найбільший інтерес у рамках даної роботи викликає метод імовірнісного латентно-семантичного аналізу тексту. Розглянемо його докладніше.

Метод імовірнісного латентно-семантичного аналізу (PLSA – Probabilistic Latent Semantic

Analysis) ставить своєю задачею виявлення латентних, схованих, факторів (тем), присутніх у колекції й пов'язаних з її документами й словами. Саме, фіксувавши число схованих факторів r , за допомогою методу PLSA можна оцінити наступні величини:

– $P(z_i)$ – імовірність того, що випадково обраний з колекції документ найбільш тісно пов'язаний з фактором (найбільшою мірою відповідає темі) z_i ;

– $P(d_j | z_i)$ – імовірність того, що найбільш тісно пов'язаний з даним фактором z_i документ – це d_j ;

– $P(w_j | z_i)$ – імовірність того, що для даного фактора z_i найбільш тісно пов'язане з ним слово – це w_j .

Тут $D = \{d_1, \dots, d_n\}$ – множина усіх проіндексованих документів, $W = \{w_1, \dots, w_n\}$ – множина усіх різних слів, що зустрілися в проіндексованих документах, $Z = \{z_1, \dots, z_r\}$ – множина латентних факторів.

Позначимо за $P(d, w)$ ймовірність спільного спостереження документа d й слова w , тобто $P(d, w) = P(d)P(w|d)$. У рамках підходу, використовуюваного при імовірнісному латентно-семантичному аналізі, величина $P(d, w)$ оцінюється за наступною формулою:

$$P(d, w) = \sum_{z \in Z} P(z)P(d|z)P(w|z).$$

Відповідно до принципу максимальної правдоподібності функції $P(z)$, $P(d|z)$ й $P(w|z)$ визначаються шляхом максимізації функції правдоподібності $L = \sum_{d \in D} \sum_{w \in W} n(d, w) \log P(d, w)$, де $n(d, w)$ є число входжень слова w у документ d .

На етапі максимізації функції L використовується метод оцінювання-максимізації. На кожній ітерації виконується крок оцінювання

$$P(z|d, w) = \frac{P(z)P(d|z)P(w|z)}{\sum_{z'} P(z')P(d|z')P(w|z')}$$

і потім крок максимізації

$$P(w|z) = \frac{\sum_d n(d, w)P(z|d, w)}{\sum_{d, w'} n(d, w')P(z|d, w')},$$

$$P(d|z) = \frac{\sum_w n(d, w)P(z|d, w)}{\sum_{d', w} n(d', w)P(z|d', w)},$$

$$P(z) = \frac{1}{R} \sum_{d, w} n(d, w)P(z|d, w),$$

$$R = \sum_{d, w} n(d, w).$$

Будемодалізаувати вектор $(P(d|z_1), \dots, P(d|z_r))$ образом документа d в просторі факторів Z і, аналогічно, вектор $(P(w|z_1), \dots, P(w|z_r))$ – образом слова w . Позначимо через D' деяку множину документів, випадково відібраних з множини D

всіх документів, у якій виконується пошук, а через W' – множину усіх слів з документів, що ввійшли до D' . Нехай також $d \in D - D'$.

Далі припускаємо, що множина D' може розглядатися у якості представницької вибірки документів з колекції D , і в ній порушені всі теми, відбиті в повній колекції. Іншими словами, довільний документ d містить значну кількість слів з W' , при цьому підмножина W' достатньо повно відображає тематичну спрямованість d .

Розглянемо систему лінійних алгебраїчних рівнянь

$$P(d, w) = \sum_{z \in Z} P(z)P(w|z)P(d|z),$$

де $w \in W'$, $n(d, w) > 0$.

У якості невідомих розглядаємо величини $P(d|z)$, $z \in Z$. Значення величин $P(z)$, $P(w|z)$, $z \in Z$, $w \in W'$ отримані в результаті застосування PLSA до множини документів D' . Величина $P(d|w)$ апроксимується по формулі $P(d, w) \approx \frac{1}{|D'|} \frac{n(d, w)}{\text{length}(d)}$, де $\text{length}(d)$ – кількість слів із W' , присутніх у документі d .

Для наближеного розв'язку даної системи ми будемо псевдообернену матрицю для матриці коефіцієнтів системи методом Гревілья. Отримане наближення є найкращим (за методом найменших квадратів) і може розглядатися як досить точна апроксимація образу документа d у просторі факторів Z .

Критерієм оцінки якості пошуку інформації є побудова усередненої за запитами кривої повноточності. Однак цей підхід неприйнятний у зв'язку із широким розкидом числа релевантних документів для різних запитів. У зв'язку із цим для порівняння результатів пошуку за вихідним й розширеним запитами використовуємо наступний підхід. Позначимо через R_q множину рангів (порядкових номерів) релевантних документів для запиту q серед перших 200 документів, отриманих у результаті пошуку. Нехай далі Q є множина запитів, а $rel(q)$ – загальне число документів в D , релевантних запиту q . Нехай

$$quality = \frac{1}{|Q|} \sum_{q \in Q} \frac{1}{(1 + rel(q))} \sum_{i \in R_q} \frac{1}{i}.$$

Саме величина $quality$ буде використовувати для оцінки ефективності пропонованого методу розширення запиту, обчислюючи її для вихідних і розширених запитів. Ця інтегральна оцінка враховує як ранги всієї сукупності релевантних документів, отриманих у відповідь на запит, так і загальне число релевантних документів у колекції. Ранги отриманих релевантних документів для запитів з невеликим загальним числом релевантних документів урахуються з більшими вагами, що відповідає інтересам користувача, що бажає бачи-

ти релевантні документи серед перших повернутих системою результатів, особливо у випадку їх мало-го числа [4].

4. Розробка інформаційної технології підтримки консультування студентів

Усю сукупність представлених на сьогоднішній день методів аналізу тексту, щодо задачі аналізу його змісту, можна розділити на дві великі групи:

а) лінгвістичний аналіз, орієнтований на добування змісту тексту за його семантичною структурою;

б) статистичний аналіз, орієнтований на добування змісту тексту за частотним розподілом слів.

Авторами було ухвалене рішення використовувати методи статистичного аналізу в силу їх зручності використання і мовної незалежності. Методи лінгвістичного аналізу хоча й дозволяють точніше аналізувати текст, виділяючи його структурні особливості, але є більш складними у використанні. Пов'язане це, насамперед, з багатством семантики й морфології природних мов. Формальний опис правил природної мови і їх реалізація – досить трудомісткий процес, що вимагає залучення фахівців з області лінгвістики. Крім того, лінгвістичний аналіз припускає орієнтацію на конкретну мову з її конкретними семантичними особливостями, це обумовлює погану міжмовну переносимість. Роботи в даному напрямку йдуть, і існують множини практичних реалізацій, але на сьогоднішній день лінгвістичний аналіз у галузі аналізу семантики досить проблематичний.

Усе це обумовило доцільність застосування статистичних методів для розв'язку задач, що були поставлені в даній статті, тому в рамках даної роботи тематичний аналіз тексту припускає розв'язання задачі семантичного пошуку інформації з використанням статистичного аналізу в силу його мовної незалежності.

5. Попередня обробка тексту

Однією з головних проблем аналізу текстів є велика кількість слів у документі. Якщо кожне із цих слів піддавати аналізу, то час пошуку нових знань різко зростає та навряд чи буде задовольняти вимогам користувачів. У той же час очевидно, що не всі слова в тексті несуть корисну інформацію. Крім того, у силу гнучкості природних мов формально різні слова насправді означають однакові поняття. Таким чином, видалення неінформативних слів, а також приведення близьких за змістом слів до єдиної форми значно скорочують час аналізу текстів. Усунення описаних проблем виконується на етапі попередньої обробки тексту.

