

ИНФОРМАЦИОННЫЕ ТЕХНОЛОГИИ

12(172)
2010

ТЕОРЕТИЧЕСКИЙ И ПРИКЛАДНОЙ НАУЧНО-ТЕХНИЧЕСКИЙ ЖУРНАЛ

Издается с ноября 1995 г.

УЧРЕДИТЕЛЬ

Издательство "Новые технологии"

СОДЕРЖАНИЕ

ОБЩИЕ ВОПРОСЫ

Гуревич И. М. Информационные методы исследования физических систем: обзор первых достижений. 2

БЕЗОПАСНОСТЬ ИНФОРМАЦИИ И УПРАВЛЕНИЕ РИСКАМИ

Алгулиев Р. М., Имамвердиев Я. Н., Деракшанде С. А. Пути повышения точности методов оценки рисков информационной безопасности. 6

Генералов Д. Н. Противодействие нарушению целостности модулей специализированного программного обеспечения бортового устройства. 12

ВЫЧИСЛИТЕЛЬНЫЕ СИСТЕМЫ И СЕТИ

Саак А. Э. Локально-оптимальный синтез расписаний для GRID-технологий. 16

Умбиталиев А. А., Шипилов Н. Н., Фахми Ш. С. Сложнофункциональный блок транскодирования видеоконтента. 21

Панько С. П., Нагорный А. В. Особенности многостанционного доступа в сеть с разделяемым ресурсом. 24

Аббасова Т. С. Повышение эффективности эксплуатации высокоскоростной кабельной системы с помощью виртуальных технологий. 28

ГЕОИНФОРМАЦИОННЫЕ СИСТЕМЫ

Струченков В. И., Козлов А. Н., Егунов А. С. Кусочно-линейная аппроксимация плоских кривых при наличии ограничений. 32

Гальяно Ф. Р. Алгоритм классификации участков поверхности Земли на основе сингулярного разложения матриц. 35

ИНФОРМАЦИОННЫЕ ТЕХНОЛОГИИ В ЭКОНОМИКЕ И УПРАВЛЕНИИ

Казаков П. В. Автоматизация синтеза оптимальных инвестиционных портфелей на основе кластерного генетического алгоритма. 38

Чучуева И. А. Модель экстраполяции временных рядов по выборке максимального подобию. 43

Гюльмамедов Р. Г. Выбор стратегии информатизации инновационного региона. Когнитивный подход. 48

ИНФОРМАЦИОННЫЕ ТЕХНОЛОГИИ В МЕДИЦИНЕ

Малыхин В. М., Меркушева А. В. Нейронные сети с вейвлет-преобразованием для классификации биосигналов. 51

Зыкин С. В., Полуянов А. Н., Чернышев А. К., Ревзин А. И. Технология подготовки и анализа данных для построения медицинских оценочных шкал. 57

ПРОГРАММНЫЕ ПРОДУКТЫ И БАЗЫ ДАННЫХ

Киселева Н. Н., Дударев В. А. База данных "Информационные ресурсы неорганической химии и материаловедения". 63

ДИСКУССИОННЫЙ КЛУБ

Вяткин В. Б. Введение в синергетическую теорию информации. 67

Указатель статей, опубликованных в журнале "Информационные технологии" в 2010 г. 74

Указатель приложений к журналу "Информационные технологии", опубликованных в 2010 г. 78

Contents. 78

Приложение. Шилов В. В. Говорящие головы. Мифы и реальность в истории механических генераторов речи

Главный редактор
НОРЕНКОВ И. П.

Зам. гл. редактора
ФИЛИМОНОВ Н. Б.

Редакционная
коллегия:

АВДОШИН С. М.
АНТОНОВ Б. И.
БАТИЩЕВ Д. И.
БАРСКИЙ А. Б.
БОЖКО А. Н.
ВАСЕНИН В. А.
ГАЛУШКИН А. И.
ГЛОРИОЗОВ Е. Л.
ДОМРАЧЕВ В. Г.
ЗАГИДУЛЛИН Р. Ш.
ЗАРУБИН В. С.
ИВАННИКОВ А. Д.
ИСАЕНКО Р. О.
КОЛИН К. К.
КУЛАГИН В. П.
КУРЕЙЧИК В. М.
ЛЬВОВИЧ Я. Е.
МАЛЬЦЕВ П. П.
МЕДВЕДЕВ Н. В.
МИХАЙЛОВ Б. М.
НЕЧАЕВ В. В.
ПАВЛОВ В. В.
ПУЗАНКОВ Д. В.
РЯБОВ Г. Г.
СОКОЛОВ Б. В.
СТЕМПКОВСКИЙ А. Л.
УСКОВ В. Л.
ФОМИЧЕВ В. А.
ЧЕРМОШЕНЦЕВ С. Ф.
ШИЛОВ В. В.

Редакция:

БЕЗМЕНОВА М. Ю.
ГРИГОРИН-РЯБОВА Е. В.
ЛЫСЕНКО А. В.
ЧУГУНОВА А. В.

Информация о журнале доступна по сети Internet по адресу <http://www.informika.ru/text/magaz/it/> или <http://novtex.ru/IT>.

Журнал включен в систему Российского индекса научного цитирования.

Журнал входит в Перечень научных журналов, в которых по рекомендации ВАК РФ должны быть опубликованы научные результаты диссертаций на соискание ученой степени доктора и кандидата наук.

УДК 004.3, 681.511

И. М. Гуревич, канд. техн. наук, ст. науч. сотр.,
Институт проблем информатики РАН, г. Москва
e-mail: iggurevich@gmail.com

Информационные методы исследования физических систем: обзор первых достижений

Приводится обзор работ, использующих информационные методы исследования физических систем. Приведены основополагающие работы данного направления исследований. Перечислены результаты, полученные автором, в том числе результаты, опубликованные в журнале "Информационные технологии". Дан обзор зарубежных работ. Показано, что работы автора и зарубежных ученых (американских, канадских, европейских, китайских ...) подтверждают первичность информационных законов. Информационные законы (законы информатики) имеют всеобщий, универсальный характер, действуют во всех возможных Вселенных, даже во Вселенных с разными физическими законами. Приведенные сведения показывают, что приоритет получения информационными методами физических результатов принадлежит автору, хотя последние результаты зарубежных ученых очень интересны и важны.

Ключевые слова: информация, информационные методы, информационный подход, информационные характеристики, физические системы, информационные и физические законы природы, Вселенная, результаты

Введение

Наряду с материей и энергией Вселенная содержит и информацию, т. е. информация является неотъемлемой частью Вселенной. Каждая физическая система наряду с физическими характеристиками имеет информационные характеристики. Информация неразрывно связана с материей и энергией. Физические системы неоднородны. Физическая неоднородность — это информация, информация — это физическая неоднородность. Наличие информации во Вселенной приводит к необходимости использования информационных методов исследования как самой информации, так и связанной с ней материи и энергии.

Основные физические понятия: пространство, время, частица, система, координаты, физическая величина (наблюдаемая), волновая функция, масса,

координаты, скорость, энергия, импульс, гамильтониан, унитарный оператор, спин, температура, энтропия, давление, объем, заряд, поле, сила... Основные информационные понятия: пространство, время, система, координаты, источник, получатель, сигнал, носитель сигнала, скорость передачи, пропускная способность, информационная энтропия, совместная информационная энтропия, информация связи, бит, кубит, достоверность, надежность, простота, сложность, объем информации, сжатие информации, производительность...

Взаимосвязь между физическими и информационными характеристиками физических систем, в частности массой и информацией, дает возможность использовать информационные оценки и методы исследования традиционных (физических) характеристик физических систем. Использование *информационного подхода* наряду с физическим позволяет получить новые, порой более общие сведения, по отношению к сведениям, получаемым на основе только физических законов. Информационные законы совместно с физическими законами являются эффективным инструментом познания Вселенной.

Обзор основополагающих результатов

Многие выдающиеся ученые отмечали важность информации. Среди них:

Дж. Уиллер: "Моя жизнь в физике представляется мне разделенной на три периода. В первый из них, растянувшийся с начала моей карьеры и до начала 1950-х годов, я был захвачен идеей, что "Все — это частицы". Я искал способы выстроить все базовые элементы материи (нейтроны, протоны, мезоны и т. д.) из самых легких, наиболее фундаментальных частиц — электронов и фотонов. Вторым периодом я называю "Все — это поля". Теперь же я захвачен новой идеей: "Все — это информация" [1];

Э. Стин написал: "...необходимо определить несколько законов, схожих с законами сохранения энергии и момента, но используемых по отношению к информации и определяющих большую часть квантовой механики" [2];

Б. Б. Кадомцев: "При переходе к изучению все более сложных систем именно структурные, информационные аспекты их поведения и развития выступают на первый план, а динамика создает лишь основу для информационного развития. С учетом

квантовых процессов в микромире картина развития мира становится еще более сложной и более богатой в смысле ее информационного поведения" [3].

В настоящей работе излагаются результаты, полученные информационными методами советскими (российскими) и зарубежными авторами. Показаны приоритеты.

Впервые анализ физических процессов с использованием информационных понятий провел *А. Эйнштейн*. Пользуясь понятием сигнала, он установил, что нужно понимать под синхронно идущими, находящимися в различных местах покоящимися часами, и определил понятия *одновременность* и *время* [4].

Л. Сциллард, анализируя мысленный эксперимент с "демоном Максвелла", показал, что энтропия, теряемая газом за счет разделения молекул на медленные и быстрые, в точности равна информации, получаемой "демоном Максвелла". Другими словами, сумма энтропии и информации в системе "газ—наблюдатель" оказалась постоянной величиной. Физическая характеристика становится мерой познания. Причем наблюдатель узнает о системе ровно столько, сколько она теряет [5].

Д. фон Нейман ввел понятие *квантовой энтропии*: $S = -\rho \ln \rho$, где ρ — матрица плотности. Квантовая энтропия, или энтропия Неймана, оценивает неопределенность смешанных состояний [6]. Энтропия Неймана чистого состояния по определению равна нулю. Но физики, как правило, для описания и исследований квантовых систем используют энтропию Неймана.

К. Шеннон ввел понятие *информационная энтропия*. Энтропия H дискретной случайной величины определяется в битах (в натах): $H = -\sum p_i \log_2 p_i$ ($H = -\sum p_i \ln p_i$). Энтропия H непрерывной случайной величины определяется в битах (натах) [7]:

$$H(x) = -\int p(x) \log_2 p(x) dx \quad (H(x) = -\int p(x) \ln p(x) dx,$$

где p_i — вероятность реализации i -го исхода. Информационная энтропия Шеннона является универсальной мерой неопределенности (информации) в классических и квантовых системах.

Е. Т. Джейнс предложил метод отыскания равновесных распределений вероятности с использованием принципа максимума энтропии при наличии дополнительных условий. Для учета дополнительных условий (их может быть несколько) применяется метод множителей Лагранжа [8].

Систематическое применение методов теории информации к анализу физических явлений и процессов, по-видимому, впервые проведено *Л. Бриллюэном*. "Мы введем теперь различие между двумя видами информации: 1) свободная информация (free information), возникающая, когда возможные

случаи рассматриваются как абстрактные и не имеющие определенного физического значения; 2) связанная информация (bound information), возникающая, когда возможные случаи могут быть представлены как микросостояния физической системы". *Л. Бриллюэн* показал, что одна двоичная единица информации соответствует энтропии $\Delta S = k \ln 2$ и должна быть оплачена негэнтропией, большей чем $k \ln 2$, и энергией, не меньшей чем $kT \ln 2$. *Л. Бриллюэн* получил оценку объема информации, содержащейся в физическом законе [9].

Х. Эверетт первым получил информационное представление соотношения неопределенности Гейзенберга [10].

Р. Пенроуз, *С. Хокинг (Hawking)*, *Ж. Бекенштейн*, *Дж. Прескилл*, *Д. А. Кирижниц* и др. использовали информационный подход применительно к процессу образования черных дыр. "Может ли исчезать информация при образовании черной дыры? Куда она может исчезать? Черная дыра искажает проглоченную информацию, но все же не разрушает ее бесследно. В процессе испарения черной дыры информация вырывается из ее объятий" [11–12].

Н. Марголис и *Л. Левитин* доказали теорему о том, что общее число элементарных действий, которые система может выполнить в секунду, ограничено энергией: $k_{\text{оп/с}} = 2E/\hbar$, где E — превышение средней энергии системы над энергией нижнего состояния; $\hbar = \hbar/2\pi = 1,0545 \cdot 10^{-34}$ с·Дж — уменьшенная постоянная Планка [13].

А. Zeilinger выдвинул следующий принцип как возможный фундамент всей квантовой теории. Им представлены две формулировки принципа: 1) элементарная система представляет истинностное значение одного суждения; 2) элементарная система несет один бит информации [14].

С. Ллойд выдвинул следующие постулаты: 1) теорема Марголиса—Левитина; 2) общее количество битов $K_{\text{бит}}$, доступных для обработки в системе, ограничено энтропией системы: $K_{\text{бит}} = S/k \log_2 2$, где S — термодинамическая энтропия системы; $k = 1,38 \cdot 10^{-23}$ Дж/К — постоянная Больцмана; 3) скорость, с которой информация может быть перемещена с места на место, ограничена скоростью света $c = 2,98 \cdot 10^{10}$ м/с. Эти три предела применены для оценки способности Вселенной обрабатывать информацию. В частности, оценены и общее число битов, доступных во Вселенной для вычисления, и число элементарных логических действий, которые могут быть выполнены на этих битах за время существования Вселенной. Общее число битов $\approx 10^{90}$ в материи (результат *И. М. Гуревича*, 1989 г. [16]), $\approx 10^{120}$, если тяготение принято во внимание), и $\approx 10^{120}$ операций.

S. Hsu и A. Zee, D. Scott и J. Zibin обсуждают проблему представления послания Создателя неоднородностями реликтового излучения и оценивают объем послания. Они считают, что 10^5 бит информации [17], 300 бит [18] могут быть закодированы в космическом микроволновом фоне.

Автор занимается систематизацией знаний по сложным системам, информационным методам их исследований в виде законов информатики и проведением исследований сложных систем на основе законов информатики [16, 19–22]. К основополагающим можно отнести следующие результаты автора: *утверждение о существовании законов природы, более общих, чем физические, — законов информатики; законов, определяющих, ограничивающих физические явления и процессы, предшествующие физическим законам. Формулировка законов информатики.* Необходимо отметить, что число ученых, использующих информационный подход и информационные методы в физических исследованиях, быстро возрастает [17–18, 24]. В начале 2010 г. появилось много интересных работ. Среди них работы [25–29].

Обзор первых достижений

Подход 1. (Verlinde, Macccone, Lisi). Исходные положения:

- Первый закон термодинамики.
- Второй закон термодинамики, Голографический принцип, формула Хокинга, эффект Унру.
- Голографический принцип — предположение, выдвинутое в 1993 г. нидерландским физиком-теоретиком Г. Хофтом (Gerard't Hooft). Оно состоит из двух основных утверждений: 1) вся информация, содержащаяся в некоей области пространства, может быть представлена как "голограмма" — теория, которая "живет" на границе этой области; 2) теория на границах исследуемой области пространства должна содержать, самое большее, одну степень свободы (один бит) на планковскую площадь.

Объем информации в физических системах оценивается на основе голографического принципа.

- Формула Хокинга: излучению вблизи горизонта событий черной дыры можно сопоставить оп-

$$\text{ределенную температуру } T_{\text{вн}} = \frac{\hbar c^3}{8\pi k G M}.$$

Здесь и далее:

$T_{\text{вн}}$ — температура излучения черной дыры;

$\hbar$ — уменьшенная постоянная планка;

c — скорость света;

G — гравитационная постоянная;

k — постоянная Больцмана;

M — масса черной дыры.

- Эффект Унру (излучение Унру) — предсказываемый квантовой теорией поля эффект наблю-

дения теплового излучения в ускоряющейся системе отсчета при отсутствии этого излучения в инерциальной системе отсчета [23]. Температура наблюдаемого излучения Унру выражается той же формулой, что и температура излучения Хокинга, но зависит не от поверхностной гравитации, а от ускорения системы отсчета:

$$T = \frac{\hbar a}{2\pi k c}, \text{ где } a — \text{ускорение, с которым дви-}$$

жется система отсчета.

Основные результаты:

- исходя из принципа максимальной энтропии определена необходимость вероятностного описания физических систем (квантовой механики) (2006 г.) [24];
- из информационных предпосылок определены: закон всемирного тяготения (2009–2010 гг.) [25–27]; второй закон Ньютона (2009–2010 гг.) [25–27]; уравнения Фридмана (2010 г.) [28];
- показана необратимость времени (2009 г.) [29].

Подход 2. (Гуревич). Исходные положения:

- Закон сохранения энергии.
- Закон сохранения неопределенности (информации).
- Основной принцип квантовой механики Цайлингера: элементарные физические системы содержат (несут) 1 бит информации.

Основные результаты: для описания физических систем используются: информационная энтропия, информационная дивергенция, совместная информационная энтропия, информация связи.

При подходе 2 для оценки объема информации в физической системе не требуется использование голографического принципа. Методика оценки объема неопределенности (информации) в физических объектах иерархической структуры заключается в следующем. Сначала оценивается объем неопределенности (информации) в объектах нижнего уровня (лептонах и кварках). Согласно основному принципу квантовой механики Цайлингера считаем, что в объектах нижнего уровня — фундаментальных частицах — содержится 1 бит неопределенности (информации). Далее оценивается объем неопределенности (информации) в объектах второго уровня. Он равен сумме неопределенности (информации) объектов нижнего уровня плюс объем неопределенности (информации), заключенной в структуре объекта второго уровня иерархии (мезоны, барионы). Объем информации в структуре объекта второго уровня оценивается по волновой функции объекта второго уровня и по графу, отображающему структуру объекта второго уровня. Аналогично оценивается объем неопределенности (информации) в объектах следующих уровней. Даны прямые оценки объема информации в физических системах (2009 г.) [22].

- Показано, что объем информации, формирующейся в системе отсчета, движущейся с ускорением, равен $I(x') = -\ln J = -\ln \sqrt{-g'} = -\ln \sqrt{1 - \frac{2ax}{c^2}} \approx \frac{ax}{c^2}$ [22, второе издание].

Здесь и далее:

I — объем информации; J — якобиан; g' — определитель метрического тензора.

- Обратим внимание на аналогию с эффектом Унру [23]. Появление теплового излучения в ускоряющейся системе отсчета при отсутствии этого излучения в инерциальной системе отсчета есть появление дополнительной информации в ускоряющейся системе отсчета при отсутствии этой информации в инерциальной системе отсчета.

Из информационных предпосылок:

— определен вид гравитационного потенциала $\propto \frac{1}{r}$ (2008, 2009 гг.) [21, 22], где r — расстояние до гравитирующего объекта;

— определен вид напряженности гравитационного поля $\propto \frac{1}{r^2}$ (2008, 2009 гг.) [21, 22].

- Показана необратимость времени (1989 г.) [16].
- Показано, что законы информатики определяют действие физических законов сохранения (энергии, импульса, момента импульса).
- Открыто существование нескольких типов материи с разной зависимостью объема информации от массы (в том числе линейная для обычного вещества $I \propto M$, квадратичная для черных дыр $I \propto M^2$, линейно-логарифмическая для нейтронных дыр и белых карликов $I \propto M \log_2 M$) (2007—2009 гг.) [20—22].

Здесь M — масса объекта.

- Показано существование пятого типа взаимодействия — информационного взаимодействия.
- Разработаны информационные модели космологических объектов (черных дыр, нейтронных звезд, белых карликов, звезд солнечного типа) (2007—2009 гг.) [20—22].
- Выведена формула Хокинга для черных дыр (информационный спектр излучения) (2007 г.) [20].
- Выведена формула для информационного спектра излучения нейтронных звезд и белых карликов (2009 г.) [22].
- Разработана методика и даны оценки объема информации в космологических объектах (2007—2009 гг.) [20—22].
- Получены информационные ограничения на образование и слияние черных дыр (2009 г.) [22].
- Открыто существование и исследованы характеристики оптимальных черных дыр (минимизи-

рующих объем информации в области Вселенной, Вселенной в целом) (2007 г.) [20].

- Дана оценка массы начальных неоднородностей Вселенной (2009 г.) [22].
- Показано, что расширение Вселенной является причиной и источником формирования информации. Разнообразные физические процессы в расширяющейся Вселенной обеспечивают формирование информации (1989—2009 гг.) [16, 19—22].
- Даны оценки максимально и минимально возможного, текущего объемов информации во Вселенной (1989 г.) [16], оценки основных информационных характеристик Вселенной (1989—2009 гг.) [16, 19—22].
- Из информационных предпосылок определена необходимость описания физических систем (квантовой механики) (2009 г.) неклассической вероятностной логикой [22].
- Показано, что во всех возможных Вселенных действуют законы информатики и тем самым физические законы сохранения.