Звичайно використовують наступні способи видалення неінформативних слів і підвищення строгості текстів:

– Видалення стоп-слів. Стоп-словами називаються слова, які є допоміжними та несуть мало інформації про зміст документа. Звичайно заздалегідь складаються списки таких слів, і в процесі попередньої обробки вони видаляються з тексту. Типовим прикладом таких слів є допоміжні слова й артиклі, наприклад: “тому що”, “крім того” тощо.

– Стеммінг – морфологічний пошук. Він полягає в приведенні кожного слова до його нормальної форми, виділення основи слова. Нормальна форма виключає відмінювання слова, множинної форми, особливості усної мови тощо. Наприклад, слова “стиск” і “стислий” повинні бути перетворені в нормальну форму слова “стискати”. Алгоритми морфологічного розбору враховують мовні особливості та внаслідок цього є мовно залежними алгоритмами.

– Використання TV-грам. TV-грами – це альтернатива морфологічному розбору та видаленню стоп-слів. TV-грама – це частина рядка, що полягає з TV символів. Наприклад, слово “дата” може бути презентовано 3-грамою “да”, “дат”, “ата”, “та” або 4-грамою “дат”, “дата”, “ата_”, де символ підкреслення заміняє попередній або замикаючий слово пробіл. У порівнянні зі стеммінгом або видаленням стоп-слів, TV-грами менш чутливі до граматичних і типографічних помилок. Крім того, діаграми не вимагають лінгвістичного представлення слів, що робить даний спосіб більш незалежним від мови. Однак TV-грами, дозволяючи зробити текст більш строгим, не вирішують проблему зменшення кількості неінформативних слів.

– Приведення регістру. Цей спосіб полягає в перетворенні всіх символів до верхнього або нижнього регістру. Наприклад, усі слова “текст”, “Текст”, “ТЕКСТ” приводяться до нижнього регістру “текст”.

6. Архітектура системи

Система передбачає використання архітектури клієнт-сервер – обчислювальної або мережної архітектури, у якій завдання або мережне навантаження розподілене між постачальниками послуг (сервісів), що називаються серверами, і замовниками послуг, що називаються клієнтами. Нерідко клієнти та сервери взаємодіють через комп’ютерну мережу та можуть бути як різними фізичними обладнаннями, так і програмним забезпеченням.

У якості архітектури пошукової системи була обрана трьохрівнева архітектура клієнт-сервер, яка передбачає наявність клієнтського застосування (звичайно говорять «тонкий клієнт» або термінал), підключене до сервера застосувань, який у свою чергу підключений до сервера бази даних. Така архітектура забезпечує масштабованість, конфігурованість (завдяки ізольованості рівнів один від одного), високу безпеку та надійність, низькі вимоги

до продуктивності та до технічних характеристик терміналів.

На рис.1 представлена архітектура пошукової системи у вигляді діаграми взаємодії компонентів системи.

Опис представлених компонентів на діаграмі:

– SearchManager – двигок пошукової системи, що забезпечує взаємодію всіх компонентів системи;

– ProbabilisticLatentSemanticAnalyzer – компонент, який виконує імовірнісний латентно-семантичний аналіз тексту;

– StopWordsRemover – компонент, який відповідає за видалення словникових стоп-слів;

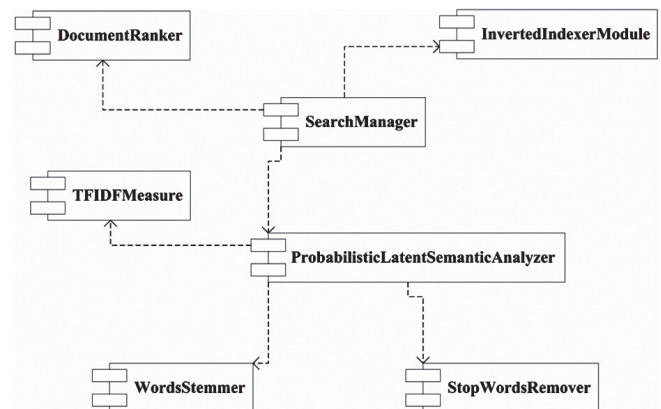


Рис. 1. Діаграма взаємодії компонентів системи

– WordsStemmer – компонент, який виконує стеммінг слів (знаходження основи слова), і використовується для знаходження найбільш інформативних слів;

– TFIDFMeasure – компонент, який відповідає за обчислення ступеня важливості слова в контексті документів по алгоритму TF-IDF;

– DocumentRanker – компонент, який виконує ранжирування пошукових образів документів, відповідно до кількості та важливості слів, що в ньому зустрічаються;

– InvertedIndexerModule – модуль, який виконує відкладене зворотне індексування документів, забезпечуючи тим самим формування бази існуючих відповідей.

Етапи інтелектуального аналізу текстової інформації з використанням імовірнісного латентно-семантичного аналізу представлені на рис. 2 у вигляді діаграми послідовності, яка зображує процес взаємодії 5 основних об’єктів пошукової системи:

- користувача системи (User);
- двигка пошукової системи (Search Manager);
- попереднього оброблювача слів (Words Processor);
- імовірнісного латентно-семантичного аналізатора тексту (Probabilistic Latent Semantic Analyzer);
- ранжиру документів (Document Ranker).

Інтелектуальний пошук текстової інформації можна представити у вигляді декількох етапів: пошук документів, їх попередня обробка, добування інформації, застосування методу імовірнісного латентно-семантичного аналізу, інтерпретація результатів.

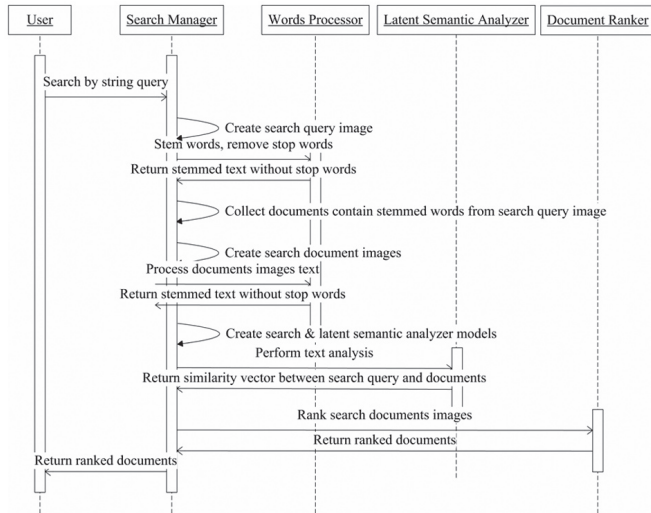


Рис. 2. Діаграма послідовності

Висновки

Дистанційне навчання ще недостатньо впроваджене в систему освіти України, хоча, судячи з останніх тенденцій, має великі шанси міцно увійти в життя як студентів, так і фахівців старших вікових категорій, надаючи їм можливість реалізувати концепцію постійного підвищення кваліфікації протягом всього життя.

Об'єктом дослідження у рамках роботи був процес консультування студентів дистанційної форми навчання. Методом дослідження виступало поліпшення процесу консультування для зменшення об'єму роботи викладача та підвищення швидкості отримання студентом відповіді на поставлене запитання.

У ході аналізу предметної області було зроблено загальний огляд системи дистанційного навчання, дистанційних навчальних технологій, процесу консультування студентів та його можливої автоматизація.