Заключение

Работы автора и зарубежных ученых (американских, канадских, европейских, китайских ...) *подтверждают первичность информационных законов.*

- Информационные законы (законы информатики) определяют и ограничивают физические законы.
- Информационные законы (законы информатики) имеют всеобщий, универсальный характер, действуют во всех возможных Вселенных, даже во Вселенных с разными физическими законами.
- Приведенные сведения показывают, что приоритет получения информационными методами физических результатов принадлежит России, хотя последние результаты зарубежных ученых очень интересны и важны.
- Использование информационных законов (законов информатики) совместно с физическими законами позволит раскрыть все тайны природы, в частности, построить теорию квантовой гравитации.

Список литературы

- Wheeler J. Geons, Black Holes & Quantum Foam: A Life in Physics. New York; W. W. Norton & Company, 1998. P. 63—64.
- Стин Э. Квантовые вычисления. М.-Ижевск: Научно-издательский центр "Регулярная и хаотическая динамика". 2000. 112 с.
- Каломцев Б. Б. Динамика и информация / М.: Изд. редакции журнала "Успехи физических наук". 1999. 396 с.
- Энштейн А. К электродинамике движущихся тел. Собр. науч. тр. Т. 1. М.: Наука. 1965. С. 7—35. 5. Szillard L. Physik. 1929. V. 53. P. 840.
- Нейман Д. Математические основы квантовой механики. М.: Наука, 1964. 366 с.

7. Шеннон К. Математическая теория связи. Работы по теории информации и кибернетики. М.: Изд-во иностр. лит. 1963. С. 243—332.
8. **Janes E.** Information theory and statistical mechanics. I. Phys. Rev. 1957. V. 106. N 4. P. 620—630. II. Phys. Rev. 1957. V. 108. N 2. P. 171—190.
9. **Бриллюэн Л.** Наука и теория информации. М.: Гос. изд-во физ.-мат. лит. 1960. 392 с.
10. **Everett H.** The theory of the universal wave function / De. Witt and N. Graham, editors // The many world interpretation of quantum mechanics. Princeton: Princeton University Press. 1973.
11. **Пенроуз Р.** Новый ум короля. М.: УРСС, 2003. 384 с.
12. **Хокинг С., Пенроуз Р.** Природа пространства и времени. Ижевск: НИЦ "Регулярная и хаотическая динамика". 2000. 160 с.
13. **Marzolis N., Levitin L.** Phys. Сотр. 96 / T. Toffoli, M. Bifare, J. Leao, eds. Boston: NECSI. 1996.
14. **Zeilinger A.** A Foundational Principle for Quantum Mechanics // Foundations of Physics. 1999. Vol. 29. N 4. P. 631—643.
15. **Lloyd S.** Computational capacity of the universe. arXiv:quant-ph/0110141 v1 24 Oct 2001.
16. **Гуревич И. М.** Законы информатики — основа исследований и проектирования сложных систем связи и управления: Метод. пособие. М.: ЦООНТИ "Экос". 1989. 60 с.
17. **Hsul S. and Zee A.** Message in the Sky. arXiv:physics/0510102 v2 6 Dec 2005.
18. **Scott D. and Zibin J.** The Real Message in the Sky. arXiv:physics/0511135 v1 15 Nov 2005.
19. **Гуревич И. М.** Законы информатики — основа строения и познания сложных систем. М.: Антиква. 2003. 176 с.
20. **Гуревич И. М.** Законы информатики — основа строения и познания сложных систем". Изд. второе, уточненное и доп. М.: Торус Пресс. 2007. 400 с.
21. **Гуревич И. М.** Оценка основных информационных характеристик Вселенной // Информационные технологии. Приложение. 2008. № 12. 32 с.
22. **Гуревич И. М.** Информационные характеристики физических систем. М.: 11-й ФОРМАТ. Севастополь "Кипарис". 2009. 170 с.
23. **Эффект Унру.** URL: http://ru.wikipedia.org/wiki/%D0%AD%D1%84%D1%84%D0%B5%D0%BA%D1%82_%D0%A3%D0%BD%D1%80%D1%83.
24. **Garrett A.** Quantum mechanics from a universal action reservoir. 0605068v1 [physics.pop-ph] 8 May 2006. February 2, 2008.
25. **Verlinde E.** On the Origin of Gravity and the Laws of Newton. ar-Xiv:1001.0785v1 [hep-th] 6 Jan 2010. Institute for Theoretical Physics University of Amsterdam / Valckenierstraat 65. 1018 XE, Amsterdam. The Netherlands.
26. **Smolin L.** Newtonian gravity in loop quantum gravity. 1001.3668v1 [gr-qc] 20 Jan 2010. Perimeter Institute for Theoretical Physics, 31 Caroline Street North, Waterloo, Ontario N2J 2Y5, Canada. January 21, 2010.
27. **Makela J.** Notes Concerning "On the Origin of Gravity and the Laws of Newton" by E. Verlinde. 1001.3808v1 [gr-qc] 21 Jan 2010. Vaasa University of Applied Sciences, Wolffmtie 30, 65200 Vaasa, Finland.
28. **Rong-Gen Caia, Li-Ming Caob, and Nobuyoshi Ohta.** Friedmann Equations from Entropic Force. arXiv:1001.3470v1 [hep-th] 20 Jan 2010. Key Laboratory of Frontiers in Theoretical Physics, Institute of Theoretical Physics, Chinese Academy of Sciences, P. O. Box 2735, Beijing 100190, China Department of Physics, Kinki University, Higashi-Osaka, Osaka 577-8502, Japan.
29. **Maccione L.** A quantum solution to the arrow-of-time dilemma. ar-Xiv:0802.0438v3 [quant-ph] 25 Aug 2009. QUIT, Dip. A. Volta, 27100 Pavia, and Institute for Scientific Interchange, 10133 Torino, Italy.

БЕЗОПАСНОСТЬ ИНФОРМАЦИИ И УПРАВЛЕНИЕ РИСКАМИ

УДК 004.056

Р. М. Алгулиев¹, д-р техн. наук, чл.-корр. НАНА, директор,
Я. Н. Имамвердиев², канд. техн. наук, зав. отделом,
С. А. Деракшанде³, аспирант,

Институт информационных технологий национальной академии наук Азербайджана, г. Баку
 e-mail: ¹director@iit.ab.az; ²yadigar@lan.ab.az; ³smdk364@yahoo.com

Пути повышения точности методов оценки рисков информационной безопасности

Приведен сравнительный анализ существующих подходов к оценке рисков информационной безопасности (ИБ). Традиционные подходы к оценке рисков ИБ базируются в основном на качественных оценках, которые являются в высокой степени субъективными и не обладают необходимой достоверностью. Для разработки адекватных моделей оценки рисков необходимо учесть природу неопределенности.

Анализируются вероятностный, нечетко-множественный и экспертный подходы к моделированию различных типов неопределенностей в контексте ИБ. Структурированы основные преимущества и недостатки методов оценки рисков ИБ на основе этих подходов и намечены важные направления исследований в этой области.

Ключевые слова: информационная безопасность, угроза, уязвимость, оценка риска, нечеткая логика

Введение

В современных условиях широкого применения информационных технологий риски информационной безопасности (ИБ) являются неотъемлемым фактором деятельности любой организации. Информационные технологии не только позволяют повысить эффективность бизнес-процессов, но и могут стать источником огромного ущерба, поэтому управление рисками ИБ имеет решающее значение для поддержания бизнес-процессов организации и как один из видов операционных рисков становится все более значимым разделом общей системы управления рисками [1].

Для управления рисками ИБ на уровне предприятий разработаны несколько стандартизованных методологий, например международные стандарты ISO/IEC 13335-3, ISO/IEC 27005, а также национальные стандарты NIST SP 800-30, BSI, COBIT, SAS 55/78 и некоторые другие [1].

Сегодня признанным международным стандартом по управлению рисками ИБ является стандарт ISO/IEC 27005 [2]. Этот стандарт носит концептуальный характер и описывает четырехфазную процессную модель управления рисками ИБ. В стандарте не навязываются конкретные методы и методики анализа и оценивания риска, допускаются качественные и количественные методики. Исторически сложилось так, что большинство существующих методик оценки рисков базируются не на стандартах управления рисками ИБ, а на стандарте управления информационной безопасностью ISO/IEC 17799 (ныне ISO/IEC 27002). На основе этого стандарта были разработаны и реализованы разные методики и около 20 инструментальных средств оценки и управления рисками ИБ [1–7]. В настоящее время самыми популярными методиками управления рисками являются методики OCTAVE (*Operationally Critical Threat, Asset, and Vulnerability Evaluation*) [4] и CRAMM (*Commercial Risk Analysis and Management Methodology*) [5]. Наиболее известными российскими программными продуктами в области анализа рисков ИБ являются программные продукты "Гриф" компании Digital Security (Санкт-Петербург) [6] и "Авангард" Института системного анализа РАН [1].

Следует отметить, что исследователи провели большую работу по классификации, систематизации рисков, нахождению методов их оценки и анализа [8]. Вместе с тем, существует ряд проблем, требующих дальнейшей разработки. До настоящего времени нет единого подхода к определению факторов, определяющих риски ИБ, и общепринятых методов оценки рисков. Оценка рисков ИБ базируется в основном на экспертных оценках, которые являются в высокой степени субъективными и по этой причине не обладают необходи-

мой достоверностью. Величины рисков, вычисленные разными методами, плохо коррелируют между собой, порой они различаются на несколько порядков [8–9].

Исследования показывают, что практические подходы, используемые при анализе рисков ИБ, часто не имеют хорошо проработанного формального обоснования и соответствующего математического аппарата. Кроме того, при разработке новых методик анализа рисков ИБ и модификации существующих опираются не на фундаментальные результаты в предметной области, а на особенности практических задач. С возрастанием роли ИБ для жизни и деятельности людей на этапе перехода к информационному обществу особенно актуальной становится разработка точных метрик для оценки рисков ИБ [10]. Актуальность данной проблемы растет с усилением факторов неопределенности внешней среды, серьезно усложняющих процесс управления рисками и увеличивающих вероятность принятия необоснованных и недостоверных управленческих решений. Исходя из этого в данной работе анализируются основные недостатки существующих подходов к оценке рисков ИБ и проблемы применения современных математических подходов для повышения точности оценок рисков ИБ.

Качественный и количественный подходы к оценке риска

В существующих методиках риск ИБ оценивается по двум (угроза, потенциальное воздействие) или по трем (угроза, уязвимость, потенциальное воздействие) факторам [1]. Эти факторы можно измерить качественно или количественно, поэтому большинство методов оценки рисков ИБ основаны на качественном и количественном подходах.

Качественные оценки используются в случаях, когда количественные оценки по ряду причин затруднены [1]. Как правило, качественные показатели оцениваются с помощью балльных шкал. При качественном подходе сначала определяются шкалы измерения для факторов риска. Кроме того, для уровней риска устанавливается шкала и определяется таблица для расчета рисков по факторам риска. Приведем пример, использующий трех- и пятибалльные оценочные шкалы.

Уровни вероятности угрозы оцениваются по трехбалльной шкале: "низкий", "средний", "высокий".

Уровни ущерба оцениваются по пятибалльной шкале: "незначительный", "низкий", "средний", "высокий", "очень высокий".

Для оценки рисков ИБ устанавливается шкала из трех значений: "низкий риск", "средний риск", "высокий риск".

Определение рисков в зависимости от двух факторов

Ущерб	Вероятность		
	Низкая	Средняя	Высокая
Незначительный	Низкий риск	Низкий риск	Средний риск
Низкий	Низкий риск	Высокий риск	Высокий риск
Средний	Низкий риск	Высокий риск	Высокий риск
Высокий	Средний риск	Высокий риск	Высокий риск
Очень высокий	Средний риск	Высокий риск	Высокий риск

По заданной таблице рассчитываются уровни рисков.

Подобный подход к оценке рисков ИБ достаточно распространен. Недостатком этого подхода является то, что полученные оценки всегда зависят от субъективных суждений эксперта, его знаний и опыта. Результаты оценки невозможно интерпретировать однозначно, относительные оценки, полученные в ходе качественного оценивания, могут не устраивать тех, кто принимает решения. Действительно, как видно из таблицы, редко происходящие события ИБ с большим ущербом и часто происходящие события ИБ с маленьким ущербом представляют одинаковый риск, однако часто они требуют различных подходов в определении защитных мер.

Первая количественная методика для анализа рисков информационной безопасности была предложена в середине 1970-х годов [11]. В этой методике риск представляет собой произведение трех параметров:

- стоимости актива (*Asset Value, AV*);
- меры уязвимости актива к угрозе (*Exposure Factor, EF*);
- вероятности реализации угрозы за определенный период времени (*Annual Rate of Occurrence, ARO*).

На основе переменных *AV* и *EF* оценивается ожидаемый ущерб от единичной реализации угрозы (*Single Loss Exposure, SLE*):

$$SLE = AV \times EF.$$

Ожидаемый ущерб от угрозы или риск за определенный период времени (*Annual Loss Exposure, ALE*) рассчитывается по формуле

$$ALE = SLE \times ARO$$

или

$$ALE = (AV \times EF) \times ARO.$$

Преимуществами количественного подхода являются:

- научное обоснование оценок;
- возможность предоставления результатов оценки в терминах управления (финансовые показатели и вероятности в процентах);
- увеличение точности оценок с накоплением статистических данных.

К недостаткам количественного подхода, в первую очередь, относится отсутствие или сложность эффективных формальных методов, позволяющих точно определить основные числовые показатели.

Как отмечается в работе [12], сообщество информационной безопасности широко применяет парадигму вычисления *ALE*, "но существует только одна проблема с *ALE*: ... числа настолько "убоги", что не могут даже лгать". Автор приводит три основные причины этого:

- врожденная трудность моделирования выбросов;
- отсутствие статистики для оценки вероятностей возникновения угроз или потенциальных ущербов;
- чувствительность модели *ALE* к малым изменениям в предположениях.

Качественная и количественная оценки рисков могут использоваться по отдельности или вместе при необходимости. Методы анализа рисков нацелены в конечном итоге на лиц, принимающих решения, и количественные оценки, обеспечивая научную поддержку в процессе принятия решения, имеют огромное значение. Без достоверных количественных характеристик рисков ИБ нельзя успешно сопоставлять и однозначно интерпретировать полученные результаты, выбрать оптимальные защитные меры по критерию стоимость—эффективность, снизить влияние субъективных факторов при принятии решений. Для разработки адекватных моделей оценки рисков ИБ необходим, прежде всего, глубокий анализ природы этих рисков.

Риск и неопределенность

Существуют разные взгляды на природу рисков. Анализ литературы по управлению рисками в разных областях деятельности показывает, что в настоящее время существует множество определений понятия "риск" [13, 14]. Большинство определений связывают существование рисков с наличием неопределенности. Неопределенность выступает необходимым и достаточным условием риска в принятии решений. Именно необходимость принимать решения в условиях неопределенности порождает риск; если нет необходимости принимать решение, то и риска нет.

Отождествление понятий "неопределенность" и "риск" совершенно недопустимо с теоретической и практической точек зрения. Ф. Найт в 1920-х годах первым дал четкое разграничение понятий "неопределенность" и "риск" [15]. Согласно концепции Найта риск — измеримая неопределенность, объективная вероятность наступления события, он может быть оценен и просчитан. Неопределенность же означает, что вероятность наступления этого события нельзя измерить или выразить количественно.

Решения, связанные с вопросами ИБ, часто принимаются в условиях неопределенности состояний внешней и внутренней среды, неопределенности целей, источников и последствий угроз, неопределенности защитных мер по ИБ и т. д. Поэтому изучение явления "неопределенность" в контексте ИБ, источников его возникновения, структуры, последствий, а также классификация факторов, влияющих на неопределенность, представляют большую актуальность для оценки рисков ИБ.

Среди различных подходов моделирования неопределенности можно выделить три основных подхода: вероятностный, нечетко-множественный и экспертный.

Как отмечают разные исследователи [16—18], эффективность применения подходов на основе вероятностных, нечетко-множественных и экспертных описаний к решению различных задач зависит от уровня и характера неопределенности, связанной с конкретной задачей. По мере увеличения неопределенности классические вероятностные описания уступают место, с одной стороны, субъективным вероятностям, основанным на экспертной оценке, а с другой стороны — нечетким описаниям, выраженным в виде функций принадлежности [16].

Вероятностная оценка риска

Исследование рисков тесно связано с развитием теории вероятностей. Теоретические аспекты экономического риска стали объектом научного анализа с XVIII века. Появление и развитие во второй половине XX века новых технологий в области обработки информации, расширение финансовых рынков и активное развитие финансовых инструментов, усиление конкуренции, начало глобализации послужили предпосылками к возникновению новой науки об управлении рисками. В настоящее время в области оценки техногенного и экономического рисков накоплен огромный научный опыт и знания [19—23]. Использование научных подходов, сформированных в этих областях для оценки рисков ИБ, сопровождается порой неопределимыми проблемами. Функционирование информационной системы характеризуется большим разнообразием типов неопределенности, активностью и целями внешней и внутренней среды. Поэтому вопросы оценки рисков различаются для экономических и информационных систем. Несмотря на это, предпринимаются попытки использования идей и подходов оценки экономических рисков. Организация Computing Research Association в 2003 г. определила пять самых актуальных проблем в области надежных вычислений на ближайшие 10 лет. Третьим вызовом в этом списке указана проблема количественной оценки рисков ИБ [24].

Авторы надеются, что исследователям удастся разработать за ближайшие 10 лет эффективные методы оценки рисков ИБ на основе идей и подходов количественной оценки финансовых рисков.

Существует множество методов оценки финансовых рисков на основе вероятностного подхода [25]. Среди них можно выделить метод *VaR* (*Value-at-Risk*), он появился сравнительно недавно (в 1990-х годах) и первоначально использовался только в финансовой сфере для расчета различных видов рисков: ценового риска, валютного риска, кредитного риска, риска ликвидности. *VaR* определяет предельный уровень, который с заданной вероятностью не превысит возможные потери. В последнее время некоторые исследователи по ИБ предложили использовать метод *VaR* для оценки рисков ИБ [26].

Но несмотря на привлекательность и достоинства вероятностного подхода, данный подход встречает серьезные препятствия при практическом применении, одним из них является высокая чувствительность получаемого результата к законам распределения вероятностей. Кроме того, оценить их достоверность в практическом исследовании обычно не представляется возможным, так как получение качественной статистической информации не представляется возможным по самым различным причинам [27].

В вероятностном подходе неопределенность независима от ее природы отождествляется со случайностью. Но как отмечают некоторые авторы, не все угрозы ИБ носят вероятностный характер [27]. Поэтому понятие риска как функции от вероятности угрозы требует переосмысления.

Кроме того, точность оценки вероятностей (объективных и субъективных) зависит от множества факторов, в том числе от качества статистической информации или качества экспертных оценок. Это послужило росту недоверия к получаемым на их основе оценкам и решениям.

Вероятностный подход предполагает выполнение вероятностных предпосылок, в частности, повторяемости опытов. В случае ИБ повторяемость может трактоваться только как возможность осуществления многих однородных угроз на длительном временном интервале при неизменных условиях, но вряд ли это возможно в реальной практике.

Вследствие вышеописанных причин точность оценок рисков, полученных по вероятностным методам, в значительной степени зависит от качества исходных предположений, что может привести к значительным ошибкам в полученных результатах (например, переоценке или недооценке риска ИБ).

Очевидно, если факторы риска ИБ характеризуются репрезентативной статистикой или имеются достаточные основания полагать, что исходные параметры подчиняются определенному вероятност-

стному закону, то в данной ситуации применение вероятностного подхода вполне оправдано и эффективно. Однако, как правило, при моделировании ИБ статистика либо недостаточно репрезентативна, либо отсутствует вовсе, тогда применение вероятностного подхода затруднительно или невозможно вовсе.

Сети Байеса и оценка риска

Как уже отмечалось, одной из основных проблем при оценке рисков ИБ является отсутствие статистических данных. Стандартным подходом в этом случае является использование экспертного мнения для дополнения данных. Далее с появлением новых данных можно обновлять, модернизировать или калибровать модель, а также возможно полное замещение экспертного мнения. Привлекательным и перспективным в этом направлении видится использование сетей Байеса [28]. Сети Байеса являются развитием вероятностного подхода на основе теоремы Байеса.

Сеть Байеса — это графическая модель для представления вероятностных взаимоотношений на множестве переменных. По структуре сеть является ориентированным графом, в котором каждая вершина имеет некоторые значения вероятностей. Использование сетей Байеса дает ряд преимуществ для оценки рисков ИБ. Она легко справляется с ситуациями, когда часть данных отсутствует. Также сеть Байеса является идеальным средством для совместного представления экспертных знаний и статистических данных.

Возможности сетей Байеса работать с вероятностями, данными разной природы (экспертные оценки, числовые данные об объеме нанесенного ущерба) и моделировать причинно-следственные зависимости между событиями предоставляют подходящий формализм для использования их при анализе рисков. Анализ литературы по применениям сетей Байеса показывает, что начиная с 2000 г., сети Байеса успешно применяются в различных сферах деятельности для оценки рисков [29—31].

В последние годы делаются попытки применить сети Байеса для оценки рисков ИБ [31]. Существующие работы сосредоточивают свое внимание на измерении воздействий отдельных уязвимостей взамен измерения их совместного воздействия. Они не приспособлены для применения к более реальным случаям, когда использование одной уязвимости делает эксплуатацию другой более легкой. Для решения проблемы учета взаимной зависимости уязвимостей можно применить графы атак [32], описывающие угрозы, нацеленные на бизнес-процессы, реализованные информационной инфраструктурой.

Нечеткий подход к оценке риска

Теория нечетких множеств, основные идеи которой были предложены Лютфи Заде (Lotfi Zadeh) в начале 1960-х годов [33], позволяет описывать качественные, неточные понятия и знания об окружающем мире, а также оперировать этими знаниями.

В последнее время нечеткое моделирование является одним из наиболее активных и перспективных направлений исследований в области управления и принятия решений. Нечеткое моделирование успешно применяется для управления рисками в техногенных системах [34—36], в финансовом секторе [18, 37, 38].

Нечеткий подход позволяет совмещать в методах оценивания рисков качественные аспекты, не имеющие точной числовой оценки, и количественные признаки, что резко повышает уровень применяемых методов. Ниже перечислены некоторые преимущества нечеткого подхода к оценке рисков ИБ по сравнению с вышеперечисленными методами:

- нечеткий подход позволяет формально описать и учесть различные виды неопределенности;
- нечеткий подход характеризуется простотой выявления экспертных знаний;
- нечеткий подход позволяет формализовать в единой форме и использовать неоднородную информацию (детерминированную, интервальную, статистическую, лингвистическую) [17];
- вычисление оценок риска на основе нечеткого метода оказывается эффективным в ситуациях, когда вероятностные оценки не могут быть получены;
- результат, получаемый на основе нечеткого подхода, характеризуется низкой чувствительностью к изменению вида функций принадлежности, поэтому нечеткий подход не требует абсолютно точного задания функций принадлежности.

Анализ литературы показывает, что имеются попытки применить нечеткое моделирование также для оценки рисков ИБ [17]. К сожалению, в большинстве таких работ неоднократно высказывается возможность использования нечеткого подхода для оценки рисков ИБ и приводятся простые примеры для иллюстраций. Приходится признать, что отсутствуют работы в области оценки рисков ИБ, использующие разнообразные инструменты и достижения нечеткого моделирования.

Заключение

Проведенный анализ традиционных методов оценки рисков ИБ свидетельствует об их ограниченной практической применимости вследствие относительности и большой субъективности оценок. Сравнительный анализ традиционных мето-

дов оценки рисков ИБ и нечеткого подхода показывает, что теория нечетких множеств является одной из наиболее эффективных математических теорий, направленных на формализацию и обработку разнородной неопределенной информации, характерной для задач анализа рисков ИБ. Более привлекательными и перспективными видятся гибридные подходы, комбинирующие преимущества вероятностного и нечеткого подходов для оценки рисков ИБ. Такой подход позволил бы использовать различные по природе данные (экспертные оценки, числовые данные, качественные оценки), снизить субъективность данных и был бы пригодным для повышения достоверности оценок риска.

Список литературы

1. Петренко С. А., Симонов С. В. Управление информационными рисками. Экономически оправданная безопасность. М.: ДМК Пресс, 2004. 384 с.
2. ISO/IEC 27005:2002. Information Technology — Security Techniques — Information Security Risk Management. 2007.
3. Лопарев С., Шелупанов А. Анализ инструментальных средств оценки рисков утечки информации в компьютерной сети предприятия // Вопросы защиты информации. 2003. № 4.
4. Alberts C. J., Behrens S. G. et al. Operationally Critical Threat, Asset and Vulnerability Evaluation (OCTAVE) Framework. Pittsburgh: Carnegie Mellon, 1999. P. 1—69.
5. Barber B., Davey J. The Use of the CCTA Risk Analysis and Management Methodology CRAMM // Proc. of the 7-th World Congress on Medical Informatics. North Holland. 1992. P. 1589—1593.
6. Куканова Н. Современные методы и средства анализа и управления рисками информационных систем компаний. URL: http://www.dsec.ru/about/articles/ar_compare/.
7. Астахов А. М. Искусство управления информационными рисками. М.: ДМК Пресс, 2010. 312 с.
8. Hamdi M., Boudrigha N. Computer and Network Security Risk Management: Theory, Challenges, and Countermeasures // International Journal of Communication Systems. 2005. Vol. 18. N 8. P. 763—793.
9. Марков А. С., Цирлов В. Л. Управление рисками — нормативный вакуум // Открытые системы. 2007. № 7. URL: <http://www.osp.ru/os/2007/08/44928737/>
10. McHugh J. Quality of protection: Measuring the unmeasurable? // Proceedings of the 2nd ACM Workshop on Quality of Protection (QoP'06). 2006. P. 1—2.
11. Guidelines for Automated Data Processing Physical Security and Risk Management, NIST FIPS PUB-65. 1974.
12. Jaquith A. Security Metrics: Replacing Fear, Uncertainty, and Doubt. NJ: Addison-Wesley Pearson Education, 2007. P. 336.
13. Vlek C., Stallen J. Judging Risk and Benefits in the Small and in the Large Organization // Organization Behavior and Human Performance. 1981. Vol. 28. Issue. 2. P. 235—271.
14. Рогов М. А. Риск-менеджмент. М.: Финансы и статистика, 2001. 120 с.
15. Найт Ф. Х. Риск, неопределенность и прибыль. М.: Дело, 2003. 358 с.
16. Борисов А. Н., Алексеев А. В., Меркурьева Г. В. и др. Обработка нечеткой информации в системах принятия решений. М.: Радио и связь, 1989. 304 с.
17. Деревянко П. М. Сравнение нечеткого и имитационного подхода к моделированию деятельности предприятия в условиях неопределенности // Современные проблемы экономики и управления народным хозяйством: Сб. науч. статей. СПб.: СПбГИЭУ. 2005. С. 289—292.
18. Царев В. В. Оценка экономической эффективности инвестиций. СПб.: Питер, 2004. 464 с.
19. Шапкин А. С., Шапкин В. А. Теория риска и моделирование рискованных ситуаций: Учебник. М.: Дашков и Ко, 2007. 880 с.
20. Владимиров В. А., Измалков В. И., Измалков А. В. Оценка риска и управление техногенной безопасностью. М.: Деловой экспресс, 2002. 184 с.
21. Рябинин И. А. Надежность и безопасность структурно-сложных систем. СПб.: 2-е изд. Изд. Санкт-Петербургского университета, 2007. 278 с.
22. Быков А. А., Мурзин Н. В. Проблемы анализа безопасности человека, общества и природы. СПб.: Наука, 1997. 247 с.
23. Быков А. А., Акимов В. А., Фалеев М. И. Нормативно-экономические модели управления риском // Проблемы анализа риска. 2004. Т. 1. № 2. С. 125—137.
24. Computing Research Association Four Grand Challenges in Trustworthy Computing URL: <http://www.cra.org/reports/trustworthy.computing.pdf>
25. Бартон Т., Шенкир У., Уокер П. Комплексный подход к риск-менеджменту. М.: Издат. дом "Вильямс". 2003. 208 с.
26. Jaisingh, J. and Rees. J. Value at Risk: A Methodology for Information Security Risk Assessment // In Proc. of the INFORMS Conference on Information Systems and Technology. Miami, 2001. P. 101—106.
27. Нечунаев В. М. Оценка рисков информационной безопасности // Докл. ТУСУРа. —2009. № 1. С. 51—53.
28. Neil M., Fenton N., Tailor M. Using Bayesian Networks to Model Expected and Unexpected Operational Losses // Risk Analysis: An International Journal. 2005. Vol. 25, N 4. P. 963—972.
29. Medina-Oliva G., Weber P., Simon C., Iung B. Bayesian Networks Applications on Dependability, Risk Analysis and Maintenance // 2-nd IFAC Workshop on Dependable Control of Discrete System. Italy, 2009. P. 245—250.
30. Ram P., Prakash K. Risk Management Using Behavior Based Bayesian Networks // IEEE international Conference on Intelligence and Security Informatics. Lecture Notes in Computer Science. 2005. Vol. 3495. P. 115—126.
31. Frigault M., Wang L., Singhal A., Jajodia S. Measuring Network Security Using Dynamic Bayesian Network // Proc. of the 4-th ACM Workshop on Quality of Protection (QoP). Alexandria, 2008. P. 23—30.
32. Sheyner O., Jha S., Wing J., Lippmann R., Haines J. Automated Generation and Analysis of Attack Graphs // Proc. IEEE Symposium on Security and Privacy. 2002. P. 273—284.
33. Zadeh L. Fuzzy Sets / Information and Control June 1965. N 8 (3). P. 338—353.
34. Chen Shi-Jay, Chen Shyi-Ming. Fuzzy Risk Analysis Based on Similarity Measures of Generalized Fuzzy Numbers // IEEE Trans. on Fuzzy Systems. 2003. Vol. 11. N 1. P. 45—56.
35. Karimi I., Hüllermeier E. Risk Assessment System of Natural Hazards: A New Approach Based on Fuzzy Probability // Fuzzy Sets and Systems. 2007. N 158 (9). P. 987—999.
36. Dikmen I., Birgonul M. T., Han S. Using Fuzzy Risk Assessment to Rate Cost Overrun Risk in International Construction Projects // International Journal of Project Management. 2007. N 25 (5). P. 494—505.
37. Недосекин А. О. Нечеткий финансовый менеджмент М.: Аудит и финансовый анализ, 2003. 184 с.
38. Недосекин А. О. Нечетко-множественный анализ риска фондовых инвестиций. СПб.: Сизм, 2002. 181 с.
39. Alguliev R. M., Imamverdiev Y. N. About One Method of Risk Measurement of Maintenance of Information Security of Corporate Networks Because of Fuzzy Sets // Proc. of the 4-th International Conference on New Information Technologies. Minsk, 2000. P. 76—81.
40. Авдошин А. С. Оценка защищенности информационной системы методами нечеткой логики // Вестн. Самарского гос. техн. ун-та. 2005. N 32. С. 191—193.
41. Балашов П. А., Кислов Р. И., Безгузиков В. П. Оценка рисков информационной безопасности на основе нечеткой логики // Безопасность компьютерных систем. Конфидент. 2003. № 5. С. 56—59.
42. Доценко С. М., Зайчиков А. А., Малыш В. Н. Повышение объективности исходных данных как альтернатива методу нечеткой логики при оценке риска информационной безопасности // Защита информации. Конфидент. 2004. N 5. С. 83—85.
43. Кашенко А. Г., Остапенко А. Г., Кашенко Г. А. Применение нечеткой логики к оценке рисков информационной безопасности автоматизированных систем // Междунар. научно-техн. конф. "Кибернетика и высокие технологии XXI века". Воронеж: НПФ "Саквеев". 2005. С. 249—254.

Д. Н. Генералов, аспирант,
Поволжский государственный
университет сервиса, г. Тольятти
e-mail: kaf_pmii@tolgas.ru

Противодействие нарушению целостности модулей специализированного программного обеспечения бортового устройства

Рассматривается порядок основных операций выбора системы противодействия нарушению целостности модулей специализированного программного обеспечения бортового устройства. Приводится разработанный алгоритм защиты от несанкционированного доступа в момент запуска, завершения сеанса работы, при возможной ликвидации защиты специализированного программного обеспечения на основе распределения данных по контентам.

Ключевые слова: бортовое устройство, специализированное программное обеспечение, защита информации, нарушение целостности данных

Введение

При разработке, модернизации, эксплуатации специализированного программного обеспечения бортовых устройств возникают вопросы, связанные с обеспечением целостности данных и защитой от несанкционированного изменения промежуточных информационных баз данных. Промежуточной (буферной) информационной базой является база данных для временного хранения необработанной информации, поступающей от преобразователей, промежуточных значений баз данных систем бортовых устройств, а также открытых данных о доступе пользователя к ресурсам специализированного программного обеспечения. Под специализированным программным обеспечением (СПО) в статье понимается совокупность модулей программного обеспечения бортовых устройств, обеспечивающих обработку или передачу данных. Функциями бортовых устройств (БУ) являются сбор, управление и преобразование информационных сигналов для организации хранения, передачи, преобразования, обработки, контроля целостности и защиты информации от действий внутреннего нарушителя [1]. Внутренним нарушителем (ВН) является субъект или группа субъектов, обладающих правом доступа на объект и к конфиденциальной информации в силу выполнения служебных или иных обязанностей. Угрозой нарушения

целостности данных является нарушение состояния информации, при котором ее изменение осуществляется только преднамеренно субъектами, имеющими на него право [2].

Для стабильного функционирования СПО БУ необходимо определить механизмы контроля целостности и защиты данных от несанкционированных воздействий ВН. В процессе функционирования БУ информация первичных, вторичных преобразователей может несанкционированно изменяться ВН [3]. Это способствует возникновению ошибочных действий, возможных отказов БУ, нарушению целостности данных отдельных модулей СПО.

Актуальность настоящей работы обусловлена необходимостью выбора подходящего варианта защиты, алгоритма обеспечения целостности данных и модулей СПО БУ. Необходимо выбрать вариант системы противодействия нарушению целостности модулей СПО БУ, который позволяет предотвратить несанкционированное изменение буферных баз данных, снизить критичность и вероятность реализации угрозы, затруднить использование имеющихся уязвимостей СПО для осуществления угроз. В рамках рассмотрения различных известных систем защиты [1, 2, 4, 6] СПО необходимо решить задачу выбора подходящего варианта системы защиты данных от несанкционированного доступа для разных стадий (при возможной ликвидации защиты, в момент запуска и завершения сеанса) функционирования СПО.

Выбор варианта защиты от атак внутреннего нарушителя

Для предотвращения несанкционированного доступа к СПО БУ на базе бортового компьютера типа "ШТАТ", контроллера типа BOSH M. 1.5.4 автомобиля семейства ВАЗ-2110 был образован порядок основных операций выбора и разработки подходящего варианта защиты для модулей СПО. Данный порядок операций был использован при разработке авторской программы Auto Security Mobile v.1.0 для защиты СПО БУ. Разработанная программа защиты и противодействия нарушению целостности модулей СПО может быть интегрирована совместно с серийными СПО БУ.

Основные операции разработки и выбора подходящего варианта защиты от атак ВН для СПО БУ приведены на рис. 1 в виде схемы. Представленная схема включает несколько стадий определения выбора подходящего варианта для системы защиты. Возможные атаки ВН можно представить в виде одномерного массива $A = \{a_1, \dots, a_j, \dots, a_n\}$, при этом элементы

$$a_j = I^S(B_I, S_K, \lambda_C, \beta_R, C_{SM}), \quad (1)$$



Рис. 1. Операции выбора и разработки подходящего варианта защиты от атак ВН

где I^S — функциональная зависимость возможных видов атак a_j (постоянная, единовременная, периодическая, единичная, множественная атака) от параметров функционирования БУ. Данные параметры можно представить в виде множества $B_I = \{b_1, \dots, b_j, \dots, b_m\}$, где b_1 — число режимов работы, b_2 — число задействованных модулей, b_3 — параметры быстрогодействия обработки информации, и множества $S_K = \{s_1, \dots, s_j, \dots, s_n\}$, где s_1 — число буферных баз данных, s_2 — параметры сопоставления с базой данных сигнатур, s_3 — параметры реализации угрозы и пр. В равенстве (1) учитываются параметры времени жизни (актуальности) λ_C оперируемой информации и число β_R задействованных модулей СПО при обработке результатов опроса первичных преобразователей информационных сигналов. Информационные сигналы вторичных преобразователей и исполнительных устройств автомобиля формируются в виде матрицы C_{SM} для топливных форсунок двигателя внутреннего сгорания, форсунок для жидкостного смыва загрязнений с плафонов головных фар, соленоидов управления механизмом запираания дверей, моторедуктора стеклоподъемников, исполнительных реле [5].

Для соответствующего варианта защиты данных БУ можно записать систему уравнений, характеризующую цель информационно-программного воздействия ВН на отдельные модули СПО:

$$\begin{cases} B_I := CLSASS(b_i); \\ CND := DIFF(S_K, \beta_R, \lambda_C, b_i) \cup COND(b_i); \\ SCR := SCRIPT(S_K, \beta_R, \lambda_C, CND, ACT). \end{cases} \quad (2)$$

Содержание системы (2) включает классификацию $CLASS$ цели атаки ВН, направленной на соответствующие аппаратные модули b_i БУ. Необходимо отметить, что единичная атака CND и возможная серия атак ACT в пределах функционирования одной сессии СПО зависит от формирования

$DIFF$ информационных процессов, целостности S_K программных модулей, актуальности λ_C информации, числа β_R активных модулей СПО в соответствующих аппаратных частях $COUND(b_i)$ БУ. При локализации несанкционированного воздействия используются готовые сценарии SCR противодействия, задекларированные в политике безопасности $SCRIPT$ СПО БУ в системе защиты Auto Security Mobile v.1.0.

На определенной стадии (рис. 1) разработки и выбора варианта защиты от атак ВН осуществляется моделирование механизмов восстановления модулей СПО после несанкционированного воздействия. Модель системы механизмов восстановления при изменении целостности модулей СПО можно представить в виде

$$\begin{cases} SCR_{PS} := SCR_{PS} \cup SCR; \\ \forall SCR_{BI} \in SCR_{PS}; \\ \forall CND \in COND(SCR_{PS}); \\ F_{PS} := F_{PS} \cup GOAL(CND), \end{cases} \quad (3)$$

где SCR_{PS} — готовые сценарии политики безопасности; SCR_{BI} — сценарий противодействия при соответствующих атаках ВН; $COND(SCR_{PS})$ — новый сценарий противодействия, добавленный в политику безопасности; F_{PS} — элементы процесса обновления политики безопасности при успешно локализованной атаке $GOAL(CND)$.

В зависимости от функциональности БУ может выполняться весь комплект этапов противодействия несанкционированному доступу или выполняться частично с ограниченным числом этапов и механизмов защиты. Состав и число этапов защиты информации для СПО БУ зависит от комплектации пакета программы Auto Security Mobile v.1.0.

Порядок основных операций выбора и разработки подходящего варианта защиты для СПО, использованный при создании программы Auto Security Mobile v.1.0. защиты для бортового компьютера типа "ШТАТ", контроллера типа BOSCH M. 1.5.4 автомобиля семейства ВАЗ-2110, позволяет снизить критичность и вероятность реализации угроз. Снижение угроз достигается за счет устранения возможности использования возникающих уязвимостей в процессе функционирования отдельных модулей СПО.

Алгоритм противодействия нарушению целостности модулей специализированного программного обеспечения

Одним из моментов стабильного функционирования СПО на базе бортового компьютера типа "ШТАТ" является обеспечение целостности модулей, предотвращение несанкционированного изменения ВН буферных баз данных. Число модулей СПО, буферных информационных баз Ψ_n зависит

от комплектации БУ, объема обрабатываемых и поступающих данных с первичных, вторичных преобразователей систем автомобиля. Функциональная схема модулей СПО бортового компьютера типа "ШТАТ" с блоком аудита событий БУ представлена на рис. 2.

В процессе функционирования СПО и отдельных модулей (1—8) формируются буферные информационные базы Ψ_i , которые могут быть представлены в виде одномерного массива $B = \{\Psi_1, \dots, \Psi_i, \dots, \Psi_n\}$. Внешние случайные или намеренные воздействия ВН могут привести к нарушению целостности модулей, внутренним отказам соответствующих модулей СПО и нарушению работоспособности БУ [7].

Для решения задачи противодействия несанкционированному доступу к модулям СПО БУ рассмотрено несколько вариантов алгоритма защиты, использованных при создании авторской программы защиты Auto Security Mobile v.1.0. Рассмотренные алгоритмы защиты различаются порядком, логикой выполнения действий, составом механизмов защиты и процедур.