В результаті було спроектовано та розроблено інформаційну технологію підтримки консультування студентів дистанційної форми навчання у вигляді веб-застосування з використанням технології ASP.NET, яка оптимізує процес консультування студентів, зменшуючи об'єм роботи викладача та підвищуючи швидкість отримання студентом відповіді на поставлене запитання.

Отриманий програмний продукт може бути використано у закладах, що забезпечують навчання

та перепідготовку з використанням дистанційної освіти. Крім того, він може бути використаний на фірмах, що хотіли б збільшити швидкість навчання своїх нових співробітників.

Розроблена інформаційна технологія може бути розширена за наступними напрямками:

- застосування морфологічного аналізу при пошуку текстової інформації;
- інтегрування перевірки орфографічних помилок у пошуковому запиті;
- реалізація надання пошуку мультимедійної інформації, наприклад, формул у вигляді рисунків;
- розгляд можливості використання онтологій;
- розширення локалізації засобів попередньої обробки даних.

Список літератури: 1. *Deerwester, S.* Indexing by Latent Semantic Analysis [Текст] / S. Deerwester, S. T. Dumais, G. W. Furnas and other // Journal of the American Society for Information Science. – 1990. – Т. 41, №6. – PP. 391-407. 2. Latent semantic indexing (LSI) and TREC-2 [Текст]: In Proc. of the Second Text Retrieval Conference / D. Harman. – 1994. – PP. 105-116. 3. Using latent semantic indexing for information filtering [Текст] : In ACM Conference on Office Information Systems (COIS) / P. W. Foltz. – 1990. – PP. 40-47. 4. *Hofmann, T.* Probabilistic latent semantic indexing [Текст] / T. Hoffman. – In Proc. of the SIGIR'99. – 1999. – PP 50-57

Надійшла до редколегії 15.09.2010.

УДК 519.863.5

Информационная технология поддержки консультирования студентов дистанционной формы обучения / Н.С. Лесная, А.В. Найденко, В.В. Сокол // Бионика интеллекта: науч.-техн. журнал. – 2010. – № 3 (74). – С. 138–143.

Статья посвящена решению актуальной задачи – повышению качества консультирования в системах дистанционного обучения. В процессе решения были проанализированы существующие подходы и определен наилучший – метод вероятностного латентно-семантического анализа. В результате было получено программное средство, которое позволяет ускорить процесс консультирования и повысить качество поиска в базе существующих ответов.

Ил. 2. Библиогр.: 4 назв.

UDC 519.863.5

Information technology of support for counseling students of distance learning / Lesna N.S., Naydenko A.V., Sokol V.V. // Bionics of Intelligence: Sci. Mag. – 2010. – № 3 (74). – P. 138–143.

The main goal of the article is solving the actual problem - improving the quality of counseling in the systems of distance learning. A method of probabilistic latent-semantic analysis was chosen to solve the problem. As a result of the software tool which allows improving the performance of counseling and decreasing the searching time of correct answer.

Fig. 2 Ref.: 4 items

УДК 004.08



МУРАВЬИНЫЙ АЛГОРИТМ И УЧЕТ РИСКА В ТРАНСПОРТНОЙ ЗАДАЧЕ

А.Я. Кузemin, Я. Даюб

ХНУРЭ, г. Харьков, Украина, kuzy@kture.kharkov.ua

В статье произведен анализ возможности применения муравьиного алгоритма как наиболее перспективного из метаэвристических алгоритмов к задаче поиска наикратчайшего маршрута в условиях наличия риска и ограничения по риску. Показана возможность эффективного применения муравьиного алгоритма в случае предсказания риска и удовлетворительная скорость сходимости алгоритма в любых условиях.

ТРАНСПОРТНАЯ ЗАДАЧА, РИСК, МУРАВЬИНЫЙ АЛГОРИТМ

Введение

В настоящее время все большее значение приобретают разработка и внедрение методов снижения рисков перевозок при транспортировке опасных грузов (ОГ). Необходимы мероприятия, направленные на снижение времени прохождения маршрута и уменьшения различных рисков, связанных с прогрессирующим усложнением среды движения с точки зрения потенциальной опасности или, другими словами, возникновения чрезвычайных ситуаций (ЧС) на маршруте движения, интенсификацией перевозочного процесса и характеристиками надежности транспортных средств. Очевидная сложность задачи заключена уже в самом названии, т.к. подразумевает одновременную оптимизацию по двум параметрам, к тому же оба этих параметра являются функциями времени. Класс задач поиска оптимального маршрута, для которого за время прохождения любой его части могут измениться условия прохождения его участков, получил название «динамических транспортных задач» [1]. Принято условно разделять процесс планирования и выполнение прохождения маршрута [2]. Именно последний и будет представлять собой непосредственно динамическую транспортную задачу в чистом виде. До недавнего времени процессу планирования (статическая задача) уделялось более пристальное внимание, нежели процессу прохождения маршрута. Необходимо подчеркнуть, что в процессе прохождения маршрута могут возникать различные ЧС. Эти ситуации не всегда можно предсказать или прогнозировать с заданной точностью (т.е. оценить риски от их возникновения). Таким образом, в случае возникновения ЧС при прохождении транспортного средства с ОГ по планируемому маршруту (за заданное время) может потребоваться полная переработка этого маршрута. Другими словами, уже после начала движения транспортного средства с ОГ при возникновении ЧС необходимо изменение его маршрута [2].

1. Постановка задачи

Требуется найти путь из точки A в точку B произвольного графа G , в котором каждое ребро характеризуется двумя функциями: $R(t)$ — риск при

прохождении данного ребра и $T(t)$ — время, которое будет затрачено при перемещении между точкой начала S и конца D данного ребра. При этом $t = t_0 + \sum t_i$, t_i — время, затраченное на прохождение каждого из ранее пройденных ребер от A до S . Необходимо найти минимальное время t_{\min} прохождения дерева при гарантированном риске не более R_{\max} , при чем риск изменяется в диапазоне $R \in (0; 1)$.

Для решения задачи следует:

- разработать методологию оценки риска при возникновении ЧС и транспортных происшествий и экономических ущербов от них для транспорта в целом и конкретных маршрутов перевозки ОГ;
- обеспечить формирование схемы вариантных расчётов последствий возможных ЧС с их отображением на картографической основе программными средствами геоинформационных (ГИС) технологий;
- получить и систематизировать результаты оценки рисков возникновения аварийных ситуаций и ущербов от них при перевозке ОГ для транспортной сети с учётом наличия взаимодействия в местах опасных пересечений с автомобильными дорогами на переездах и системами трубопроводного транспорта;
- разработать экспертную и обучающую системы для мобильной среды управления перевозками ОГ с минимумом рисков и тренинга персонала для транспортировки ОГ в условиях ЧС.

2. Анализ существующих методов

В общем виде данная задача является задачей многокритериальной оптимизации. Для подобных задач [1] существуют «точные», эвристические и метаэвристические методы решения. Точное решение [1] данной задачи дает «brute force», точнее — метод простого полного перебора, при котором рассматриваются все варианты маршрутов, строится полный их перечень, после чего отбрасываются маршруты с $R_n > R_{\max}$, а среди оставшихся производится поиск наименее затратного по времени. Этот метод найдет идеальный вариант из всего множества возможных маршрутов. Однако такой вариант

неприемлем даже при современных уровнях развития вычислительной техники, поскольку на его проработку потребуется память объемом $n!$ ячеек, где n – число точек в маршруте. По причине неприемлемости точных методов, основной упор в разработке и использовании приходится на метаэвристические [1], точнее, стохастические методы, принцип работы которых можно сформулировать как [3] «множество, выполняющее простые действия, и характеризующееся сложным поведением».