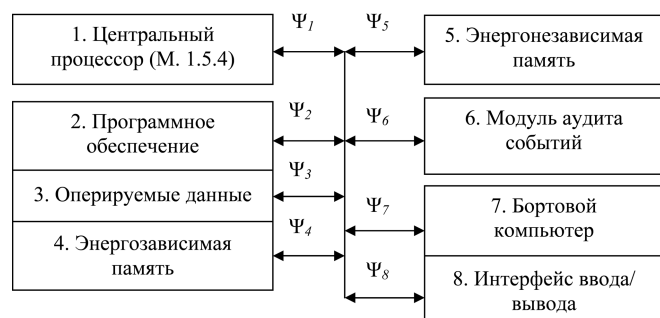


Рис. 2. Функциональная схема модулей СПО бортового компьютера типа "ШТАТ" с блоком аудита событий

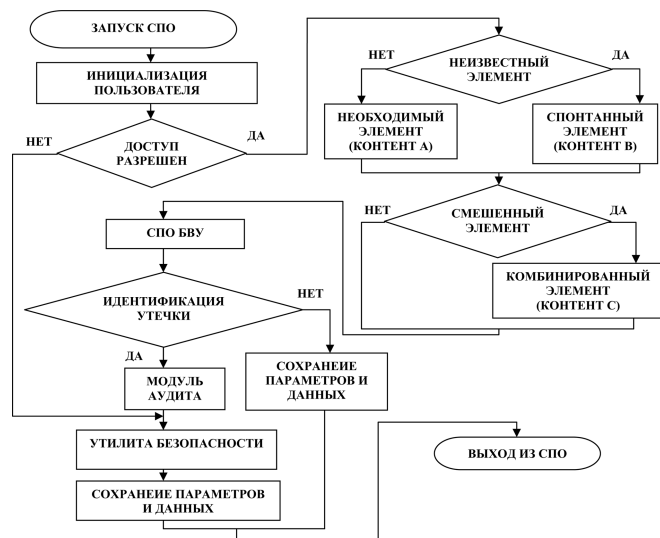


Рис. 3. Фрагмент варианта алгоритма предотвращения несанкционированного доступа на разных стадиях функционирования СПО

Функционирование фрагмента варианта алгоритма (рис. 3), обеспечивающего защиту от несанкционированного доступа и целостность модулей СПО, происходит в момент запуска, завершения сеанса работы и при возможной ликвидации защиты СПО.

Принцип функционирования алгоритма заключается в обработке поступающих данных от первичных, вторичных преобразователей и сопоставлении полученной информации со значениями элементов баз данных сигнатур. База данных сигнатур содержит перечень задекларированных элементов данных, разрешенных к исполнению, которые указаны в политике безопасности программы защиты Auto Security Mobile v. 1.0.

При запуске сеанса СПО БУ данные, получаемые от преобразователей, разбиваются на контенты трех типов: контент *A* (необходимый элемент данных), контент *B* (спонтанный элемент данных), контент *C* (комбинированный элемент данных). Если от преобразователей поступил элемент данных, который занесен в базу сигнатур, то такой элемент включается в контент *A*. При отрицательном результате сопоставления элемента данных с базой сигнатур элемент включается в контент *B*. Существуют моменты функционирования СПО, при которых один элемент данных в различных пропорциях (измеряется в %) может состоять из необходимого и спонтанного элементов данных. При возникновении таких моментов осуществляется сопоставление состава элемента с базой данных сигнатур и происходит включение в контент *C*.

Перед завершением сеанса работы и сохранением параметров функционирования СПО осуществляется проверка данных на возможность утечки информации и нарушения целостности модулей. При идентификации утечки данных и нарушении целостности модулей формируется информационный сигнал, который поступает в модуль утилиты безопасности, где принимается решение о ликвидации возможной опасности несанкционированного воздействия. Все санкционированные и несанкционированные события, выполняемые во время работы сеанса СПО, регистрируются модулем аудита событий. Модуль аудита событий в системах БУ необходим для получения объективных данных о текущем состоянии систем автомобиля, целостности модулей СПО и событиях санкционированного, несанкционированного доступа.

На основе статистических данных с модуля аудита событий выполнен расчет рисков R для различных режимов работы СПО БУ: момент запуска сеанса, рабочий режим, завершение работы. Значение возможного риска определялось для трех режимов работы СПО при четырех базовых угрозах: R_c — риск в момент запуска сеанса; R_i — риск в рабочем режиме; R_g — риск в момент завершения сеанса работы; R_j — риск при возмож-

ной ликвидации защиты систем управления обору-
дованием автомобиля.

Уровень уязвимости нарушения целостности модулей СПО определяется на основе критичности и вероятности реализации угроз при различных режимах работы в момент запуска сеанса (Th_c), в рабочем режиме (Th_i), при завершении работы (Th_a), при возможной ликвидации защиты (Th_l):

$$Th_c = 0,01 \cdot ER_c \cdot P(V)_c; Th_i = 0,01 \cdot ER_i \cdot P(V)_i;$$

$$Th_a = 0,01 \cdot ER_a \cdot P(V)_a; Th_l = 0,01 \cdot ER_l \cdot P(V)_l,$$

где ER_c, ER_i, ER_a, ER_l — критичность реализации угрозы несанкционированного доступа в момент запуска, завершения сеанса, в рабочем режиме СПО и при возможной ликвидации системы защиты (указывается в %); $P(V)_c, P(V)_i, P(V)_a, P(V)_l$ — вероятность реализации угрозы несанкционированного доступа в момент запуска, завершения сеанса, в рабочем режиме СПО и при возможной ликвидации системы защиты (указывается в %).

Риск R для различных режимов (Th_c, Th_i, Th_a, Th_l) работы СПО БУ при возможной реализации одной из угроз определяется формулой

$$R = CThR \cdot D,$$

где D — общая критичность СПО (задается в уровнях); $CThR$ — общий уровень реализуемых угроз по соответствующим модулям СПО.

Расчет риска уровня реализуемых угроз $CThR$ для СПО осуществляется по формуле (3). В качестве значения критичности D оценки уровня реализуемых угроз определена оценка возможной угрозы на соответствующие модули СПО. Данные расчета для СПО системы защиты информации приведены в таблице.

В результате обработки статистических данных, сформированных модулем аудита при проверке работоспособности программы защиты Auto Security Mobile v.1.0, получено, что суммарное сниже-

ние общего уровня реализуемых угроз для СПО составляет $\Delta CThR = 2,0 + 2,8 + 2,4 + 3,7 = 10,9 \%$.

Полученные соотношения снижения общего уровня реализуемых угроз $\Delta CThR$ свидетельствуют об эффективности выбранного алгоритма на основе распределения данных по контентам, обеспечивающего целостность данных и модулей СПО БУ при различных режимах работы.

Разработанный алгоритм позволяет снизить критичность и вероятность исполнения угрозы, реализовать надлежащую защиту от несанкционированного изменения промежуточных информационных баз соответствующих модулей СПО.

Заключение

Учет сформированных условий выбора и разработки подходящего варианта защиты с реализацией соответствующих механизмов защиты модулей СПО позволяет снизить вероятность реализации уязвимостей осуществления угроз. Вариант защиты с учетом трех режимов работы СПО и распределения данных по контентам был использован при создании программы защиты Auto Security Mobile v.1.0 для противодействия нарушению целостности модулей СПО для бортового компьютера типа "ШТАТ", контроллера типа BOSCH M. 1.5.4 автомобилей семейства ВАЗ-2110.

Использование разработанного алгоритма функционирования системы по предотвращению несанкционированного доступа в момент запуска, завершения сеанса работы, при возможной ликвидации защиты СПО в составе программы защиты Auto Security Mobile v.1.0 позволяет снизить на 10,9 % общий уровень реализуемых угроз.

Список литературы

1. ГОСТ ИСО/МЭК 15408—2002. Методы и средства обеспечения безопасности. Критерии оценки безопасности информационных технологий. Ч. III. Требования доверия. М.: Изд-во стандартов, 2002.
2. Рекомендации по стандартизации (Р 50.1.053—2005). Информационные технологии. Основные термины и определения в области технической защиты информации. М.: Изд-во стандартов, 2005.
3. Душкин А. В. Распознавание и оценка угроз несанкционированного воздействия на защищенные информационно-телекоммуникационные системы // Информационные технологии. 2008. № 3. С. 71—75.
4. Валуцких С. А., Рог А. И. О подходе к классификации угроз безопасности автоматизированных систем // Информатика и безопасность. 2003. № 2. С. 76—79.
5. Генералов Д. Н., Шлегель О. А. Идентификация скрытых каналов утечки информации при инсталляции инсайдера в мобильное устройство // Вестник Поволжского государственного университета. 2009. № 6.
6. Kotenko I., Man'kov E. Based Modeling and Simulation of Computer Network Attacks // Proceedings of Fourth International Workshop Agent-Based Simulation 4 (ABS 4), Montpellier, France. 2003. P. 121—126.
7. Molasy M., Erokhhina L., Shlegel O., Shlegel A., Molasy M. Transfer and Processing, Coding of industrial // SYSTEMS, Journal of Transdisciplinary Systems Science, Wroclaw University of Technology, Poland. 2007. Vol. 12. N 3.

Значение критичности оценки реализуемых угроз СПО

Наименование периода функционирования СПО	Значение $CThR$, % (при наличии системы защиты информации)	Значение $CThR$, % (при отсутствии системы защиты информации)	Значение снижения общего уровня реализуемых угроз $CThR$, %
Момент запуска сеанса	31,0	33,0	2,0
Рабочий режим	47,0	49,8	2,8
Момент завершения сеанса	54,0	66,4	2,4
Возможная ликвидация системы защиты информации	82,0	85,7	3,7

УДК 004.7

А. Э. Саак, канд. техн. наук, доц.,
Технологический институт
Южного федерального университета, г. Таганрог,
e-mail: saak@tti.sfedu.ru

Локально-оптимальный синтез расписаний для GRID-технологий

Решается задача оптимального распределения ресурсов в Grid-системах, многопроцессорных вычислительных системах. Задания пользователей интерпретируются прямоугольниками, где стороны равны заявленным ресурсам процессоров и времени обслуживания. Предлагается целевая функция для оптимизации функционирования — минимизация площади объемлющего ресурсного прямоугольника аддитивной графики первоначального семейства ресурсных прямоугольников — заявок пользователей. Детально рассматривается парный синтез ресурсных прямоугольников. Общий случай заданий сводится к парной оптимизации как последовательными, так и параллельными алгоритмами.

Ключевые слова: Grid-система, многопроцессорная вычислительная система, диспетчирование, локальное расписание, оптимальное расписание, принцип минимума площади объемлющего ресурсного прямоугольника

Постановка задачи

Эффективность функционирования Grid-систем, многопроцессорных вычислительных систем (МВС) во многом зависит от качества распределения процессорных и временных ресурсов [1–5]. Заявка пользователя на обслуживание как в центр Grid-технологий, так и диспетчеру операционной системы МВС геометрически может быть представлена в виде координатного прямоугольника, по сторонам которого откладываются значения (далее — измерения), соответствующие заявленным ресурсам обоюдо рода [6]. Например, горизонтальное измерение равно числу единиц ресурса процессоров, а вертикальное — времени, требуемого заявке пользователя, — ресурсному прямоугольнику.

Требуется оптимально распределить в первом координатном квадранте заданное множество ресурсных прямоугольников с соблюдением координатности и аддитивности, т. е. параллельно осям

координат и без наложений. Упомянутую оптимизацию предлагается осуществлять с помощью целевой функции — минимальной площади объемлющего прямоугольника. Таким образом, получается минимизация площади объемлющего прямоугольника аддитивной графики первоначального семейства ресурсных прямоугольников.

Решение задачи

Декларативная задача планарного расписания содержит в себе следующие целевые критерии:

- расходование ресурсов при стадийном обслуживании в сторону минимума;
- рамочное ограничение расходоустойчивости ресурсов [6] в сторону минимизации площади объемлющего прямоугольника, в сторону симметрии;
- полнота использования ресурсов обслуживания в сторону минимизации внутренних пустот, внешних пустот;
- снижение диспетчерских затрат.

Перечисленные целевые критерии в декларативной форме до некоторой степени независимы друг от друга и приводят к мысли об отсутствии определенной совокупности формальных критериев оптимального планарного расписания. Поэтому мы различаем глобально-оптимальное расписание и частично-оптимальные расписания — по некоторому набору целевых критериев. Последним посвящена данная статья.

Пусть имеется линейная полиэдраль k модулей из k координатных прямоугольников каждый. Суммарные длины оснований прямоугольников модуля ограничены одной и той же величиной $N < M$, где M — общий ресурс. Аналогичное условие налагается на вертикальные измерения.

Планарным расписанием называем задачу об оптимальной аддитивной графике в операционном

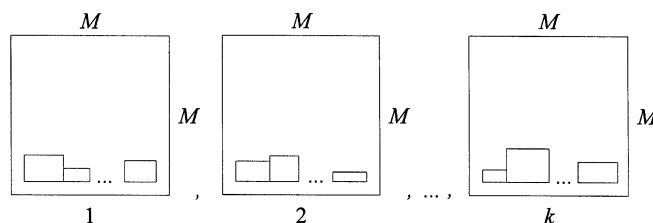


Рис. 1. Множество рамок, подлежащее минимизации

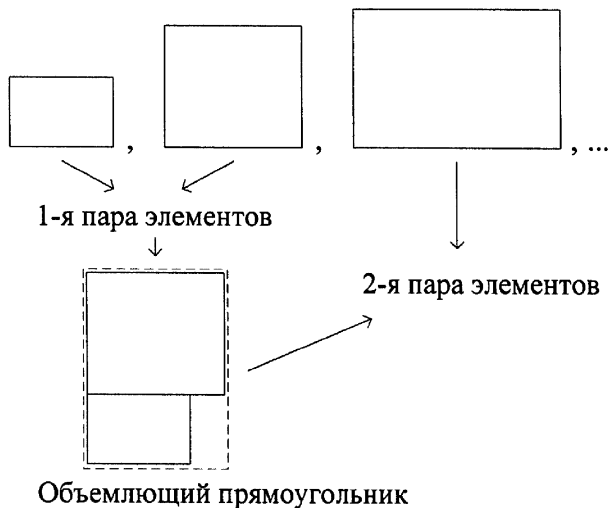


Рис. 2. Последовательно-парный синтез

поле $M \times M$ множества k модулей координатных прямоугольников, аддитивно расположенных внутри квадрата $M \times M$ (рис. 1).

Требуется минимизировать число блоков $M \times M$ так, чтобы в них содержались аддитивно и координатно все указанные ресурсные прямоугольники.

Для каждого модуля находим точный локатор — объемлющий прямоугольник минимальной площади. Получим k прямоугольников

$$\mu(j) \times m(j) \subset M \times M, j = 1, 2, \dots, k,$$

с меньшими параметрами $\mu(j) \times m(j)$ по сравнению с общим ресурсом M .

В случае массива ресурсных прямоугольников

$$\{\mu(j) \times m(j)\}_{j=1}^k, \text{mes}^2[\mu(j) \times m(j)] = \mu(j)m(j) \uparrow$$

(где mes — эвклидова площадь прямоугольника), упорядоченного по возрастанию ($\uparrow$) площади элементов, предлагается последовательная оптимизация аддитивного синтеза. А именно, к оптимальной суперпозиции первой пары в виде объемлющего прямоугольника минимальной площади присоединением третий элемент большей площади, нежели предыдущие, и приходим к новой задаче оптимальной суперпозиции прямоугольной пары и т. д. (рис. 2).

Переходим к параллельной оптимизации аддитивного синтеза. Пусть в элементах линейной полиэдральной

$$\mu(j) \times m(j), \mu(j) \geq m(j), m(j) \downarrow, j \uparrow, j = 1, 2, \dots, k,$$



Рис. 3. Параллельно-парный синтез

превалируют горизонтальные измерения над вертикальными и последние упорядочены по убыванию. Строим кусочно-равномерную мажорантную полиэдраль по целевому критерию минимизации площади пустот, дополняющих прямоугольники исходной полиэдральной до мажорантных прямоугольников-локаторов. С этой целью выделяем часть заданной полиэдральной динамическим выбором вида

$$\sum_{j=1}^q \mu(j)m(j) \leq \frac{1}{2} \sum_{j=1}^k \mu(j)m(j),$$

$$\sum_{j=1}^{q+1} \mu(j)m(j) > \frac{1}{2} \sum_{j=1}^k \mu(j)m(j)$$

числа $q \in Z^1$ в индексном массиве $[1, k] \subset Z^1$ элементов полиэдральной. Далее строим локатор

$$\left(\sum_{j=1}^q \mu(j) \right) \times m(1) \cup \left(\sum_{j=q+1}^k \mu(j) \right) \times m(q+1)$$

горизонтальной суперпозицией указанной пары прямоугольников-блоков (рис. 3).

В силу свойств исходного массива данный синтез объемлющей прямоугольной пары дает кусочно-равномерный локатор искомой структуры.

Таким образом, мы показали, что синтез произвольного числа ресурсных прямоугольников может быть сведен к парному синтезу, который и рассматривается далее.

Рассмотрим задачу суперпозиции двух идентичных ресурсных прямоугольников $a \times b \cup a \times b$ относительно минимума площади объемлющего координатного прямоугольника. Здесь знак $\times$ означает умножение координатных измерений. Локальность массива ресурсных прямоугольников позволяет найти оптимум непосредственным сравнением возможных вариантов задачи.

Тогда вертикальный синтез на рис. 4, слева, дает объемлющий прямоугольник $a \times (b + b)$, а горизонтальный синтез на рис. 4, справа, приводит к объемлющему прямоугольнику $(a + a) \times b$ одинаковой площади.

В более сложном случае суперпозиции двух ресурсных взаимно транспонированных прямоугольников $a \times b \cup b \times a$ имеем два локатора равной площади: горизонтального синтеза на рис. 5, слева,

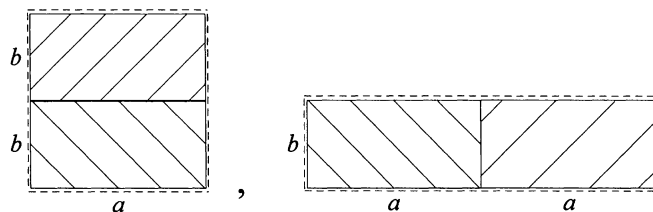


Рис. 4. Варианты аддитивности идентичных прямоугольников

и вертикального синтеза на рис. 5, справа, с площадью

$$\text{mes}^2[(a + b) \times a] = \text{mes}^2[a \times (a + b)] = a^2 + ab.$$

Обобщая далее предыдущую задачу, рассмотрим суперпозицию двух ресурсных прямоугольников со свойством идентичности коизмерений $a \times b \cup \beta \times a$.

Пусть имеет место горизонтальное превалирование измерений и площадей прямоугольников:

$$a \geq b, b \geq \beta.$$

Тогда горизонтальный синтез на рис. 6, слева, дает объемлющий прямоугольник $(a + \beta) \times a$, а вертикальный синтез на рис. 6, справа, приводит к объемлющему прямоугольнику $a \times (b + a)$ большей площади. При условиях

$$a \geq b, b < \beta$$

свойство оптимальности переходит к вертикальному синтезу. Задача для вертикального превалирования

$$a < b, b \leq \beta,$$

допускает аналогичное рассмотрение вариантов.

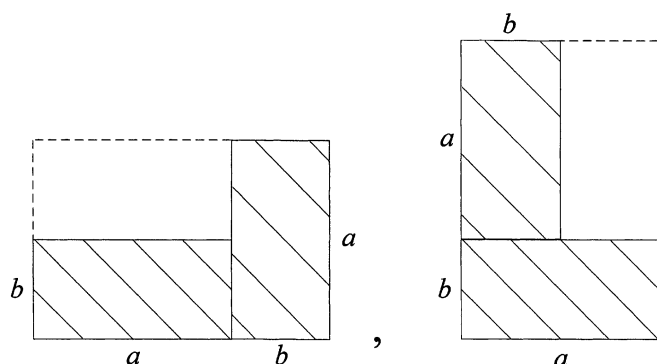


Рис. 5. Варианты аддитивности взаимно транспонированных прямоугольников

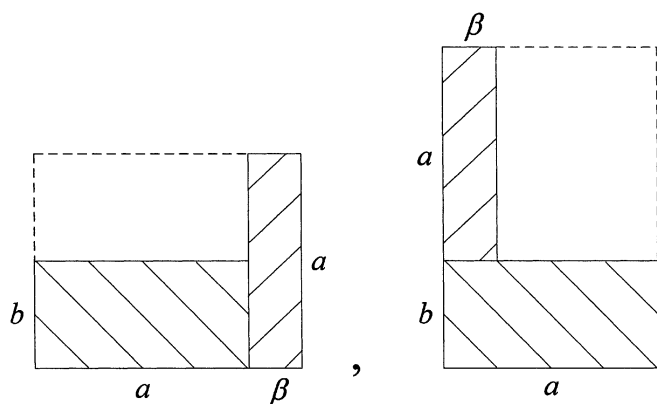


Рис. 6. Варианты аддитивности прямоугольников со свойством идентичности коизмерений

Переходим к общему случаю суперпозиции пары ресурсных прямоугольников с различными измерениями по обоим родам. Рассмотрим задачу оптимального расписания двух прямоугольников

$$a \times b, \alpha \times \beta, a \geq b.$$

В зависимости от соотношения остальных параметров оптимальный синтез двух элементов имеет вид горизонтальной или вертикальной суперпозиции (рис. 7).