В «муравьином» алгоритме [3] производится рассылка «муравьев» из точки А по направлению к узлу В. На каждом шаге муравьи выбирают маршрут следования вероятностным образом, основываясь на количестве оставленного ранее «феромона», но не назад. По достижении конечной точки, в зависимости от эффективной длины маршрута, пройденные участки отмечаются «феромоном», который испаряется после каждой итерации по заданному закону. Алгоритм проходит заданное число итераций до выдачи маршрута, на котором осталась наибольшая «концентрация феромона» в качестве рекомендуемого. Основная формула для выбора маршрута муравьем (1) выглядит как [3]:

$$p_{ij}^k(t) = \frac{(\tau_{ij}(t))^\alpha (v_{ij})^\beta}{\sum_{l \in J_i^k} (\tau_{il}(t))^\alpha (v_{il})^\beta}, \quad (1)$$

здесь k – номер муравья; i – точка, в которой муравей находится; j – точка, в которую может перейти муравей; p_{ij}^k – вероятность выбора перехода $i - j$ для муравья k ; t – момент (этап) развития колонии муравьев; τ – концентрация феромона на дистанции; v – скорость муравья на переходе $i - j$, обратно пропорциональна длине перехода; J_i^k – список разрешенных для перехода точек; α и β – коэффициенты, показывающие вес феромона и расстояния при выборе перехода.

Известны работы [4], в которых показано применение алгоритма «роя частиц» для задачи управления беспилотным поисково-транспортным средством. В работе рассматривается также поиск «роем» одной цели. Эта цель является «маячком», который притягивает частицы роя. Примечательно, что в методе «роя частиц» каждая частица имеет некоторый вес, подстройкой которого можно добиться прироста эффективности алгоритма, а также два параметра, которые учитывают вес индивидуальных параметров частицы относительно параметров роя в момент принятия решений об ускорении и направлении. Там же доказано, что метод «роя частиц» легко может быть адаптирован к задаче с одним объектом, месторасположение которого в каждый конкретный момент известно. Разница между задачей, рассмотренной в [4], и рассматриваемой авторами состоит в отсутствии в первом случае ограничения по риску, которое накладывает ограничения на выбор маршрута.

О перспективности развития методов «роя частиц» и аналогичных им применительно к задачам логистики говорится также в [5]. Авторами отчета [5] произведен анализ применения «муравьиного» алгоритма и алгоритма «роя частиц» в решении задачи коммивояжера, маршрутизации, и произведена оценка возможности применения стохастических метаэвристических алгоритмов в логистике Вооруженных сил Швеции. Авторы указанного отчета пришли к выводу, что наиболее перспективным является «муравьиный» алгоритм решения задач оптимизации логистики этой отрасли. В [5] показана возможность применения «муравьиного алгоритма» непосредственно для оценки риска в логистике и оценена вычислительная сложность алгоритмов, которая, по оценке Свенсона, для «муравьиного алгоритма» составляет $O(n)$. Из [5] можно сделать заключение о хорошем качестве «муравьиного алгоритма» как метода решения статических задач. Однако, к сожалению, о динамической задаче в [5] не говорится.

Применение муравьиного алгоритма уже для динамической транспортной задачи показано в [6]. При этом подход к решению следующий: все время, которое теоретически максимально коммивояжер может находиться в пути (рабочий день) разбивается на участки, для которых решается статическая транспортная задача и находится оптимальное решение.

Обратим внимание, что часто метаэвристические методы применяются для улучшения решений, найденных иными методами. Для этого может использоваться любая из т.н. метаэвристических методов (к примеру, «метод роя частиц», «имитации отжига» или «муравьиный» алгоритм). Все эти алгоритмы носят вероятностный характер, что обуславливает их относительно медленную сходимость. Для улучшения скорости сходимости алгоритма и исключения блужданий вводятся [7] т.н. «гибридные» методы, в которых первоначальное решение строится с использованием, в частности, нейронной сети, а в дальнейшем улучшается с использованием (мета)эвристических алгоритмов.

3. Неразрешенные части проблемы

Специфика рассматриваемой задачи состоит в том, что, в отличие от классической транспортной задачи в нашем случае имеем точки, которые необходимо обойти. Их всего две – начало и конец маршрута, все же остальные могут добавляться к нему или удаляться по мере необходимости (однако их число может составить сотни и более, что исключает применение точных и эвристических методов), а также существование ограничения по уровню риска, которое нельзя превысить, и ограничение в одну транспортную единицу. Также задача включает ограничение по уровню риска.

Следует учесть факт, что в сформулированном выше виде задача ни в одной из рассмотренных работ поставлена не была, несмотря на то, что все рассмотренные задачи близки ей. Для динамических транспортных задач также не формулировались рекомендации по выбору алгоритма поиска решения, тем более, в условиях переменных и риска и эффективной длины ребра, когда сведение к одному параметру для оптимизации становится невозможным.

Целью данного исследования является оценка возможности решения задачи с применением «муравьиного» алгоритма, как наиболее перспективного из стохастических методов, а также возможность его видоизменения для адаптации к динамическим задачам.

4. Результаты испытания

Для решения динамической задачи была построена модель по методике, описанной в [6], в которой стандартное правило (2) испарения феромона

$$\tau_{ij}(t+1) = (1-\rho)\tau_{ij}(t), \quad (2)$$

где ρ – коэффициент, показывающий скорость испарения феромона, для учета риска было модифицировано в виде (3):

$$\tau_{ij}(t+1) = \begin{cases} (1-\rho)(1-R_{ij})\tau_{ij}(t), & R_{ij} < R_{\max} \\ 0, & R_{ij} \geq R_{\max} \end{cases}, \quad (3)$$

где R_{ij} – уровень риска на соответствующем ребре: $R_{ij} \in (0;1)$.

Модель была проверена на сети из 256 и 1024 точек, в которых количество муравьев, участвующих в поиске маршрута, соответствовало числу узлов, а число итераций соответствовало 1/3 от числа узлов. Результаты представлены в табл. 1.

Таблица 1

Число узлов	Характер функций R(t) и T(t)	Время расчета, с	Среднее время достижения цели (усл. ед. времени)	Оценка входа в тупик
256	периодические	1,86	315,52	1,4 %
1024	периодические	35,84	962,13	2,2 %
256	стохастические	2,09	426,84	7,0 %
1024	стохастические	43,26	1266,38	3,8 %

Сеть представляла собой матрицу точек, в которых связность выбиралась на запуске программы произвольно, но с условием максимальной связности узла не более 5. Уровень риска изменялся при первом прогоне псевдослучайно, при втором – как тригонометрическая функция с периодом, соответствующим максимальному предполагаемому

времени прохождения системы (как и в [7]). Оценивались скорость расчета (на ЦП Intel Core 2 Duo E7500, 2,93 GHz), среднее время прохождения сети в условных единицах (при максимальном времени прохода ребра 5,0 ед.). Вероятность захода в тупик (ситуация, когда вокруг точки нет узлов с риском ниже порогового) определялась как процент тупиковых решений из общего числа запусков (всего 50 с одинаковыми параметрами).

5. Экспертная обучающая система

Для мобильного управления перевозками ОГ с минимумом рисков была разработана информационная система. Эта система может быть использована как экспертная для принятия решений в случае возникновения ЧС, так и для тренинга или обучения персонала, который обеспечивает транспортировку ОГ с определенным риском при возникновении ЧС. В программном обеспечении использовался предложенный выше алгоритм.

В состав системы (рис. 1) входит сервер, на котором алгоритм. Система поддерживает связь с диспетчерским пунктом, водителями транспортных средств и имеет геоинформационную привязку транспортного средства, координаты возникших или возможных ЧС (для обучения персонала).

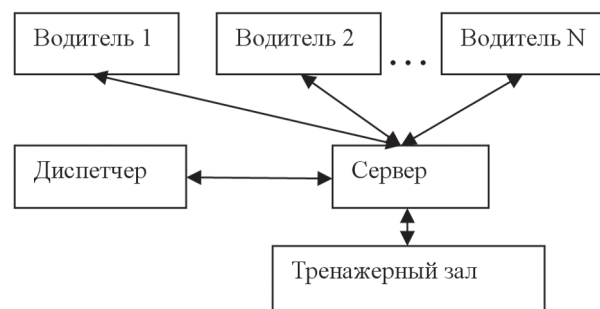


Рис. 1. Возможный вариант системы для мобильного управления перевозками ОГ с минимумом рисков

Система позволяет при помощи мобильного телефона, оборудованного камерой и имеющего доступ в Интернет, получить информацию о текущем месторасположении объекта (отображаемую на одной из Интернет-карт), информацию об объектах, находящихся поблизости.

Такой результат достигается благодаря использованию технологии Microsoft Tag, которая представляет собой базу данных, содержащую множество URL (Universal Resource Locator). Каждому такому URL ставится в соответствие некоторый ключ, а затем этот ключ кодируется специальным изображением – тегом. Изображения, которыми кодируются ключи (теги), похожи на штрих-коды, которые часто размещают на упаковке товара для автоматизированной обработки (например, в больших магазинах).

В отличие от штрих-кодов, для расшифровки которых нужны специальные сканеры, теги могут

быть расшифрованы с помощью камеры невысокого разрешения и специального программного обеспечения. Используя мобильный телефон с камерой и доступом в Интернет, можно отсканировать тег и получить соответствующий URL из базы данных Microsoft Tag. Каждому такому URL ставятся в соответствие описание и географические координаты объекта, на которых расположен тег (в случае, если этот объект не перемещается в пространстве, например здания, памятники и др.). Эту технологию можно использовать и для движущихся объектов, тогда нельзя определить местоположение объекта, но можно узнать о нем дополнительную информацию (например, по тегу на трамвае можно узнать полный маршрут, телефон диспетчерской и т.п.).

Система мобильного управления перевозками ОГ с минимумом рисков состоит из нескольких подсистем:

- клиентской части, имеющей веб-браузер (для мобильного или стационарного использования);
- подсистемы, позволяющий создавать теги для произвольных объектов и обеспечивать связь с объектами;
- базы данных, в которой каждому тегу соответствует необходимая информация;
- веб-сервера мобильных и стационарных устройств для тренинга и обучения персонала;
- геоинформационной он-лайн системы, позволяющей получать пользователю картографическую информацию о местоположении и планируемых маршрутах;
- подсистемы безопасности и администрирования.

Система имеет базу знаний прецедентов с возможными сценариями ЧС. База знаний позволяет сократить время на принятие решений как диспетчеру, так и водителю транспортного средства в случае возникновения ЧС. Клиентская часть системы, установленная на мобильном устройстве, сканирует тег при помощи цифровой фотокамеры, затем тег расшифровывается при помощи соответствующих сервисов. Расшифровав изображение тега, система запускает браузер и передает ему URL (связанный с тегом), по которому веб-сервер, в свою очередь, отыскивает необходимую информацию в базе данных и он-лайн геоинформационной системе, после чего формирует соответствующую HTML-страницу и передает ее на клиентскую часть в мобильное устройство.

Выводы

По анализу табл. 1 можно сделать выводы, что в целом результат работы муравьиного алгоритма в динамической транспортной задаче с риском – удовлетворительный. Время поиска решения оказывается во вполне приемлемых пределах (едини-

цы секунд, редко минуты). Резкое возрастание (до 7%) заходов в тупик на небольшой сети с низкой связностью и стохастическими характеристиками обусловлено тем, что предлагаемое решение в этом случае требует переключения в системе (рис.1) на экспертную оценку принятого решения. В то же время, на большой сети есть вероятность того, что между ребром с резко возросшим риском и положением точки в системе есть промежуточное ребро, за время прохода по которому риски уменьшаются. В то же время применение муравьиного алгоритма для сети с 1000 узлами вызывает возрастание затрат времени на поиск приемлемого решения. В этом случае для решения транспортных задач больших размерностей необходимо использовать для сервера многоядерный или кластерный вычислительный ресурс.

Необходимо также обратить внимание на тот факт, что муравьиный алгоритм находит хорошее по времени решение при заданном риске, о чем свидетельствует возрастание времени прохождения системы, отстающее от возрастания числа точек.

Намного лучшие результаты для условий, когда уровень риска изменяется по известному закону. В этом случае оптимальное решение текущего шага может оказаться достаточно удовлетворительным для следующего шага. Это говорит о том, что для повышения эффективности системы для мобильного управления перевозками ОГ с минимумом рисков необходимо особое внимание уделять при системном анализе изменения риска и его учете в системе. Иными словами, проблема состоит не в выборе алгоритма для поиска маршрута, а в выборе алгоритма для предсказания риска следующих шагов.

Научная новизна исследования заключается в следующем:

- разработаны методы и модели системного анализа транспортной задачи в целом и конкретных маршрутов перевозки для оценки рисков возникновения транспортных происшествий и экономических ущербов от них;
- впервые предложена модель оптимизации принятия решений в условиях неопределенности рисков транспортных ситуаций с использованием муравьиного алгоритма;
- впервые, с учётом взаимодействия с другими видами транспорта (автомобильный, трубопроводный), определены основы принятия решений при управлении транспортными объектами и обучения персонала для условий неопределенности с оценкой рисков возникновения транспортных происшествий (крушения, аварии, теракта, возникновения ЧС) и ущербов от них при перевозке в транспортной сети.

В дальнейшем предполагается усовершенствовать алгоритм учета риска таким образом, чтобы

исключить возможность входа в тупик как таковую (в частности, применением муравьиного же алгоритма для оценки риска, как это сделано в [6]).

Список литературы: 1. *Емельянова, Т.С.* Эвристические и метаэвристические методы решения динамической транспортной задачи [Текст] / *Т.С. Емельянова* // Перспективные информационные технологии и интеллектуальные системы. – 2007. – № 3 (31). 2. *Minimizing logistics risk through real-time vehicle routing and mobile technologies* Research to date and future trends. *G.M. Giaglis, I. Minis, A. Tatarakis, V. Zeimpekis* // *International Journal of Physical Distribution & Logistics Management* Vol. 34 No. 9, 2004. PP. 749-764 3. *Swarm Intelligence for logistics: Background.* *Svenson P., Mertenson Ch., Sidenbladh H., Malm M.* Linköping, 2004. 46 p. 4. *Sheetal, Venayagamoorthy G.K.* Unmanned vehicle navigation using swarm intelligence // *Intelligent Sensing and Information Processing (ICISIP), Proceedings of International Conference on.* 2004, PP. 249–253. 5. *Svenson, P., Sidenbladh, H.* Determining possible avenues of approach using ANTS // *6th International Conference on Information Fusion.* Cairns, Australia, 2003. PP. 1110–1117 6. *A new algorithm for a Dynamic Vehicle Routing Problem based on Ant Colony System.* *Montemanni R., Gambardella L.M., Rizzoli A.E., Donati A.V.* Technical Report IDSIA-23-02, IDSIA. 2002 7. *Solving stochastic job shop scheduling problems by a hybrid*

method. *Pervaiz A., Reza T.-M., Fariborz J., Farzaneh V.* // *University of Wolverhampton Business School.* 2004.