Задача двух элементов полностью исследована, и найдены точные объемлющие прямоугольники. Пусть при $a \geq b$ выполнены парные условия:

1) $b \leq \beta, a \ll \alpha$, а именно

$$a \leq \frac{b}{\beta} \alpha;$$

2) $b > \beta, a \gg \alpha$, а именно

$$a > \frac{b}{\beta} \alpha.$$

При любом из условий 1), 2) оптимальным является горизонтальный синтез (рис. 8) с точным объемлющим прямоугольником $(a + \alpha) \times \beta \vee (a + \alpha) \times b$.

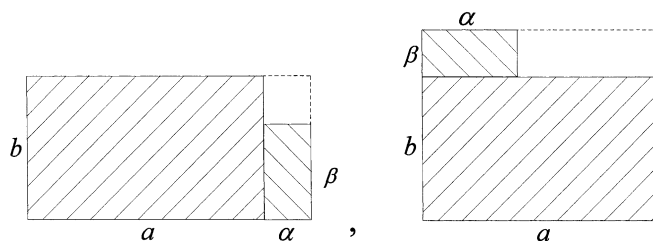


Рис. 7. Варианты оптимального синтеза двух прямоугольников

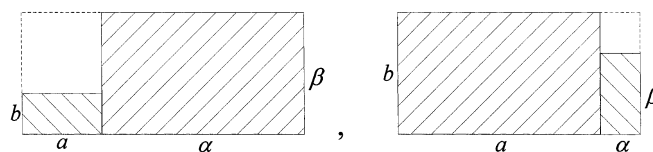


Рис. 8. Оптимальность горизонтального синтеза

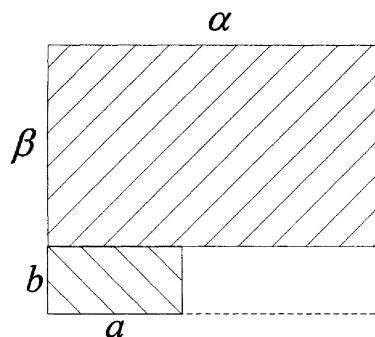


Рис. 9. Вариант вертикального синтеза в условиях $a \leq \frac{b}{\beta} \alpha$

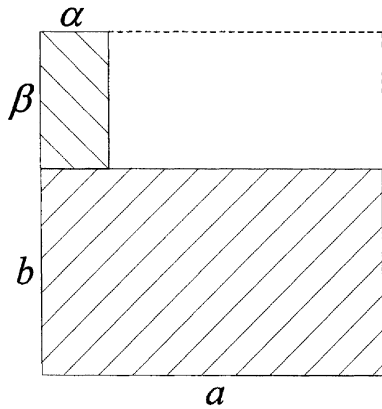


Рис. 10. Вариант вертикального синтеза в условиях $a > 2\frac{b}{\beta}\alpha$

Для доказательства проведем сравнение с вертикальным синтезом. В условиях 1) будем иметь (рис. 9)

$$\rightarrow \rightarrow \text{mes}^2 = a(b + \beta) \rightarrow \rightarrow a(b + \beta) - (a + \alpha) \times \beta = ab - \alpha\beta \geq 0.$$

В условиях 2) аналогично получаем (рис. 10)

$$\rightarrow \rightarrow \text{mes}^2 = a(b + \beta) \rightarrow \rightarrow a(b + \beta) - (a + \alpha) \times b = a\beta - \alpha b > 0.$$

При анализе предпочтительности вертикального синтеза измерения прямоугольников меняются ролями. А именно, пусть при прежнем условии $a \geq b$ выполнены парные неравенства

1) $a \leq \alpha$, $b \ll \beta$, т. е.

$$b \leq \frac{a}{\alpha}\beta;$$

2) $a > \alpha$, $b \gg \beta$, т. е.

$$b > \frac{a}{\alpha}\beta.$$

При любом из условий 1), 2) оптимален вертикальный синтез (рис. 11).

Дополним решение указанной общей задачи рассмотрением специальных частных случаев.

Пару ресурсных прямоугольников

$$a \times b, \alpha \times \beta$$

называем горизонтально упорядоченной при выполнении неравенств превалирования горизонтальных измерений

$$(a + \alpha) \geq (b + \beta), \max\{a, \alpha\} \leq \max\{b, \beta\}.$$

Оптимальное расписание пары с горизонтальным превалированием имеет вид вертикальной

графики (рис. 12) с объемлющим прямоугольником минимально возможной площади

$$\rightarrow \rightarrow \max\{a, \alpha\} \times (b + \beta).$$

Действительно, горизонтальная аддитивность данной пары элементов дает графику вида, представленного на рис. 13, и объемлющая мера

$$\text{mes}^2[(a + \alpha) \times \max\{b, \beta\}] = (a + \alpha)\max\{b, \beta\} \geq (b + \beta)\max\{a, \alpha\} = \text{mes}^2[\max\{a, \alpha\} \times (b + \beta)]$$

превалирует над предыдущей величиной.

В случае вертикальной упорядоченности превалируют вертикальные измерения:

$$(a + \alpha) \leq (b + \beta), \max\{a, \alpha\} \geq \max\{b, \beta\},$$

оптимальная аддитивность становится горизонтальной (рис. 13) и объемлющий прямоугольник минимальной площади приобретает измерения $(a + \alpha) \times \max\{b, \beta\}$ в отличие от объемлющего прямоугольника $\max\{a, \alpha\} \times (b + \beta)$ на рис. 12.

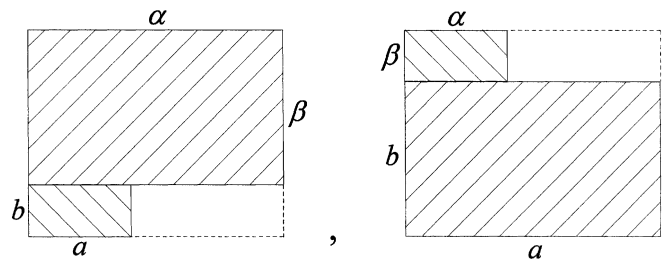


Рис. 11. Оптимальность вертикального синтеза

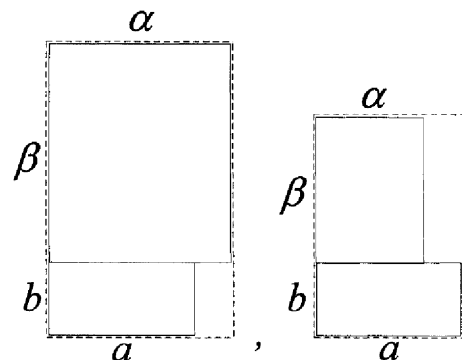


Рис. 12. Оптимальность вертикального синтеза горизонтально упорядоченной пары ресурсных прямоугольников

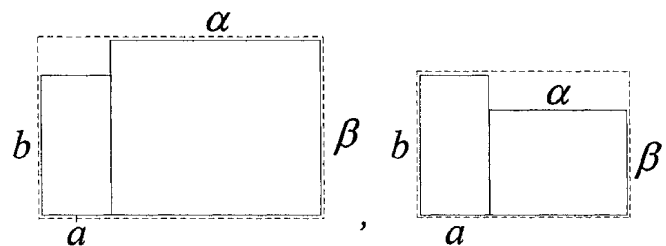


Рис. 13. Горизонтальная аддитивность

Пару ресурсных прямоугольников

$$a \times b, \alpha \times \beta$$

называем относительно горизонтально упорядоченной, если горизонтальные измерения относительно превалируют

$$\left\lceil \frac{a}{\alpha} \right\rceil > \left\lceil \frac{b}{\beta} \right\rceil$$

и имеется диодная сравнимость измерений одного рода

$$\left\lceil \frac{a}{\alpha} \right\rceil = 0; 1, \left\lceil \frac{b}{\beta} \right\rceil = 0; 1.$$

При горизонтальном превалировании оптимальной является вертикальная аддитивность (рис. 14).

Действительно,

$$\frac{(a + \alpha) \max\{b, \beta\}}{(b + \beta) \max\{a, \alpha\}} = \frac{a + \alpha}{b + \beta} \frac{\max\left\{\beta \left(\frac{b}{\beta}, 1\right)\right\}}{\max\left\{\alpha \left(\frac{a}{\alpha}, 1\right)\right\}} = \frac{\frac{a}{\alpha} + 1}{\frac{b}{\beta} + 1} > 1$$

при $\left\lceil \frac{a}{\alpha} \right\rceil > \left\lceil \frac{b}{\beta} \right\rceil$.

Объемлющий ресурсный прямоугольник минимальной площади образован измерениями $\max\{a, \alpha\} \times (b + \beta)$.

При вертикальном относительном превалировании имеем противоположное неравенство

$$\left\lceil \frac{a}{\alpha} \right\rceil < \left\lceil \frac{b}{\beta} \right\rceil,$$

и те же выкладки приводят к оптимальности горизонтальной аддитивности (рис. 15) и объемлюще-

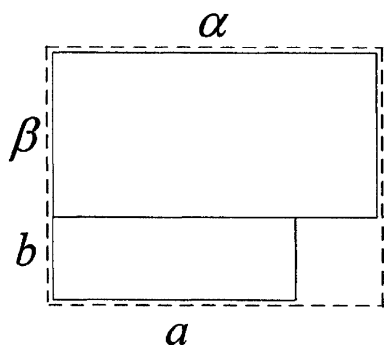


Рис. 14. Оптимальность вертикального синтеза для относительно горизонтально упорядоченной пары прямоугольников

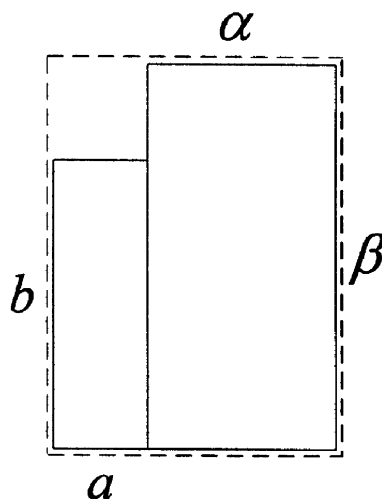


Рис. 15. Оптимальность горизонтального синтеза для относительно вертикально упорядоченной пары прямоугольников

му ресурсному прямоугольнику минимальной площади с измерениями $(a + \alpha) \times \max\{b, \beta\}$, в отличие от предыдущего случая.

Заключение

Таким образом, рассмотрены оптимальные планарные расписания для пары ресурсных прямоугольников общего вида и специальных частных случаев: идентичности, транспонированности, идентичности коизмерений, горизонтальной и вертикальной упорядоченности, относительной горизонтальной и вертикальной упорядоченности.

Значимость оптимального парного синтеза ресурсных прямоугольников автор усматривает в возможности использования в качестве решающего правила диспетчера центра Grid-технологий. В статье приведены соответствующие построения последовательного и параллельного видов синтеза.

Список литературы

1. Барский А. Б. Параллельные информационные технологии. М.: ИНТУИТ; БИНОМ. Лаборатория знаний, 2007. 503 с.
2. Хорошевский В. Г. Архитектура вычислительных систем. Изд-во МГТУ им. Н. Э. Баумана, 2005. 512 с.
3. Воеводин В. В., Воеводин Вл. В. Параллельные вычисления. СПб.: БХВ-Петербург, 2002. 608 с.
4. Барский А. Б. Параллельные информационные технологии в основе Grid-системы // Информационные технологии. 2006. № 12. С. 54–60.
5. Каляев И. А., Левин И. И., Семерников Е. А., Шмойлов В. И. Реконфигурируемые мультимониторные вычислительные структуры / Под общ. ред. И. А. Каляева. Ростов-на-Дону: Изд-во ЮНЦ РАН, 2008. 320 с.
6. Бакенрот В. Ю., Чефранов А. Г. Эффективность приближенных алгоритмов распределения программ в однородной вычислительной системе // Изв. АН СССР. Техн. кибернетика. 1985. № 4. С. 135–148.

А. А. Умбиталиев, д-р техн. наук, проф.,
директор ФГУП "НИИТ",

Н. Н. Шипилов, канд. техн. наук, проф.,
зам. директора ФГУП "НИИТ",

Ш. С. Фахми, канд. техн. наук, доц.,
СПбГЭТУ "ЛЭТИ",
e-mail: shakeebf@mail.ru

Сложнофункциональный блок транскодирования видеоконтента

Рассматривается новый подход к распространению компрессированного видеоконтента с одновременным использованием двух альтернативных технологий сжатия движущихся изображений: на основе трехмерного дискретного косинусного преобразования (не стандартизованная на сегодняшний день технология) и на основе двумерного дискретного косинусного преобразования цифрового телевизионного сигнала (технология, предусмотренная стандартами MPEG-2, MPEG-4/AVC, H.264). В рамках указанного подхода предлагается метод и сложнофункциональный блок быстрого транскодирования потока видеоданных, сжатых по технологии трехмерного дискретного косинусного преобразования, в компрессированный поток стандартного формата MPEG-2.

Ключевые слова: дискретное косинусное преобразование, СФ-блок, транскoder

По данным исследования Cisco® Visual Networking Index Forecast and Methodology [1], суммарный поток видеоданных (IPTV, видео по требованию, обмен индивидуальными видеоматериалами) составит в 2013 г. не менее 90 % пользовательского телекоммуникационного трафика, а на долю только лишь Интернет-видео придется более 60 % всего пользовательского трафика сети. В связи с указанной тенденцией чрезвычайно актуальной становится проблема эффективного сжатия видеоданных в магистральных телекоммуникационных сетях с использованием новых прогрессивных технологий компрессии телевизионного сигнала при сохранении обширной инсталлированной базы цифровых абонентских приемников, телеприставок и видеоплееров. В успешном решении данной проблемы заинтересованы практически все субъекты мультимедийного рынка: магистральные операторы, Интернет-провайдеры, операторы кабельного телевидения и конечные пользователи. Действительно, повышение степени сжатия передаваемого видеоматериала без ухудшения его качества может обеспечить существенное увеличение доходов магистральных операторов и Интернет-

провайдеров, предоставляющих мультимедийные услуги по сбалансированным ценам, поскольку в этом случае становится возможным значительный рост удельного объема услуг, приходящихся на единицу сетевого трафика. В то же время окажутся защищенными инвестиции операторов кабельного телевидения и интересы миллионов потребителей видеоконтента, эксплуатирующих абонентское оборудование стандартов MPEG-2, MPEG-4/AVC, H.264.

К числу современных решений в сфере многоуровневого кодирования распространяемого видеоконтента следует отнести транскoderы Cisco ART [1], позволяющие преобразовывать видеопоток MPEG-4/AVC в формат MPEG-2, не выходя за рамки общей для указанных стандартов технологии компрессии видео на основе двумерного дискретного косинусного преобразования отсчетов (2ДКП-технология). 2ДКП-технология компрессии видео применяется обычно в сочетании с теми или иными методами предсказания и компенсации движения [2]. С теоретической точки зрения предсказание движения относится к классу так называемых трудноформализуемых задач, для которых характерна следующая особенность. При формализации любой задачи для уточнения понятия цели явно или неявно должна задаваться некоторая оценочная функция (функция полезности), выражающая степень приближения к цели или устанавливающая порядок предпочтения возможных вариантов достижения цели.

Для трудноформализуемых задач полностью определить такую функцию либо невозможно, либо связанные с ней вычисления чрезвычайно громоздки. Указанная особенность характерна для видеокодеков, поддерживающих стандарты MPEG-2, MPEG-4/AVC, H.264. Так, полный перебор векторов движения для видео стандартного разрешения в соответствии с H.264 требует производительности вычислений более 10^{22} операций в секунду. Поэтому на практике обычно применяют упрощенные алгоритмы предсказания движения, которые позволяют снизить сложность вычислений. Однако затраты на их реализацию все равно остаются значительными и составляют 60–80 % от всех вычислительных затрат. При этом упрощенные алгоритмы поиска более чем в 3 раза ухудшают эффективность кодирования за счет пропуска так называемых локальных минимумов [3, 4].

Вместе с тем, в контексте декорреляции видеосигнал как многомерная функция имеет оптимальное разложение по всем своим аргументам. Другими словами, для устранения временной корреляции нет существенных формальных оснований использовать средства, отличающиеся от приме-

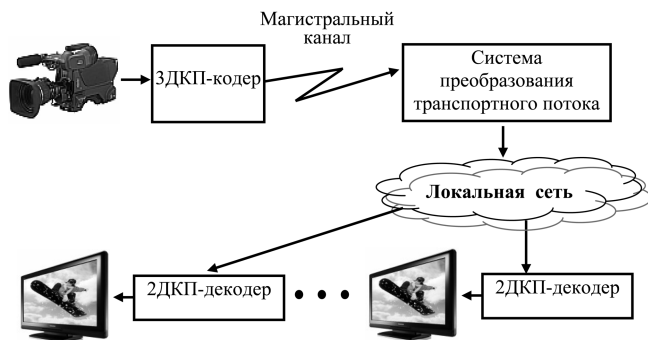


Рис. 1. Распространение видеоконтента с использованием двух технологий компрессии изображений

няемых для пространственной декорреляции сигнала. Кроме того, большинство практически значимых классов изображений характеризуются экспоненциальной автокорреляционной функцией [5], а для модели стационарного сигнала с такой автокорреляционной функцией асимптотически (при большом обрабатываемом блоке) оптимальным разложением, обеспечивающим максимальную скорость сходимости ряда, является дискретное косинусное преобразование [6]. Результаты сравнительной оценки альтернативных технологий [3] показывают, что 3ДКП-технология обеспечивает примерно полтора-двакратный выигрыш в степени сжатия видеoinформации при сохранении субъективного качества изображений.

С учетом вышеизложенного наиболее рациональным на сегодняшний день представляется распространение видеоконтента по схеме: "исходный видеоматериал → 3ДКП-кодирование → магистральная передача → преобразование магистрального транспортного потока, включающее 3ДКП/2ДКП-транскодирование → локальное распределение → 2ДКП-декодирование → отображение" (рис. 1).

Для кодирования исходного контента может быть использован российский 3ДКП-кодер, первый СБИС-прототип которого экспонировался на международной выставке NATEXPO-2009 [7]. Для 2ДКП-декодирования применим любой декодер, поддерживающий стандарт MPEG-2. Система преобразования транспортного потока может быть выполнена в виде СБИС-класса "система на кристалле" [8] и должна содержать ряд сложнофунк-

циональных блоков (СФ-блоков [9]), обеспечивающих:

- анализ магистрального транспортного потока и выделение в нем элементарных 3ДКП-потокaв;
- преобразование каждого элементарного 3ДКП-потокa в элементарный поток MPEG-2, т. е. собственно 3ДКП/2ДКП-транскодирование;
- формирование транспортного потока MPEG-2.

В зависимости от структуры магистрального транспортного потока и способов его обработки в состав системы преобразования может входить один или несколько одинаковых СФ-блоков 3ДКП/2ДКП-транскодирования.

Тривиальным решением задачи 3ДКП/2ДКП-транскодирования является последовательное 3ДКП-декодирование и 2ДКП-кодирование элементарного потока (рис. 2). Однако такое решение характеризуется неоправданно высокой аппаратной и временной сложностью, поскольку оно эквивалентно пятикратному одномерному дискретному косинусному преобразованию отсчетов (тройному обратному в 3ДКП-декодере и двойному прямому в 2ДКП-кодере).

В целях упрощения структуры 3ДКП/2ДКП-транскодера предлагается метод ускоренного преобразования элементарного 3ДКП-потокa в элементарный поток MPEG-2, основанный на учете свойства разделяемости многомерного дискретного косинусного преобразования [10] и совмещении операций деквантования/квантования трансформант с их вычислением. Суть предлагаемого метода поясняется схемой на рис. 3 и состоит в следующем.

В результате энтропийного декодирования и обратного сканирования каждой очередной порции элементарного 3ДКП-потокa формируется трехмерный массив, содержащий $8 \times 8 \times M$ квантованных 3ДКП-трансформант $F_{3ДКП}^{KB}(x, y, t)$, где $x = \overline{0, 7}$ и $y = \overline{0, 7}$ — пространственные координаты, а $t = \overline{0, M-1}$ — временная координата трансформанты $F_{3ДКП}^{KB}(x, y, t)$. В типичном случае M фиксировано и равно 8. В более общем случае [3] значение M варьируется в диапазоне от 1 до 8.