Поступила в редколлегию 17.09.2010.

УДК 004.08

Мурашиний алгоритм та врахування ризику у транспортній задачі / О.Я. Кузьмін, Я. Даюб // *Біоніка інтелекту: наук.-техн. журнал.* – 2010. – № 3 (74). – С. 144–148.

Розглянута специфічна динамічна транспортна задача в умовах ризику. Проводиться моделювання вибору маршруту доставки для умов фіксованого старту й цілі з пошуком оптимального за часом та рівнем ризику маршруту із використанням мурашиного алгоритму для різних умов зміни рівня ризику і тривалості проходження ділянок маршруту.

Табл. 1. Лл. 1. Бібліогр.: 7 найм.

UDC 004.08

Ant pheromone algorithm with risk consideration in vehicle routing task / О. Я. Кузмин, Я. Даюб // *Bionics of Intelligence: Sci. Mag.* – 2010. – № 3 (74). – P. 144–148.

The specific dynamic transportation problem in terms of risk was considered. A modeling of the process of finding the optimal by risk and time route with ant pheromone algorithm under the conditions of fixed start and goal and various conditions of change of the level of risk and duration of passage route parts has been done.

Tab. 1. Fig. 1. Ref.: 7 items.

ОБ АВТОРАХ

Аврунин Олег Григорьевич	99	канд. техн. наук, доцент кафедры биомедицинских электронных устройств и систем Харьковского национального университета радиоэлектроники
Аксак Наталья Георгиевна	94	канд. техн. наук, доцент кафедры ЭВМ Харьковского национального университета радиоэлектроники
Алексеев Владимир Олегович	52	д-р техн. наук, старший научный сотрудник, профессор кафедры мехатроники автотранспортных средств Харьковского национального автомобильно-дорожного университета
Алексеев Олег Павлович	52	д-р техн. наук, профессор, заведующий кафедрой мехатроники автотранспортных средств Харьковского национального автомобильно-дорожного университета
Бескорвайный Владимир Валентинович	37	д-р техн. наук, профессор кафедры системотехники Харьковского национального университета радиоэлектроники
Бондаренко Михаил Федорович	3, 8	член-корреспондент НАН Украины, д-р техн. наук, профессор, ректор Харьковского национального университета радиоэлектроники
Валенда Наталья Анатольевна	14	канд. техн. наук, доцент кафедры программного обеспечения ЭВМ Харьковского национального университета радиоэлектроники
Вечирская Ирина Дмитриевна	14	канд. техн. наук, старший научный сотрудник кафедры программного обеспечения ЭВМ Харьковского национального университета радиоэлектроники
Гороховатский Владимир Алексеевич	58	канд. техн. наук, профессор кафедры информатики Харьковского национального университета радиоэлектроники
Губаренко Евгений Витальевич	26	аспирант кафедры системотехники Харьковского национального университета радиоэлектроники
Губин Вадим Александрович	109	ст. преподаватель кафедры искусственного интеллекта Харьковского национального университета радиоэлектроники
Гуляева Наталья Анатольевна	62	ассистент кафедры прикладной математики и информационных технологий в бизнесе Института бизнеса, экономики и информационных технологий, Одесский национальный политехнический университет
Даюб Ясир	144	аспирант Харьковского национального университета радиоэлектроники
Довбыш Анатолий Степанович	129	д-р техн. наук, профессор, заведующий кафедрой информатики Сумского государственного университета
Жемчужкина Татьяна Владимировна	99	канд. техн. наук, старший преподаватель кафедры биомедицинских электронных устройств и систем Харьковского национального университета радиоэлектроники
Зайцев Сергей Алексеевич	70	аспирант кафедры программных средств Запорожского национального технического университета
Зыкова Ольга Владимировна	37	студентка кафедры системотехники Харьковского национального университета радиоэлектроники

Калмыков Андрей Викторович	42	канд. техн. наук, докторант Национального аэрокосмического университета им. Н.Е.Жуковского «ХАИ»
Кашубин Станислав Петрович	120	научный сотрудник Института проблем машиностроения НАН Украины
Келеберда Игорь Николаевич	133	канд. техн. наук, доцент кафедры программного обеспечения ЭВМ Харьковского национального университета радиоэлектроники
Кирпуть Юрий Анатольевич	133	студент кафедры программного обеспечения ЭВМ Харьковского национального университета радиоэлектроники
Клочкова Александра Александровна	83	студентка кафедры программных средств Запорожского национального технического университета
Козленко Николай Викторович	19	магистрант Ecole Polytechnique, Париж, Франция
Котов Андрей Сергеевич	76	канд. техн. наук, старший научный сотрудник кафедры системотехники Харьковского национального университета радиоэлектроники
Кругликова Наталья Павловна	3, 8	аспирант кафедры программного обеспечения ЭВМ Харьковского национального университета радиоэлектроники
Крылов Виктор Николаевич	62	д-р техн. наук, профессор кафедры прикладной математики и информационных технологий в бизнесе Института бизнеса, экономики и информационных технологий, Одесский национальный политехнический университет
Куземин Александр Яковлевич	144	д-р техн. наук, профессор кафедры информатики, начальник инновационно-маркетингового отдела Харьковского национального университета радиоэлектроники
Лесная Наталия Советовна	133, 138	канд. техн. наук, профессор кафедры программного обеспечения ЭВМ Харьковского национального университета радиоэлектроники
Лещинская Ирина Александровна	3, 8	аспирант кафедры программного обеспечения ЭВМ Харьковского национального университета радиоэлектроники
Маслянюк Павел Павлович	19	канд. техн. наук, доцент, зам. декана факультета прикладной математики Национального технического университета Украины «КПИ»
Мартыненко Сергей Сергеевич	129	аспирант кафедры информатики Сумского государственного университета
Мирошниченко Татьяна Александровна	90	старший преподаватель кафедры менеджмента Харьковского института экономики рыночных отношений и менеджмента
Найденко Андрей Вячеславович	138	студент кафедры программного обеспечения ЭВМ Харьковского национального университета радиоэлектроники
Новосельцев Игорь Валериевич	94	аспирант кафедры ЭВМ Харьковского национального университета радиоэлектроники
Носова Татьяна Витальевна	99	канд. техн. наук, доцент кафедры биомедицинских электронных устройств и систем Харьковского национального университета радиоэлектроники

Олейник Алексей Александрович	83	ассистент кафедры программных средств Запорожского национального технического университета
Олейник Андрей Александрович	83	канд. техн. наук, старший преподаватель кафедры программных средств Запорожского национального технического университета
Петров Эдуард Георгиевич	26	д-р техн. наук, профессор, заведующий кафедрой системотехники Харьковского национального университета радиоэлектроники
Половенко Карина Григорьевна	105	стажер-исследователь кафедры биомедицинских электронных устройств и систем Харьковского национального университета радиоэлектроники
Полякова Марина Вячеславовна	62	канд. техн. наук, доцент кафедры прикладной математики и информационных технологий в бизнесе Института бизнеса, экономики и информационных технологий, Одесский национальный политехнический университет
Русакова Наталья Евгеньевна	3, 8	аспирант кафедры программного обеспечения ЭВМ Харьковского национального университета радиоэлектроники
Рябушенко Андрей Виталиевич	19	аспирант кафедры прикладной математики факультета прикладной математики Национального технического университета Украины «КПИ»
Сокол Виталий Викторович	138	ассистент кафедры программного обеспечения ЭВМ Харьковского национального университета радиоэлектроники
Субботин Сергей Александрович	70	канд. техн. наук, доцент кафедры программных средств Запорожского национального технического университета
Тянянский Сергей Станиславович	112	канд. техн. наук, доцент кафедры ЭВМ Харьковского национального университета радиоэлектроники
Хабаров Владимир Остапович	52	аспирант кафедры мехатроники автотранспортных средств Харьковского национального автомобильно-дорожного университета
Четвериков Григорий Григорьевич	52	д-р техн. наук, профессор кафедры программного обеспечения ЭВМ Харьковского национального университета радиоэлектроники
Шабанов-Кушнарченко Юрий Петрович	3, 8	д-р техн. наук, профессор кафедры программного обеспечения ЭВМ Харьковского национального университета радиоэлектроники
Шергин Вадим Леонидович	90	канд. техн. наук, доцент кафедры искусственного интеллекта Харьковского национального университета радиоэлектроники
Шрестха Светлана Николаевна	133	канд. техн. наук, доцент кафедры программного обеспечения ЭВМ Харьковского национального университета радиоэлектроники

ПРАВИЛА ОФОРМЛЕНИЯ РУКОПИСЕЙ ДЛЯ АВТОРОВ НАУЧНО-ТЕХНИЧЕСКОГО ЖУРНАЛА «БИОНИКА ИНТЕЛЛЕКТА»

Научно-технический журнал «Бионика интеллекта» принимает для опубликования написанные специально для него оригинальные рукописи, которые нигде раньше не издавались. Структура рукописи должна быть следующей: индекс УДК, заглавие, сведения об авторах, аннотация, ключевые слова, введение, основной текст статьи, выводы, список использованной литературы.

Согласно Постановлению ВАК Украины от 15.01.2003 №7-05/1 (Бюллетень ВАК, №1, 2003, с. 2), статья должна иметь следующие необходимые элементы: постановка проблемы в общем виде и ее связь с важными научными или практическими заданиями; анализ последних исследований, публикаций и выделение нерешенных ранее частей общей проблемы в данной области; формулирование целей и задач исследования; изложение основного материала исследований с полным обоснованием полученных научных результатов; выводы по данному исследованию и перспективы дальнейших исследований в данном направлении.

Статьи должны быть выполнены в редакторе Microsoft Word Формат страницы – А4 (210x297 мм), поля: верхнее – 25 мм, нижнее – 20 мм, левое, правое – 17 мм. Количество колонок – 2, с интервалом между ними 5 мм, основной шрифт Times New Roman, кегль основного текста – 10 пунктов, междустрочный интервал – множитель (1,1), абзацный отступ – 6 мм. Объем рукописи – от 4 до 12 страниц (языки: русский, украинский, английский).

УДК печатается с первой строки, без отступа, выравнивание по левому краю.

Название рукописи печатается прописными буквами; шрифт прямой, полужирный, кегль 12.

Название разделов нумеруют арабскими цифрами, выделяют жирным шрифтом. Отступы для названия рукописи, инициалов и фамилий авторов, сведений об авторах, названий разделов, введения и выводов, списка литературы сверху – 6 пт, снизу – 3 пт.

Аннотация (на языке рукописи, абзац 4-10 строк, кегль 9) помещается в начале статьи и содержит информацию о результатах описанных исследований.

Ключевые слова (4-10 слов из текста статьи, которые с точки зрения информационного поиска несут смысловую нагрузку, на языке рукописи, через запятую в именительном падеже, кегль 9)

Рисунки и таблицы (черно-белые, контрастные) помещаются в текст после первой ссылки в виде встроенных объектов и отдельно нумеруются арабскими цифрами сквозной нумерацией, при наличии более одного объекта. Все надписи в рисунках и таблицах должны быть выполнены шрифтом Times New Roman, кегль в рисунках – 10, в таблицах – 9. Табличный заголовок располагается справа над таблицей (кегль 9 пт).

Формулы, символы, переменные, которые нельзя набрать в Microsoft Word, должны быть набраны в редакторе формул MathType или Microsoft Equation. Формулы располагаются по центру и нумеруются при наличии ссылок на них в рукописи. Шрифт – Times New Roman. Высота переменной – 10 пунктов, индексов и подиндексов – 8 пт, основной математический символ – 12 (10) пт. Переменные, обозначенные латинскими буквами, набирают курсивом, греческие буквы, сокращения русских слов и цифры – прямым начертанием.

Список литературы включает опубликованные источники, на которые имеются ссылки в тексте, заключенные в квадратные скобки, печатается без абзацного отступа, кегль 9 пунктов, отступ сверху – 6 пт.

После списка литературы с отступами сверху – 6 пт указывается дата поступления статьи в редколлегию. Число и месяц задаются двузначными числами через точку. Размер шрифта – 9 пт, курсив, выравнивание по правому краю.

Рефераты (Times New Roman, кегль - 9 пунктов, 3–4 предложения) на украинском и английском языках. Реферат не должен дублировать текст аннотации.

Вместе с рукописью (на листах белой бумаги формата А4 плотностью 80–90 г/м², напечатанный на лазерном принтере, в 2-х экземплярах) необходимо подать следующие документы: заявку утвержденного образца, подписанную всеми авторами; акт экспертизы о возможности опубликования материалов в открытой печати; рецензию; сведения об авторах; электронный вариант рукописи, реферата и сведений об авторах; оплату за публикацию.

Редакция журнала «Бионика интеллекта» обращает внимание авторов на необходимость составления списка литературы в соответствии с новым ДСТУ ГОСТ 7.1:2006.

Примеры библиографического описания

Книга одного автора

Андреев, В. В. Как организовать делопроизводство на предприятии [Текст] / В. В. Андреев. – М. : ИНФРА-М, 1997. – 94 с.

Книга двух авторов

Белов, А. В. Финансы и кредит [Текст] : учеб. / А. В. Белов, В. Н. Николаев ; КНУ им. Т. Г. Шевченко. – К. : Университет, 2004. – 215 с. – Библиогр.: с. 213–215. – ISBN 5-7042-1441-X.

Гроссе, Э. Химия для любознательных [Текст] : пер. с нем. / Э. Гроссе, Х. Вайсменталь. – М.: Химия, 1980. – 392с.

Книга трех авторов

Агафонова, Н. Н. Гражданское право [Текст] : учеб. пособие / Н. Н. Агафонова, Т. В. Богачева, Л. И. Глушкова; под общ. ред. А. Г. Калпина; МОН Украины. – 2-е изд., перераб. и доп. – Х. : Фактор, 2000. – 542 с. – (Университетская книга).

Бахвалов, Н. С. Численные методы [Текст] : учеб. пособие для физ.-мат. специальностей вузов / Н. С. Бахвалов, Н. П. Жидков, Г. М. Кобельков ; под общ. ред. Н. И. Тихонова. — 2-е изд. — М.: Физматлит : Лаб. базовых знаний; СПб. : Нев. диалект, 2002. — 630 с. : ил. ; 25 см. — (Технический университет. Математика). — Библиогр.: с. 622–626. — Предм. указ.: с. 627–630. — 30000 экз. — ISBN 5-93208-043-4 (в пер.).