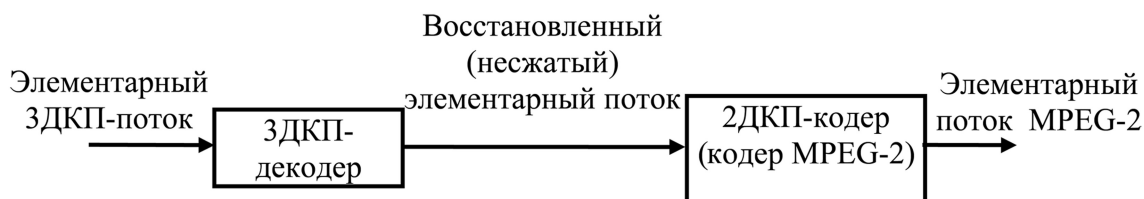


Рис. 2. Тривиальная схема 3ДКП/2ДКП-транскодера

Полученные квантованные 3ДКП-трансформанты подвергаются *модифицированному* обратному одномерному ДКП по временной координате t согласно формуле

$$F_{2\text{ДКП}}^{\text{КВ}}(x, y, z) = \sum_{t=0}^{M-1} \gamma(x, y, z, t) F_{3\text{ДКП}}^{\text{КВ}}(x, y, t) \cos\left(\frac{(2z+1)\pi t}{2M}\right), \quad (1)$$

где $z = \overline{0, M-1}$, $\gamma(x, y, z, t) = \alpha(t) \frac{Q_{3\text{ДКП}}(x, y, z)}{Q_{2\text{ДКП}}(x, y)}$,

$$\alpha(t) = \begin{cases} \frac{1}{\sqrt{M}} & \text{при } t = 0, \\ \sqrt{\frac{2}{M}} & \text{при } t \neq 0, \end{cases}; Q_{3\text{ДКП}}(x, y, z) — \text{коэффициент}$$

квантования 3ДКП-трансформанты $F_{3\text{ДКП}}(x, y, z)$; $Q_{2\text{ДКП}}(x, y)$ — коэффициент квантования 2ДКП-трансформанты $F_{2\text{ДКП}}(x, y)$.

Несложно заметить, что используемые в формуле (1) обобщенные коэффициенты $\gamma(x, y, z, t) = \alpha(t) \frac{Q_{3\text{ДКП}}(x, y, z)}{Q_{2\text{ДКП}}(x, y)}$ реализуют со-

вмещенную функцию нормировки и переквантования трансформант. Значения этих коэффициентов могут быть рассчитаны заранее, а их число равно $128M$ ($8 \times 8 \times M$ значений для $t = 0$ и столько же значений для $t \neq 0$). При этом результатом *модифицированного* обратного одномерного ДКП квантованных 3ДКП-трансформант являются M двумерных массивов, каждый из которых содержит 8×8 квантованных 2ДКП-трансформант.

Двумерные массивы квантованных 2ДКП-трансформант, относящиеся к M последовательным видеокадрам типа Intra [2], загружаются в буфер транскодера (рис. 3), а затем сканируются и подвергаются энтропийному кодированию в соответствии со стандартом MPEG-2.

Следует подчеркнуть, что рассмотренная выше модификация обратного одномерного дискретного косинусного преобразования 3ДКП-трансформант не отличается по своей сложности от обычного одномерного ДКП и позволяет совместить во времени сразу три процедуры транскодирования (рис. 3), а именно: деквантования 3ДКП-трансформант, обратное

одномерное ДКП по времени (т. е. вычисления восьми двумерных массивов 2ДКП-трансформант) и квантования 2ДКП-трансформант. Дополнительное ускорение процесса 3ДКП/2ДКП-транскодирования может быть достигнуто за счет заблаговременного расчета и хранения в памяти транскодера

всех произведений вида $\gamma(x, y, z, t) \cos\left(\frac{(2z+1)\pi t}{2M}\right)$.

Число таких произведений равно $128M^2$, что при $M = 8$ составляет 2^{13} . В случае представления указанных произведений в виде 32-разрядных чисел объем памяти, потребной для их хранения, составит 32 Кбайт, что по современным меркам вполне приемлемо для "систем на кристалле" средней сложности [11].

Таким образом, в отличие от тривиального решения (см. рис. 2), предлагаемый СФ-блок транскодера (см. рис. 3) обеспечивает высокоскоростное преобразование элементарного 3ДКП-видеопотока в элементарный поток MPEG-2 без использования трудоемких многомерных преобразований отсчетов. Полученные авторами результаты сравнительной оценки VHDL-моделей показывают, что

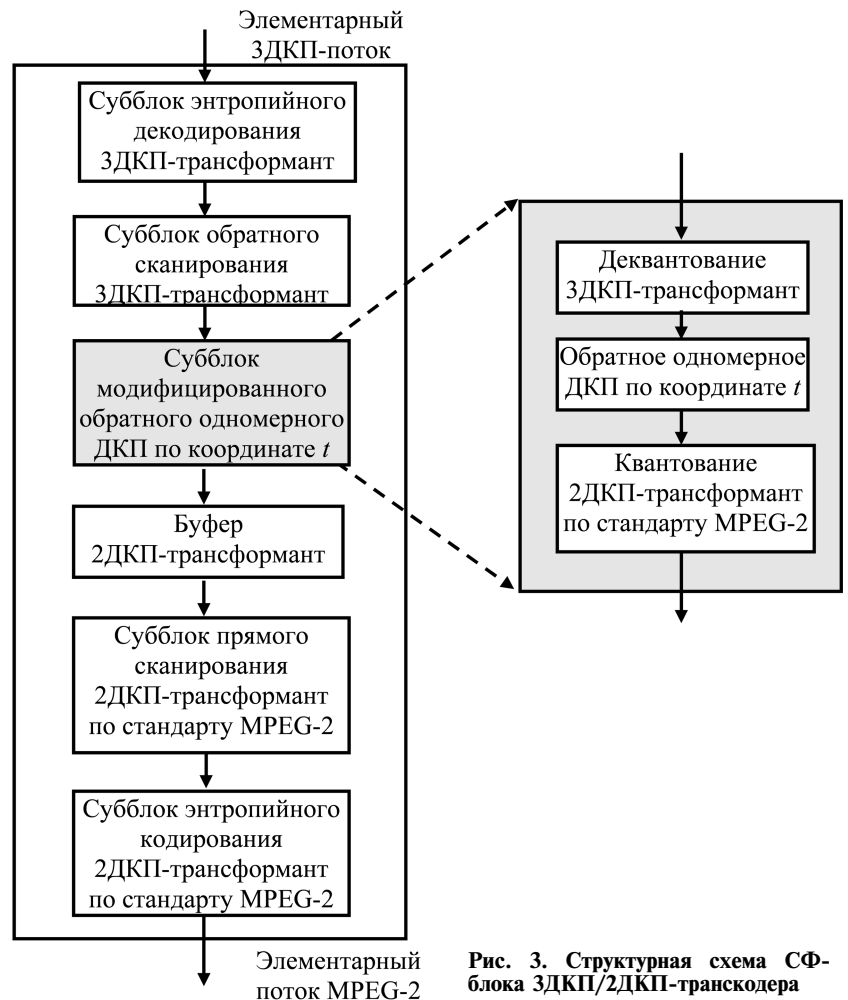


Рис. 3. Структурная схема СФ-блока 3ДКП/2ДКП-транскодера

сложность СФ-блока 3ДКП/2ДКП-транскодера примерно на порядок ниже совокупной сложности 3ДКП-декодера и 2ДКП-кодера, а задержка преобразуемого элементарного потока меньше соответствующей совокупной задержки в 2–3 раза.

Список литературы

1. <http://www.cisco.com>.
2. Ричардсон Я. Видеокодирование. Н. 264 и MPEG-4 — стандарты нового поколения. М.: Техносфера, 2005. 368 с.
3. Umbitaliev A. A., Fahmi Sh. S., Shipilov N. N. et al. A Versatile Real Time Video Codec Based On Three-Dimensional Discrete Cosine Transform. Ibc 2008, RAI International Congress and Exhibition Centre Amsterdam. The Netherlands. Conference 11–15 September 2008. P. 386–391.
4. Lee M., Chan R., Adjieroh D. Quantization of 3D-DCT coefficients and scan order for video compression // J. Vis. Commun. Image Represent. Dec. 1997. 8. P. 405–422.
5. Цифровое кодирование телевизионных изображений / Под ред. И. И. Цуккермана. М.: Радио и связь, 1981. 240 с.
6. Ахмед Н., Рао К. Р. Ортогональные преобразования при обработке цифровых сигналов: Пер. с англ. / Под ред. И. Б. Фоменко. М.: Связь, 1980.

7. <http://www.natexpo.ru/archive/2009>.
8. Немудров В., Мартин Г. Системы на кристалле. Проблемы проектирования и развития. М.: Техносфера, 2004. 216 с.
9. Стешенко В., Руткевич А., Гладкова Е. и др. Проектирование СБИС типа "система на кристалле". Маршрут проектирования. Синтез схемы // Электронные компоненты. 2009. № 1. С. 14–21.
10. Raymond Westwater, Borko Furht. Real-Time Video Compression: Techniques and Algorithms. Kluwer Academic Publishers, 1996. 164 p.
11. Березин В. В., Умбиталиев А. А., Фахми Ш. С., Цыцулин А. К., Шипилов Н. Н. Твердотельная революция в телевидении. М.: Радио и связь, 2006. 312 с.
12. Фахми Ш. С., Зубакин И. Л. Адаптивный алгоритм кодирования видеоинформации на основе трехмерного дискретного косинусного преобразования // Изв. вузов России. Радиоэлектроника. 2010. Вып. 1. С. 49–54.
13. Способ кодирования и декодирования видеоинформации на основе трехмерного дискретного косинусного преобразования. № 2375838, заявка № 2008145795 (приоритет от 19.11.2008) Фахми Ш. С., Ибатуллин С. М., Ибатуллин В. Ф., Иванов К. В., Шипилов Н. Н., Умбиталиев А. А., Цыцулин Л. К. Зарегистрировано в гос. реестре изобретений РФ 10.12. 2009 со сроком действия 19.11.2028.

УДК 519.6;004.72

С. П. Панько, д-р техн. наук, проф.,
А. В. Нагорный, аспирант,
Сибирский федеральный университет
e-mail: anakras@mail.ru

Особенности многостанционного доступа в сеть с разделяемым ресурсом

Исследуется влияние коллизий на производительность информационно-телекоммуникационных сетей с многостанционным доступом рабочих станций пользователей к разделяемому ресурсу. Предлагается методика и достаточно простое техническое решение увеличения производительности таких сетей.

Ключевые слова: Ethernet, многостанционный доступ, разделяемый ресурс, пакетные сети, общая шина

Введение

Технология пакетной передачи информации в информационно-телекоммуникационных сетях получила широкое распространение благодаря возможности обслуживания большого числа пользователей практически одновременно. Основным принципом работы пакетной сети заключается в равноправном доступе в произвольный момент времени рабочих станций (РС) пользователей к общему ресурсу, или, как его еще называют, к разделяемой среде. Наиболее ярким представителем многостан-

ционного доступа является широко известный способ CSMA/CD (Carrier Sense Multiply Access with Collision Detection — метод коллективного доступа с опознаванием несущей и обнаружением коллизий) [1]. В компьютерных сетях способ CDMA реализован на основе семейства протоколов Ethernet с физической общей шиной, а также как элемент звездной топологии на участке сервер — маршрутизатор, к портам которого подключены РС пользователей. Также широко способ представлен в сетях с логической общей шиной, а именно в спутниковых сетях с многостанционным доступом к бортовому ретранслятору и в мобильной связи.

Одной из наиболее важных особенностей этого способа является возникновение коллизий (collisions), которые происходят, когда пакеты двух или более РС поступают на общую шину примерно одновременно. Наложение пакетов приводит к полному или частичному искажению информации, что нарушает работу сети. Некоторое уменьшение негативного влияния коллизий достигается тем, что доступ РС разрешается только при отсутствии сигнала другой РС в общей шине. Однако контроль несущей частоты не снимает проблему коллизий, что объясняется территориальным распределением РС сети. При возникновении коллизии попытка получения доступа к общей шине возможна через интервал времени, который формируется с помощью генератора случайных чисел. Это приводит к появлению очередей.

Задача оценки эффективности сетей с разделяемым ресурсом рассматривалась также в работе [2].

Влияние коллизий на производительность пакетной сети

Под разделяемым ресурсом понимается ресурс, который совместно используют несколько РС, в рассматриваемом случае общий канал передачи данных, который физически реализуется на основе топологии построения сетей "Общая шина".

Коэффициент использования разделяемого ресурса должен приближаться к единице, но не настолько, чтобы очереди пакетов к нему — неминуемое явление в пакетных сетях — были бы постоянно большими, приводя к задержкам и даже потерям пакетов. Только в этом случае функционирование пакетной сети можно считать эффективным.

Коэффициент использования разделяемого ресурса равен отношению трафика, который передается в текущий момент времени, к максимальной пропускной способности. Коэффициент использования обычно измеряют в относительных единицах или процентах [3]. Коэффициент использования ресурса оперирует с полным трафиком, включающим служебную информацию.

Основной недостаток сети с разделяемым ресурсом начинает проявляться при превышении некоторого порогового числа РС и состоит в следующем. Даже та доля пропускной способности разделяемого ресурса, которая должна в среднем доставаться одной РС (например, $10/N$ Мбит/с для сети Ethernet с N РС), не гарантируется. Причина заключается в увеличении вероятности появления коллизий.

В научно-технической литературе широко цитируется экспериментальный график зависимости доли полной пропускной способности сети $R = 0 \dots 10$ Мбит/с, предоставляемой по запросу пользователя, от коэффициента использования $\rho(N, V)$, который является функцией числа РС и генерируемого ими трафика V (рис. 1) [4]. Так, если в сети работает одна РС и она генерирует трафик, включающий данные и служебную информацию объемом 10 Мбит, то сеть предоставит этому пользователю $R = 10$ Мбит/с. Этот случай является идеальным вариантом использования сети, и на графике рис. 1 он отражается прямой 1. При отсутствии коллизий и повторных передач пакетов данных линейная зависимость означает выделение пропорциональной доли пропускной способности каждому пользователю в соответствии со значением текущей загрузки ρ . В реальных же условиях возникновение коллизий приводит к нерациональному использованию R , что уменьшает производительность сети, т. е. в единицу времени через сеть проходит меньше информации при той же электрической скорости (кривая 2 на рис. 1). С ростом ρ сеть начинает работать сама на себя, обрабатывая коллизии, что при $\rho \rightarrow 1$ приводит

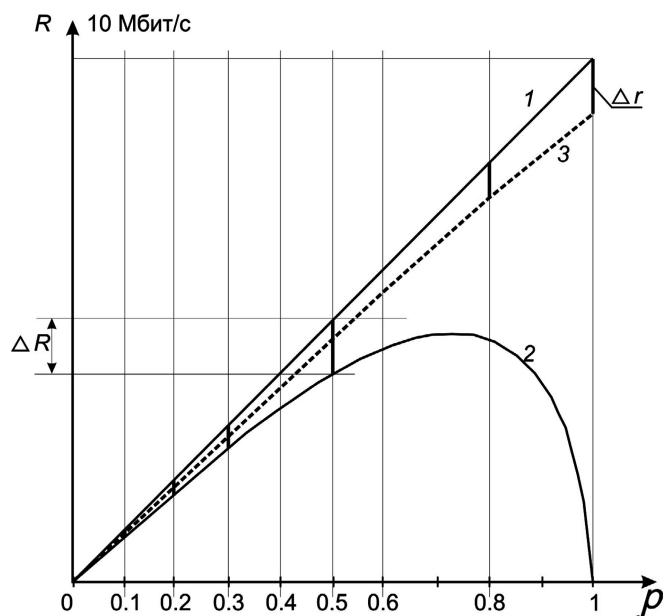


Рис. 1. Зависимость полезной пропускной способности сети Ethernet от коэффициента использования ρ

к $R \rightarrow 0$, т. е. существенному уменьшению производительности сети. Это отражается кривой 2 на рис. 1. Разность мгновенных значений прямой 1 и кривой 2 (обозначено ΔR) определяет потери производительности, вызванные обработкой коллизий.

Анализ появления коллизий

Представляет интерес теоретическое обоснование описанного явления. Рассмотрим работу сети при двух ($X1$ и $X2$) рабочих станциях. Пусть каждая из РС в интервале времени от 0 до T инициирует передачу пакета в момент времени $T1$ и $T2$ соответственно. Далее примем $T = 1$, длительность пакетов данных для всех РС примем dt . При этом возможны четыре комбинации формирования пакетов на общей шине. Две из них соответствуют удачной передаче и две — возникновению коллизии. При $T1 = T2$ происходит 100 %-ное перекрытие пакетов; в других случаях — частичное, но также приводящее к коллизии.

Для оценки вероятности коллизии удобно использовать геометрическое представление (рис. 2), когда вероятность коллизии равна отношению площади, удовлетворяющей перекрытию двух пакетов, к T^2 :

$$\frac{S_6}{S_{ABCD}} = \frac{2dt(T-dt) + dt^2}{T^2}.$$

Эта задача известна в теории вероятностей под собственным именем "задачи о встрече" [5]. Простой расчет показывает, что при двух пакетах и соотношении $dt/T \approx 0,1$ вероятность коллизий не

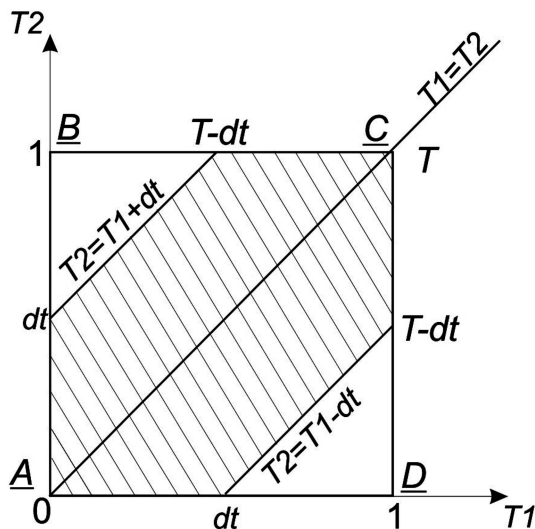


Рис. 2. Геометрическое представление вероятности возникновения коллизии

превышает 20 %. Аналогично, геометрическое представление позволяет оценить вероятность коллизии для столкновения трех пакетов в виде трехмерного объема, отнесенного к T^3 :

$$\frac{S_6}{S_{ABCD}} = \frac{2dt(T-dt) + dt^2}{T^3}.$$

Дальнейшее распространение методики геометрического представления для оценки вероятности столкновения четырех и более пакетов не представляется возможным, поэтому задача решается компьютерным моделированием в соответствии со следующей методикой.

Формируется система независимых случайных чисел A, B, C и т. д., характеризующих моменты появления пакетов, с равновероятным распределением в интервале $0 - T$. В таком случае пакеты $X_i (i = 1, 2, \dots)$ занимают на оси времени положе-

ния $(A \dots A + dt), (B \dots B + dt)$ и т. д. Следующая операция состоит в оценке перекрытия пакетов путем проверки выполнения условий $(A < B < A + dt)$ или $(B < A + dt < B + dt)$ по принципу "каждый пакет с каждым". Удовлетворение хотя бы одного неравенства из группы всех испытаний соответствует появлению коллизии. Вероятность коллизии оценивается как отношение числа неудачных обращений к общей шине к числу испытаний. Результат компьютерного моделирования сведен в таблицу.

Число источников данных	2	3	4	5	6	7
Вероятность коллизии при $dt/T \approx 0,1$	0,19	0,57	0,75	0,92	0,98	0,99
Вероятность коллизии при $dt/T \approx 0,01$	0,02	0,06	0,11	0,19	0,27	0,31

Отсюда следует, что при числе РС около 5 и $dt/T \approx 0,1$ вероятность успешного доступа к ресурсу не превышает 8 %. Единственным решением, снижающим остроту этой особенности, является уменьшение длительности пакета, как это предусматривается в технологии ATM [6]. Однако этот путь, как и увеличение электрической скорости передачи до 100 Мбит/с в Fast Ethernet или до 1 Гбит/с Gigabit Ethernet [7], не являются радикальным средством борьбы с коллизиями, поскольку не устраняют явление по существу.

Число рабочих станции, при котором коэффициент использования разделяемого ресурса начинает приближаться к 1, зависит от типа используемых приложений (потокковая передача данных, мультимедийный трафик, работа с базами данных и т. п.).

Одним из путей разрешения описанного противоречия является использование известных методов уплотнения информации, тем более что в традиционном Ethernet имеется значительный не используемый частотный ресурс кабеля, на основе которого выполнена общая шина.

На рис. 3 приведена структурная схема адаптера (обведена штриховой линией), обеспечивающего доступ рабочей станции к общей шине с частотным разделением источников информации (РС) [8]. Синтезатор частот может формировать определенное число несущих частот, которые не закрепляются за рабочими станциями сети. Предоставление несущей частоты осуществляется по запросу РС и только на время, необходимое для передачи. По умолчанию синтезатор частот формирует частоту f_1 .

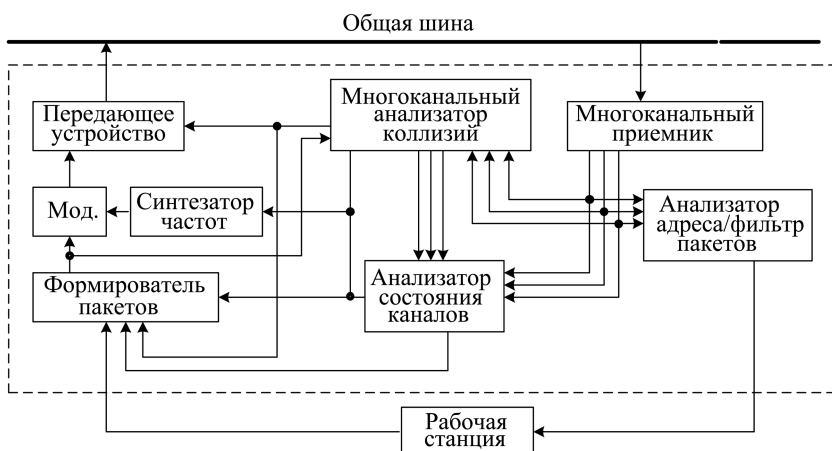


Рис. 3. Адаптер доступа к общей шине

Каждый пакет с выхода формирователя пакетов содержит служебную информацию, включающую адрес получателя и адрес отправителя. Модуляция битовым потоком несущей частоты проводится в модуляторе (Мод.). С помощью передающего устройства сигнал излучается в общую шину.

Многоканальный приемник принимает сигналы с общей шины на всех несущих частотах из набора несущих частот, которые способен формировать синтезатор частот, демодулирует принимаемые сигналы и формирует их в виде битовой последовательности до уровня пакетов данных. В анализаторе адреса/фильтре пакетов проводится вскрытие адресов пакетов, принятых по всем каналам многоканального приемника. Если адрес принятого пакета совпадает с адресом РС, то пакет передается в рабочую станцию. Если адреса не совпадают, то пакет игнорируется. Одновременно битовые сигналы с выходов многоканального приемника поступают на анализатор состояния каналов и на анализатор коллизий. Функцией анализатора коллизий является выявление коллизии на общей шине, т. е. случая совместной работы в одном частотном канале более одной рабочей станции. Это может произойти в силу распределенного характера сети, когда РС ошибочно заняла некоторую рабочую частоту, уже занятую другой РС, но сигнал которой еще не успел поступить на многоканальный приемник рассматриваемой РС. В случае коллизии сигналы двух или более РС, работающих на одной несущей частоте, суммируются на общей шине и, следовательно, информация от обоих РС искажается. Обнаружение факта коллизии проводится путем сравнения передаваемого и принимаемого пакетов. Совпадение пакетов указывает на возможность продолжения работы на занятой несущей частоте. Если пакеты не совпали, т. е. коллизия произошла, на синтезатор частот поступает команда на смену несущей частоты и попытка передачи повторяется.

Функцией анализатора состояния каналов является слежение за использованием набора несущих частот и выбор несущей частоты для сеанса передачи данных. Анализатор состояния каналов оценивает занятость или свободность частотного канала по наличию/отсутствию сигналов на выхо-

дах многоканального приемника и подготавливает выбор ближайшей свободной несущей частоты, пока РС не проявляет активности. Поэтому РС всегда начнет работу на свободной несущей частоте.

Таким путем обеспечивается одновременная работа нескольких РС при значительном уменьшении вероятности коллизий.

Вероятность коллизии в каждом частотном канале определяется числом РС и размером трафика только в этом частотном канале. Использование метода многостанционного доступа с частотным уплотнением позволяет повысить производительность сетей в m раз, где m — число несущих частот. При $N = 10$ (N — число РС в сети), скорости в канале $R = 10$ Мбит/с и $m = 10$ каждая РС имеет возможность работать в своем частотном канале с вероятностью коллизии, равной нулю. При $N/m = 2$ вероятность коллизии в каждом канале будет соответствовать значению вероятности коллизии классического Ethernet при $N = 2$ РС в сети. Обозначим Δr разность между идеальной зависимостью (см. рис. 1, прямая 1) и пропускной способностью при способе многостанционного доступа с частотным уплотнением (см. рис. 1, кривая 3). Эта кривая построена для $m = 2$, что эквивалентно увеличению r в 2 раза относительно классического варианта для всех значений ΔR . Следует отметить, что увеличение числа несущих частот до $m > 2$ позволяет увеличить r , но несущественно относительно $m = 2$.

Список литературы

1. Норенков И. П., Трудоношин В. А. Телекоммуникационные технологии и сети. М.: Изд-во: МГТУ им. Н. Э. Баумана, 2001.
2. Лукьянов С. А., Ховансков С. А., Серогодский Д. И., Горягина Т. М., Котегов М. Г. Зависимость полезной пропускной способности от размера кадра для сети стандарта Ethernet // Современные наукоемкие технологии. 2008. № 3.
3. Олифер В. Г., Олифер Н. А. Компьютерные сети. СПб.: Питер, 2001.
4. Олифер Н., Олифер В. Мосты и коммутаторы // Журнал сетевых решений "LAN". 2002. № 11.
5. Гнеденко Б. В. Курс теории вероятности. М.: Государственное издательство физико-математической литературы, 1961.
6. Мартин Д., Чапмен К. К., Либен Д. АТМ. Архитектура и реализация. 2000. 224 с.
7. Оглтри Т. Модернизация и ремонт сетей, 2005. 1328 с.
8. Способ многостанционного доступа в сеть Ethernet: пат. Российской Федерации RU (11) 2368085 (13) С. 1 / Нагорный А. В., Панько С. П.

Т. С. Аббасова, канд. техн. наук,
Российский государственный университет
туризма и сервиса, РГУТиС
e-mail: abbasova_univer@mail.ru

Повышение эффективности эксплуатации высокоскоростной кабельной системы с помощью виртуальных технологий

Рассмотрены методы повышения эффективности эксплуатации кабельной системы для центра обработки данных. Приведена методика расчета основных тестируемых параметров сети для кабеля "витая пара" категорий 6, 6а с учетом особенностей эксплуатации при работе на высоких частотах для достижения пропускной способности 10 Гбит/с. Представлена аналитическая оценка эффективности использования виртуальных технологий и статистическая оценка стоимости владения виртуальной структурой.

Ключевые слова: вычислительные мощности, электромагнитная совместимость, виртуальный режим

Цель работы: минимизация эксплуатационных затрат и сведение к минимуму электромагнитных помех в 10-гигабитных электрических трактах для обеспечения заданной пропускной способности. Объектом исследования является горизонтальная подсистема структурированной кабельной системы (КСК) на основе кабеля типа "витая пара" для крупного центра обработки данных (ЦОД).

Совместно используемая инфраструктура центров обработки данных служит для размещения данных и их обработки предприятиями не очень большого масштаба.

ЦОД — это сложное инженерное сооружение, предназначенное для централизованного размещения и обслуживания компьютерного оборудования предприятия. Предприятиям не очень большого масштаба легче арендовать вычислительное оборудование ЦОД у более крупной организации, чем иметь расходы на содержание своего собственного ЦОД (концепция *Cloud Computing*).

Услуги по размещению данных предоставляют возможность передачи всех рисков, связанных с размещением и поддержкой вычислительного оборудования, провайдеру. Для уменьшения плотности оборудования (портов оборудования) на единицу площади ЦОД (и соответствующего уменьшения инженерного оборудования для отвода теплоты), увеличения коэффициента загрузки сервера и снижения энергопотребления применяются технологии виртуализации.

Также ставится задача резервирования данных (в электронном виде) — это доставка в центр, хранение, постоянное обновление, восстановление корпоративных данных клиента и незамедлительное предоставление доступа к ним в момент наступления аварийного события. Эти услуги подразумевают создание резервного ЦОД (его также часто называют "резервный офис"). В соответствии с указанным назначением проектируются основной и резервный центры обработки. Структурная схема их организации представлена на рис. 1.

На рис. 1 приняты следующие обозначения: СБЭ — система бесперебойного электроснабжения; СВК — система вентиляции и кондиционирования; СОП — система обнаружения пожара; КСБ — комплексная система безопасности; КСК — структурированная кабельная сеть.

Web-узел отображает информацию о компании даже после аварийной ситуации при наихудшем варианте сценария. Если системы работают нормально, то с помощью вычислительного оборудования резервного ЦОД осуществляется удаленная диагностика основного ЦОД и другие задачи.

Резервный ЦОД может обслуживать несколько основных ЦОД, наиболее важные функции основного ЦОД могут быть продублированы в нескольких резервных ЦОД. При передаче обслуживания применяется децентрализованное управление для повышения надежности.

Требования к пропускной способности каналов ЦОД — 10 Гбит/с. Группой IEEE 802.3 подготовлены спецификации интерфейсов для передачи данных по кабелям из витых пар со скоростью 10 Гбит/с, для достижения которой по каждой из четырех пар горизонтального кабельного тракта (100 м, четыре точки коммутации) в двух направлениях должны передаваться сигналы со скоростью 2,5 Гбит/с. Скорость 2,5 Гбит/с позволяет организовывать стык между локальной сетью ЦОД и глобальными сетями АТМ (от англ. *asynchronous transfer mode* — асинхронный режим переноса информации). Иерархические ступени сети АТМ имеют скорость передачи около 2,5 Гбит/с. Типичная дальность передачи в ЦОД для горизонтальных каналов составляет 50...100 м, поэтому в системах "эконом-класс" нет необходимости использовать оптические варианты сетевого интерфейса длиной от сотен метров до нескольких километров для достижения заданной пропускной способности [1]. Распространены следующие конструкции кабеля "витая пара" для приложений 10GBaseT: неэкранированный кабель UTP (от англ. *unshielded twisted pair* — неэкранированная витая пара) категории 6а (диаметр ~7,9 мм); экранированный кабель FTP (от англ. *foiled twisted pair* — экранированная витая пара) категории 6 (диаметр ~6,3 мм).

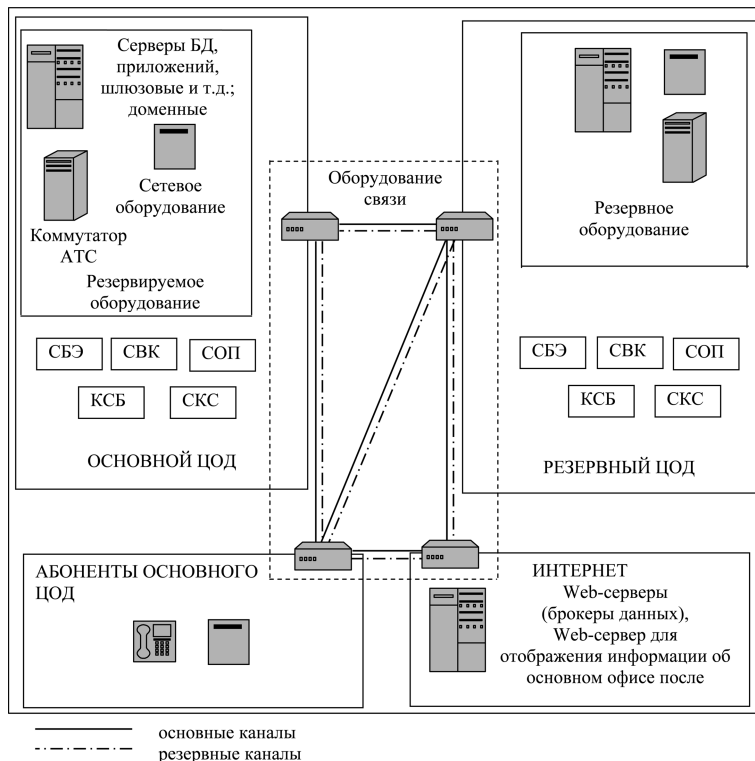


Рис. 1. Принцип построения основного и резервного центров обработки данных

Наряду с перечисленными существуют новые предложения: экранированный кабель категории 7 (диаметр ~7,8 мм); дважды экранированные кабели категории 8 с медными проводниками (диаметр ~6,0 мм).

Рассмотрим особенности эксплуатации электрических кабельных линий для передачи информационных приложений со скоростью 10 Гбит/с. Увеличение пропускной способности до 10 Гбит/с приводит к переходу на более высокие частоты передачи (250 МГц и выше), поэтому возникает необходимость соблюдения более жестких технических требований к электротехническим параметрам [2].

На частоте 300 МГц и выше наибольшее влияние в электрическом кабельном тракте оказывает межкабельная переходная помеха — АХТ (от англ. *alien crosstalk* — межкомпонентные перекрестные помехи). При двунаправленной передаче наибольший интерес представляет сигнал помехи на ближнем конце — АNEXT (от англ. *alien near end crosstalk* — межкомпонентные перекрестные помехи на ближнем конце) [3, 4]. Помеха на ближнем конце имеет большее значение, поскольку уровень сигнала во влияющей линии по направлению к концу затухает. Помехи на дальнем конце — АFEXT (от англ. *alien far end crosstalk* — межкомпонентные перекрестные помехи на дальнем конце) — определяются при использовании одно-временной передачи и приема информации по

всем четырем парам. Их также необходимо учитывать при проектировании и эксплуатации высокоскоростной кабельной системы. Межкабельные помехи на ближнем конце линии зависят от характеристик соединительных компонентов. Межкабельные помехи на дальнем конце линии зависят от характеристик кабелей.

Начиная с частоты 500 МГц рекомендовано обязательное экранирование кабеля, а в диапазоне 300...500 МГц могут использоваться менее дорогие неэкранированные конструкции. Неэкранированные кабели в сравнении с экранированными имеют следующие преимущества: меньшая стоимость; меньшая трудоемкость монтажа и эксплуатации; отсутствие повышенных требований к внутреннему заземляющему контуру здания; лучшие массогабаритные показатели; меньший радиус изгиба.

Оценим параметр защищенности (отношение сигнал/шум) при передаче электрических сигналов в диапазоне частот 250...500 МГц по неэкранированным кабельным линиям на основе "витая пара". Относительно межкабельных наводок на ближнем конце он оценивается формулой

$$A = A_{XN} - A_L, \quad (1)$$

где A_{XN} — межкабельное переходное затухание на ближнем конце (N от англ. *near* — ближний, X — указывает на характер помех (межкабельные)), дБ; A_L — погонное затухание тракта передачи длиной L , дБ.

Суммарный параметр защищенности от межкабельных помех на ближнем конце PSACR (от англ. *powersum attenuation to crosstalk ratio* — суммарное отношение погонного затухания к переходному затуханию) оценивается формулой

$$A_{X\Sigma} = \Sigma A_{XN} - A_L, \quad (2)$$

где ΣA_{XN} — суммарные переходные помехи на ближнем конце, дБ.

Чем больше переходное затухание A_{XN} и меньше погонное затухание A_L кабельной пары, тем лучше параметр защищенности. Так как затухания и наводки разъемов и кабелей при передаче в диапазоне частот 300...500 МГц становятся сопоставимыми, то для расчета допустимого затухания кабельного тракта (проброс + шнур сетевого оборудования + + кроссовый шнур + шнур рабочего места) необходимо использовать формулу

$$A_L = \Sigma A_p + A_{k/100}(L_k + 1,2\Sigma L_{ш}), \quad (3)$$

где ΣA_p — сумма максимально допустимых затуханий, вносимых всеми разъемами, дБ; $A_{k/100}$ — максимально допустимое затухание горизонтального кабеля на длине 100 м, дБ; L_k — фактическая

длина горизонтального кабеля тракта или проброса, м; 1,2 — коэффициент; $\Sigma L_{ш}$ — фактическая сумма длин всех шнуров тракта или проброса, м.

На качество высокоскоростного тракта передачи влияют возвратные потери R_L (от англ. *return loss*). Они характеризуют согласованность с кабелем отдельных элементов тракта (соединители, панели переключений, шнуры). При низких скоростях согласованность не оказывает заметного влияния. Характеристика возвратных потерь R_L в первую очередь зависит от правильного подбора компонентов проводки, а затем — от качества монтажных работ на объекте.

Для ЦОД "эконом-класс" рекомендуется использование неэкранированных конструкций кабеля "витая пара". На архитектурной фазе проектирования центра можно предусмотреть максимальную длину горизонтальной подсистемы до 55 или 100 м, чтобы использовать кабели категорий 6 (до 55 м), 6а (до 100 м). Для того чтобы уменьшить в неэкранированных системах возвратные потери, наводки в разъемах (которые становятся сопоставимыми с наводками в кабелях) и других соединительных элементах, требуется улучшение конструкции разъемов, коннекторов, штекеров, коммутационных панелей.

Уровень электромагнитных помех в электрическом тракте при совместной прокладке питающих и неэкранированных информационных кабелей можно свести к минимуму с помощью кабельного желоба с высокими краями. При этом силовые кабели должны быть распределены по одному краю, а информационные кабели — по другому, с использованием металлического разделителя [5, 6].

Поскольку заполнение коробов не должно превышать 40 % в данном диапазоне частот для 10-гигабитных неэкранированных СКС, в формулу для вычисления ширины кабельного канала необходимо ввести коэффициент пространственного разнесения в кабельном желобе k_p , который с учетом 40 %-ного заполнения равен 0,6 [7].

Таким образом, формула для расчета ширины кабельного канала для неэкранированных конструкций кабеля "витая пара" в диапазоне частот 300...500 МГц запишется в следующем виде:

$$B_k = (1 + k_p) \nu d_{и} + d_c + r, \quad (4)$$

где B_k — коэффициент пространственного разнесения в кабельном желобе; ν — число информационных кабелей; $d_{и}$ — диаметр информационного кабеля, мм; d_c — диаметр силового кабеля, мм; r — расстояние между силовым и информационным кабелями, мм.

Возвратные потери, а также переходные затухания между парами NEXT (от англ. *near end crosstalk* — перекрестные помехи на ближнем конце) и ELFEXT (от англ. *equal level far end crosstalk* — равноуровневые перекрестные наводки на дальнем

конце) можно прогнозировать и подавить с помощью средств цифровой обработки сигнала DSP (от англ. *digital signal processing* — цифровая обработка сигнала). Чтобы гарантировать низкий уровень внутренних наводок, подавляемых средствами DSP, реализуются экстраполированные значения параметров переходных наводок на ближнем конце NEXT, суммарных переходных наводок на ближнем конце PSNEXT (от англ. *powerSum near end crosstalk* — суммарное переходное затухание на ближнем конце), суммарных разноуровневых переходных наводок на дальнем конце PSELFEXT (от англ. *powerSum equal level far end crosstalk* — суммарное равноуровневое переходное затухание на дальнем конце) и возвратных потерь каналов класса EA (кабель "витая пара" категорий 6, 6а).

В таблице определены параметры, характеризующие кабельный канал с частотой передачи 300...500 МГц.

Расчетные формулы, приведенные в таблице, свидетельствуют о том, что при выборе конструкции коммутационных панелей, разъемов и другого соединительного оборудования необходимо, чтобы их категория соответствовала категории кабеля (например, для кабеля категории 6а не рекомендуется применять коммутационный шнур категории 6). При выборе способа монтажа для неэкранированных конструкций необходимо учесть, что при отмене практики фиксации кабельных каналов стяжками для каналов вертикальной кабельной подсистемы увеличиваются продольные нагрузки, приводящие к растягиванию кабелей и ухудшению их параметров [8], поэтому надо предусмотреть механические зажимы, уменьшающие продольные нагрузки (например, зажим типа *Kellems*).

При использовании виртуальных технологий при подключении кабелей связи к вычислительному оборудованию уменьшается число портов коммутационного поля, так как количество оборудования (и принадлежащих ему портов) сокращается. Это очень важно для высокоскоростных 10-гигабитных приложений на неэкранированном кабеле "витая пара", так как в этом случае недопустимо увеличивать плотность портов коммутационного поля. Если же эти порты будут пространственно разнесены (для увеличения параметра защищенности), то ухудшатся массогабаритные характеристики коммутационного оборудования.

При оценке затрат на энергопотребление серверов необходимо учесть, что переход на виртуализацию позволяет уменьшить расходы не только на электроэнергию, но и на охлаждение оборудования. Виртуальная инфраструктура оценивается исходя из тарифов, которые могут быть рассчитаны для виртуальных элементов (например, делением полной стоимости владения на число виртуальных элементов, с учетом калькуляции ТСО

(от англ. *total cost of ownership* — совокупная стоимость владения)).

Оценим статистические данные загрузки вычислительного оборудования. Средняя загрузка работающих на предприятиях серверов — около 10...15 %. Консолидация физических серверов с помощью средств виртуализации позволяет перенести около 10 физических серверов в виртуальные машины, которые запускаются на одном физическом сервере и поднимают его загрузку до 50...60 % [9]. Этот факт не означает снижения

стоимости владения серверами в 10 раз. В реальных проектах часто одновременно происходит модернизация серверного парка с консолидацией, например, на "blades"-серверы ("лезвия"). С одной стороны, стоимость новых "blades"-серверов увеличивается; с другой стороны, затраты при их эксплуатации уменьшаются, в первую очередь, затраты на электроэнергию. Общее число необходимых для серверов сетевых портов и маршрутизаторов может быть прямо пропорционально уменьшено, и, соответственно, прямо пропорционально уменьшится эффективная плотность портов. В некоторых случаях для меньшего числа серверов могут понадобиться дополнительные порты, что определяется на стадии планирования и разработки архитектуры решения. Статистические исследования показали, что показатель ТСО для ИТ-инфраструктуры снижается на 60...80 % [10], как показано на рис. 2 (см. третью сторону обложки).