Книга четырех авторов

Элементи інформатики [Текст] : довідник / В. С. Височанський, А. І. Кардаш, В. С. Костев, В. В. Черняхівський. — К. : Наук. думка, 2003. — 192 с.

Книга пяти авторов и более

Коротковолновые антенны [Текст] : учеб. пособие / Г. З. Айзенберг, С. П. Белоусов, Я. М. Журбин и др. ; под общ. ред. А. А. Стогния. — 2-е изд. — М. : Радио и связь, 2003. — 192 с.

Переводное издание

Нойман, Э. Происхождение и развитие сознания [Текст] : пер. с англ. — К. : Ваклер ; М. : Реал-бук, 1998. — 462 с.

Книги под заглавием

Информационные технологии в маркетинге [Текст] : учеб. / под ред. Г. А. Титаренко. — М. : ЮНИТИ, 2000. — 335 с. — (Textbook). — ISBN 5-238-00154-1.

Многотомное издание, отдельный том

Савельев, И. В. Курс общей физики [Текст]. Т. 1. Механика. Молекулярная физика : учеб. Пособие / И. В. Савельев. — 2-е изд., перераб. — М.: Наука, 1982. — 432 с.

Статьи из журналов

Гончаров, В. А. Численная схема моделирования дозвуковых течений вязкого сжимаемого газа [Текст] / В. А. Гончаров, В. М. Кравцов // Журн. вычисл. математики и мат. физики. — 1988. — Т. 28, № 12. — С. 1858–1866.

Сборники научных трудов

Отчет о выполнении плана научно-исследовательских работ за 2003 год [Текст] : сб. науч. тр. / Рос. Акад. мед. наук, Сиб. отд. — Новосибирск : СО РАМН, 2004. — 83 с.

Тезисы конференций

Образование, наука, производство: пути углубления интеграции и повышения качества инженерного образования [Текст] : тез. докл. науч.-практ. конф. (окт. 2000) / отв. ред. В. Г. Вдовенко. — Красноярск : САА, 2000. — 53 с.

Материалы конференций

Проблемы экономики, организации и управления реструктуризацией и развитием предприятий промышленности, сферы услуг и коммунального хозяйства [Текст] : материалы IV междунар. науч.-практ. конф., 30 марта 2005 г. Новочеркасск / редкол. : Б. Ю. Серебряков (отв. ред.). — Новочеркасск : Темп, 2005. — 58 с.

Стандарты, технико-экономические и технические документы

ГОСТ Р 517721-2001. Аппаратура радиоэлектронная бытовая. Входные и выходные параметры и типы соединений. Технические требования [Текст]. — Введ. 2002-01-01. — М. : Изд-во стандартов, 2001. — 27 с.

Инструкция по проектированию, строительству и эксплуатации гидротехнических сооружений на подра-

батываемых горными работами территориях [Текст] : СН 522-85. — Утв. Госстроем СССР 03.05.86. — Изд. офиц. — М.: Стройиздат, 1986. — 32 с.

Авторские свидетельства, патенты

А. с. 1007970 СССР, МПК В 25 J 15/00. Устройство для захвата деталей [Текст] / Ваулин В. С., Калов В. К. (СССР). — 3360585/25-08 ; заявл. 23.11.81 ; опубл. 30.03.83, Бюл. № 12. — 2 с.

Приемопередающее устройство [Текст] : пат. 2187888 Рос. Федерация: МПК7 Н 04 В 1/38, Н 04 J 13/00 / Чугаева В. И. ; заявитель и патентообладатель Воронеж. науч.-исслед. ин-т связи. — № 2000131736/09 ; заявл. 18.12.00; опубл. 20.08.02, Бюл. № 32. — 3 с.

Отчеты о НИР

Оценка эффективности автоматизированных ИПС [Текст]: отчет о НИР (промежуточ.) / ВНИЦентр ; рук. М. И. Иванов. — ОЦ 02604-И5В ; ГР 01821100006 ; Инв. Б452743. — М., 1982. — 90 с.

Диссертации

Антопольский, А. Б. Лингвистическое описание и оценка информационных языков [Текст] : дис. ... канд. филол. наук / А. Б. Антопольский. — М., 1969. — 404 с.

Белозеров, И. В. Религиозная политика Золотой Орды на Руси в XIII-XIV вв. [Текст] : дис. ... канд. ист. наук : 07.00.02 : защищена 22.01.02 : утв. 15.07.02 / Белозеров Иван Валентинович. — М., 2002. — 215 с. — Библиогр.: с. 202-213. — 04200201565.

Авторефераты диссертаций

Бутковский, О. Я. Обратные задачи хаотичной динамики и проблемы предсказуемости хаотичных процессов [Текст] : автореф. дис. ... д-ра физ.-мат. наук : 01.04.03 / О. Я. Бутковский ; [Ин-т радиотехники и электроники РАН]. — М., 2004. — 39 с.

Электронные ресурсы

Удаленного доступа

Основные направления исследований, основанные на семантическом анализе текстов [Электронный ресурс] / С.-Петерб. гос. ун-т, фак. прикладной математики — процессов управления. — Режим доступа : <http://apcp.apmath.spbu.ru/ru/staff/tuzov/onapr.html/> — 10.12.2004 г. — Загл. с экрана.

Локального доступа

Internet шаг за шагом [Электронный ресурс] : интеракт. учеб. — Электрон. дан. и прогр. — СПб. : Питер Ком, 1997. — 1 электрон. опт. диск (CD-ROM). — Систем. требования: ПК от 486 DX 66 МГц; RAM 1616 Мб; Windows 95; зв. плата. — Загл. с этикетки диска.

Художественная энциклопедия зарубежного классического искусства [Электронный ресурс]. — Электрон. текстовые, граф., зв. дан. и прикладная прогр. (546 Мб). — М. : Большая Рос. энцикл. [и др.], 1996. — 1 электрон. опт. диск (CD-ROM) : зв., цв. ; 12 см + рук. пользователя (1 л.) + открытка (1 л.). — (Интерактивный мир). — Систем. требования: ПК 486 или выше; 8 Мб ОЗУ; Windows 3.1 или Windows 95; SVGA 32768 и более цв.; 640x480; 4x CD-ROM дисковод; 16-бит. зв. карта; мыш. — Загл. с экрана. — Диск и сопровод. материал помещены в контейнер 20x14 см.

Наукове видання

БІОНІКА ІНТЕЛЕКТУ
інформація, мова, інтелект

Науково-технічний журнал

№ 3 (74)

2010

Головний редактор — *М. Ф. БОНДАРЕНКО*

Відповідальний редактор — *Ю. П. Шабанов-Кушнарєнко*

Заступник відповідального редактора — *Г. Г. Четвериков*

Відповідальний секретар — *І. Д. Вечірська*

Коректор — *Л. М. Денісова*

Комп'ютерна верстка — *О. Б. Ісаєва*

Рекомендовано Вченою Радою
Харківського національного університету радіоелектроніки
(протокол № 66 від 29.10.2010)

Адреса редакції:

Україна, 61166, Харків-166, просп. Леніна, 14,
Харківський національний університет радіоелектроніки, к. 127, 285
тел. 702-14-77, факс 702-10-13,
e-mail: bionics@kture.kharkov.ua

Підписано до друку 29.10.2010. Формат 60 × 84 ¹/₈. Друк ризографічний.
Папір офсетний. Гарнітура Newton. Умов. друк. арк. 17,7. Обл.-вид. арк. 17,0.
Тираж 100 прим. Зам. № .

Надруковано в навчально-науковому видавничо-поліграфічному центрі ХНУРЕ
61166, Харків-166, просп. Леніна, 14