Таким образом, виртуальный режим работы увеличивает эффективность эксплуатации, так как уменьшается плотность портов коммутационного поля, количество кабельных линий и вычислительного оборудования.

Способы минимизации мощности межкабельной помехи позволяют улучшить требуемое качество услуг при передаче данных со скоростью 10 Гбит/с. При виртуализации вычислительных ресурсов можно снизить эксплуатационные затраты до 60 %.

Список литературы

1. Аббасова Т. С. Антикризисный сервис информационных систем в электротехнических комплексах // Электронное периодическое издание "Сервис в России и за рубежом". 2008. Вып. 8 / Материалы IV Межвузовской научн.-техн. конф. "Проблемы развития электротехнических комплексов и информационных систем". URL: http://rguts.ru/electronic_journal
2. Аббасова Т. С., Умудумов О. Ф. Технические средства для сервисного обслуживания высокоскоростных электрических трактов СКС // Вестник МГУС. 2008. № 1 (4). С. 77—85.
3. Олер А. М., Шикетанц Д. В. Межкабельная переходная помеха: теория и измерение // LAN. 2006. № 1. С. 26—32.
4. Артюшенко В. М., Енютин К. А. Анализ межкабельных переходных помех в электрических кабельных системах, работающих в диапазоне частот свыше 500 МГц // Электротехнические и информационные комплексы и системы. 2008. Т. 4, № 3. С. 28—35.
5. Риглсуорт К., Воловдов А. Экранирование кабельных коробов // Сети и системы связи. 2000. № 1. С. 34—39.
6. EN 50174-2. Информационные технологии. Установка кабельной системы. Ч. 2: Проектирование и методы установки внутри зданий.
7. Аббасова Т. С., Артюшенко В. М. Методы инсталляции и проектирования электрических кабельных линий в 10-гигабитных системах связи // Электротехнические и информационные комплексы и системы. 2009. Т. 5, № 2. С. 8—16.
8. Аббасова Т. С., Артюшенко В. М. Электромагнитная совместимость электропроводных кабелей и коммутационного оборудования высокоскоростных структурированных кабельных систем // Электротехнические и информационные комплексы и системы. 2008. Т. 4, № 4. С. 22—29.
9. Торок Э. Реальное планирование — виртуальная работа // LAN. 2008. № 02. С. 37—44.
11. Колесов А. Переход на технологии виртуализации // PC Week/RE. 28 октября — 3 ноября 2008. № 40 (646).

Параметры передачи для 10-гигабитного кабеля типа "витая пара"

Параметр электропроводного кабеля	Обозначение, ед. изм.	Формула для расчета
Погонное затухание	A_L (Att), дБ	$A_L = \alpha L / 100$, где α — коэффициент погонного затухания, L — длина кабеля, м $A_L = k_1 \sqrt{f} + k_2 f + \frac{k_3}{\sqrt{f}}$, где k_1, k_2, k_3 — коэффициенты, зависящие от категории кабеля и фирмы-производителя; f — частота передачи, МГц; $\alpha = \frac{k_1 \sqrt{f} + k_2 f + \frac{k_3}{\sqrt{f}}}{L}$
Межкабельное затухание на ближнем конце	A_{XN} (ANEXT), дБ	Для канала класса EA длиной 100 м: $A_{XN}(f) = 62 - 15 \lg(f/100)$, для канала класса E длиной 55 м: $A_{XN}(f) = 47 - 15 \lg(f/100)$
Параметр защищенности	A (ACR), дБ $A_{X\Sigma}$ (PSACR), дБ	$A = A_{XN} - A_L$, $A_{X\Sigma} = \Sigma A_{XN} - A_L$
Возвратные потери	R_L , дБ	$R_L = k_4 - 10 \lg(f/20)$, где k_4 — коэффициент, зависящий от категории кабеля и фирмы-производителя
Максимальная разность задержек прохождения между всеми парами	S_D (Skew/Delay), мс	$S_D = \max t_i - t_j $, $i, j = 1, v_n$, где $t_{i,j}$ — моменты прихода импульсов на дальний конец витых пар при одновременной подаче зондирующих сигналов на все пары; v_n — число пар в тестируемом кабеле
Ширина кабельного канала	B_k , мм	$B_k = (1 + k_p) v d_n + d_c + r$, где k_p — коэффициент пространственного разнесения в кабельном желобе ($k_p = 0,6$ для 10-гигабитных неэкранированных СКС); v — число информационных кабелей; d_n — диаметр информационного кабеля, мм; d_c — диаметр силового кабеля, мм; r — расстояние между силовым и информационным кабелями, мм

УДК 519.857

В. И. Струченков, д-р техн. наук, проф.,
А. Н. Козлов, аспирант, **А. С. Егунов**, аспирант,
 Московский институт
 радиотехники, электроники, автоматики
 (технический университет)
 e-mail: str1942@mail.ru

Кусочно-линейная аппроксимация плоских кривых при наличии ограничений

При автоматизированном проектировании трасс линейных сооружений возникают задачи аппроксимации плоских кривых, заданных дискретно, последовательностью элементов определенного вида (отрезки прямых, дуги окружностей или парабол второй степени, а также клотоид) при наличии ограничений на параметры элементов. Число элементов аппроксимирующей кривой неизвестно. В статье рассматривается простейшая из задач поэлементной аппроксимации, в которой элементами являются отрезки прямых, при наличии ряда ограничений. Задача решается с помощью динамического программирования.

Ключевые слова: аппроксимация, ограничения, динамическое программирование

1. Постановка задачи

Аппроксимируемая кривая задана декартовыми координатами ряда точек $L_1 = \{x_i, y_i\}$ ($i = 1, 2, \dots, n$), т. е. представлена в виде ломаной линии, которая необязательно является графиком однозначной функции. Требуется найти другую ломаную с вершинами $L_2 = (X_j, Y_j)$ ($j = 1, 2, \dots, k$) при следующих условиях:

$$\min_{L_2} K(L_1, L_2), \text{ при } f_r(X, Y) \leq 0, r = 1, 2, \dots,$$

где $K(L_1, L_2)$ — заданный количественный критерий отклонения L_2 от L_1 , а система $f_r(X, Y) \leq 0$ задает ограничения на параметры элементов L_2 . Число ограничений определяется числом элементов ($k - 1$), которое неизвестно. Крайние точки (X_1, Y_1) и (X_k, Y_k) заданы. Кроме того, могут быть заданы начальное и конечное направления (элементы примыкания).

Если L_1 и L_2 соответствуют однозначным функциям $L_1(x)$ и $L_2(x)$, то

$$K(L_1, L_2) = \int_{x_1}^{x_n} (L_1(x) - L_2(x))^2 dx$$

при $X_1 = x_1$ и $X_k = x_n$ и следующих ограничениях:

$$c_j \leq (Y_{j+1} - Y_j)/(X_{j+1} - X_j) \leq d_j; \quad (1)$$

$$p_j \leq (Y_{j+1} - Y_j)/(X_{j+1} - X_j) - (Y_j - Y_{j-1})/(X_j - X_{j-1}) \leq q_j; \quad (2)$$

$$l_{\min} \leq X_{j+1} - X_j \leq l_{\max}, \quad (3)$$

где j — номер вершины искомой ломаной; c_j и d_j — заданные минимальные и максимальные значения уклонов элементов соответственно; p_j и q_j — заданные минимальные и максимальные значения разностей уклонов; $l_{\min}$ и $l_{\max}$ — заданные минимальная и максимальная проекции длин элементов на ось X . Дополнительно могут быть заданы ограничения в некоторых точках, абсциссы которых x_r могут и не совпадать с абсциссами заданных точек

$$L_2(x_r) - L_1(x_r) \leq v_r \text{ и/или } w_r \leq L_2(x_r) - L_1(x_r), \quad (4)$$

где $x_1 < x_r < x_n$.

Все величины, кроме (X_j, Y_j) , заданы.

Эта задача возникает, в частности, при автоматизированном проектировании продольного профиля железных дорог. Неравенства (1) ограничивают уклоны элементов, неравенства (2) ограничивают разности уклонов смежных элементов искомой линии. Неравенства (3) совместно с (1) ограничивают длины элементов, а неравенства (4) задают предельные отклонения искомой ломаной от заданной в отдельных точках, в частности, могут задаваться фиксированные точки. X_{j+1} , X_j изменяются дискретно, например с шагом 10 м.

Ограничения не позволяют принять в качестве решения исходную ломаную $L_1(x)$.

Вместо квадратичного могут использоваться другие критерии, в частности

$$\min_{L_2} \max_x |L_2(x) - L_1(x)|.$$

2. Метод решения

Неизвестное число элементов k не позволяет использовать методы нелинейного программирования. При дополнительном задании полосы от-

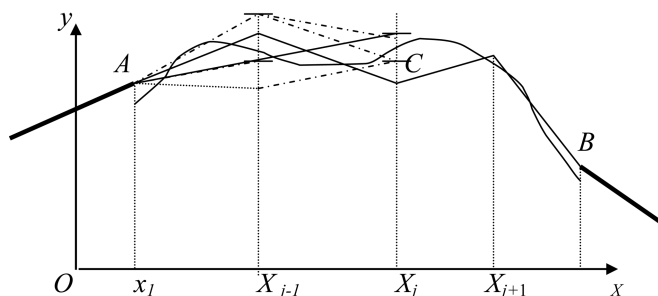


Рис. 1. Кусочно-линейная аппроксимация

клонений (зона поиска) задача решается с помощью динамического программирования.

Относительно исходной кривой в пределах зоны поиска разбивается сетка варьирования (рис. 1). Переход из заданной начальной точки A в заданную конечную точку B рассматривается как многоступенчатый процесс. На первом шаге рассматриваются все допустимые по ограничениям соединения начальной точки A и узлов сетки в пределах первого элемента $l_{\min} \leq X_2 - X_1 \leq l_{\max}$, на втором шаге — все допустимые соединения конечной точки первого элемента и узлов сетки так, чтобы $l_{\min} \leq X_3 - X_2 \leq l_{\max}$ и т. д. Если для очередного j -го элемента $x_n - X_{j+1} \leq l_{\min}$, принимается $X_{j+1} = x_n$ и $Y_{j+1} = Y_k$.

Во избежание полного перебора вводится понятие "состояние системы", и в соответствии с принципом Р. Беллмана [1] из нескольких допустимых путей, приводящих в рассматриваемое состояние, остается только один, для которого значение принятого критерия минимально.

Ввиду наличия ограничений типа (2) нельзя сравнивать варианты, сходящиеся в узле сетки. Поэтому в качестве "состояния системы" рассматривается последний из построенных элементов и сравниваются допустимые варианты достижения текущего узла, у которых последний элемент общий. При этом число элементов у сравниваемых вариантов может быть различным.

3. Алгоритм

1. Область возможных значений абсцисс концов элементов $[x_1 + l_{\min}, x_n - l_{\min}]$ разбивается на равные части с заданным дискретом d_x . Получаем массив чисел $\{X_d\}$ и соответствующих вертикалей.

2. С помощью линейной интерполяции на каждой вертикали вычисляются ординаты $\{Y_d\}$ исходной ломаной L_1 .

3. Задаются дискрет по вертикали d_y , вверх и вниз от каждой точки разбиения исходной ломаной откладывается равное число дискретов $k_d/2$.

4. Полученные точки сетки варьирования нумеруются подряд, начиная с нижней точки первой вертикали сетки и заканчивая верхней точкой последней вертикали. Запоминаются только координаты

динаты точек разбиения исходной ломаной, так как по номеру текущей точки сетки и заданным d_x , d_y и k_d легко вычислить декартовы координаты произвольной точки с номером N :

$$x_N = l_{\min} + [(N - 1)/(k_d + 1)]d_x;$$

$$y_N = y_q + (N - 1 - [(N - 1)/(k_d + 1)](k_d + 1) - k_d/2)d_y.$$

Здесь $[]$ означает целую часть, q — номер точки в массиве $\{Y_d\}$; $q = [(N - 1)/k_d] + 1$.

Все точки рассматриваются как возможные границы элементов.

5. Строятся все допустимые варианты первого элемента и запоминаются их оценки. Если задан элемент примыкания слева, то при построении первого элемента проверяется ограничение типа (2).

6. Вместо поэлементного построения траектории каждая точка при $x_1 + l_{\min} \leq x \leq x_n - l_{\min}$ рассматривается как начало следующего элемента, независимо от того, сколько элементов содержит приходящая в нее линия. При этом в цикле для каждой точки строится множество допустимых вариантов следующего элемента (рис. 2 "Выходящий веер"). При $x > x_n - 2l_{\min}$ в этом "веере" не более одного элемента с концом в точке B . "Выходящий веер" строится только из точек с ненулевым числом входящих элементов. Для каждого элемента "Выходящего веера" запоминается тот из входящих элементов, которому при выполнении ограничения (2) по сопряжению элементов соответствует минимальное значение целевой функции. Это значение суммируется с оценкой рассматриваемого элемента. Таким образом в каждой точке при $x \geq x_1 + 2l_{\min}$ формируется множество входящих элементов (см. рис. 2 "Входящий веер") и для каждого из них вычисляется оценка линии с началом в точке A и концом в данной точке. Число входящих элементов запоминается. Любое "состояние системы" задается парой целых чисел и для каждого состояния запоминается одно целое (связь) и одно вещественное число (оценка).

7. На последнем шаге при соединении с точкой B учитывается элемент примыкания справа, если он задан.

8. Обратным разворотом восстанавливается оптимальная траектория.

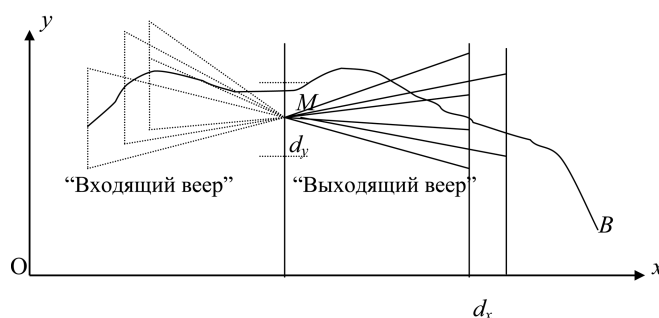


Рис. 2. Построение продолжений

4. Практическая реализация

Идея использования динамического программирования для проектирования дорог не нова [2]. Однако при ее практической реализации возникли вычислительные трудности, даже при наличии современных компьютеров, так как в реальных задачах число дискретов по вертикали более 100 и по длине 2000 и более, и трудности принципиального характера. Дело в том, что при оптимизации строительных затрат на участках пересеченного рельефа проектную линию необходимо рассматривать как единое целое [3] и динамическое программирование применить не удастся. При разработке ПО САПР задача проектирования решалась в три этапа.

1. Ограничения (3) не учитываются, абсциссы вершин исходной и проектируемой линий совпадают ($x_i = X_i, i = 1, 2, \dots, n$), ограничения (2) обеспечивают допустимую кривизну окружности, соединяющей три смежные вершины.

При известном числе элементов ($n - 1$) задача решается с помощью нелинейного программирования [3] с учетом простой структуры матрицы ограничений.

2. Полученная ломаная линия преобразуется с учетом ограничений типа (3) с целью определить число элементов. При этом отклонения в 40–50 раз меньше, чем на втором этапе, что позволяет использовать геометрические критерии близости кривых. Именно для этого преобразования использован изложенный выше алгоритм.

3. При полученном числе элементов задача решается с помощью того же алгоритма нелинейного программирования, что и на первом этапе.

Таким образом, динамическое программирование используется в САПР как метод аппроксимации, но не метод проектирования.

Экспериментально установлено, что при числе дискретов по вертикали от 40 до 100 и по длине от 2000 до 2500 существенных вычислительных трудностей не возникает.

5. Аппроксимация неоднозначных функций

Для неоднозначных функций вместо отклонений по ординатам графиков могут использоваться отклонения по нормальям к исходной кривой как при разбиении сетки поиска, так и при оценке отклонений (рис. 3).

В каждой вершине исходной ломаной вычисляется направление к центру окружности, соединяющей три смежные вершины. Это направление отождествляется с направлением нормали к исходной ломаной в рассматриваемой вершине. Вычисление отклонений очередного элемента аппроксимирующей ломаной от исходной сводится к очередному вычислению координат точек пересечения этого элемента и нормалей в тех вершинах исходной ломаной, которые захватывает этот элемент. На рис. 3 это нормали n_1, n_2, n_3 .

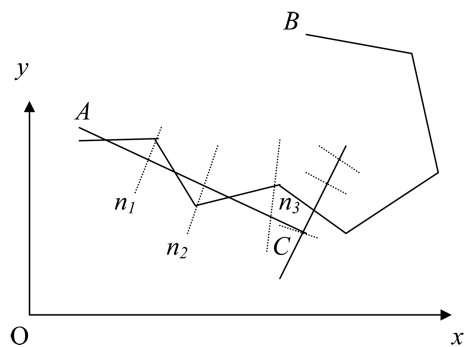


Рис. 3. Аппроксимация с использованием нормалей

Если критерием аппроксимации служит максимальное отклонение, то нормали в других точках построенного элемента вычислять не надо. Действительно, в силу линейности элементов отклонение в промежуточных точках элемента исходной ломаной не может быть больше максимального отклонения на его концах. Поэтому вычисляются расстояния точек пересечения от соответствующих вершин и берется максимальное из них (рис. 3).

Аппроксимация кривых, не являющихся графиками однозначных функций, может осуществляться и при наличии ограничений. Так, ограничения типа (4) переходят в ограничения на отклонения по нормали в отдельных точках (в частности, они могут задавать фиксированные точки), ограничения типа (1) особого смысла не имеют, а ограничения типа (2) переходят в ограничения на угол поворота в вершине искомой ломаной. В остальном в рассмотренном выше алгоритме аппроксимации однозначных функций принципиально ничего не изменяется.

Теоретически, как и при аппроксимации однозначных функций, незначительно усложняя алгоритм, можно решать задачу при дополнительном условии вписывания в углы строящейся ломаной окружностей заданного радиуса. Более того, возможен и выбор радиуса из нескольких значений. Такая задача возникает при проектировании плана трассы автомобильных дорог, магистральных трубопроводов и др. Однако в этом случае придется рассматривать существенно большие отклонения искомой линии от исходной и возникают сложности вычислительного порядка при большом числе дискретов по нормали.

Список литературы

1. Беллман Р., Дрейфус С. Прикладные задачи динамического программирования. М.: Наука, 1964.
2. Михалевич В. С., Шор Н. З. Математические основы решения задачи выбора оптимального очертания продольного профиля // Тр. Всесоюз. НИИ транспортного строительства. 1964. Вып. 51.
3. Струченков В. И. Методы оптимизации в проектировании трасс линейных сооружений // Сб. научных трудов. Искусственный интеллект в технических системах. 1999. Вып. № 20. М.: Гос. ИФТП. С. 97–111.

Ф. Р. Гальяно, аспирант,
Санкт-Петербургский институт информатики
и автоматизации Российской академии наук
e-mail: galiano@oogis.ru

Алгоритм классификации участков поверхности Земли на основе сингулярного разложения матриц

Рассматривается применение метода классификации, основанного на сингулярном разложении матриц, к задаче анализа участка земной поверхности по данным дистанционного зондирования. Предложена модификация алгоритма SVD-классификации, позволяющая увеличить ее точность за счет учета весов признаков и классов. Приводится сравнение результатов классификации данных дистанционного зондирования классическим и модифицированным алгоритмами.

Ключевые слова: дистанционное зондирование, классификация, сингулярное разложение матриц

Введение

При решении многих народнохозяйственных задач остро встает вопрос о точной и оперативной классификации участков поверхности Земли на основе данных дистанционного зондирования (ДЦЗ). При классификации исходными данными выступает набор монохромных растровых изображений, характеризующих яркость участка поверхности Земли в определенном частотном диапазоне. При этом для ряда участков заданы классы, к которым они принадлежат. Такие участки образуют обучающую выборку. Целью классификации является присвоение всем элементам (пикселям) ДЦЗ меток, определяющих класс, присвоенный соответствующему участку поверхности Земли. Класс определяется на основе вычисления сходства между свойствами классифицируемого участка и участками обучающей выборки.

В целом ряде задач для классификации применяются различные методы: искусственные нейронные сети [1]; подходы, основанные на учете энтропии [2]; правила байесовской классификации [3] и др. В том числе применяется и метод классификации на основе сингулярного разложения матриц [4], который хорошо себя зарекомендовал в части оперативной классификации. Однако ему свойственен существенный недостаток. Он проявляется в том, что метод не позволяет учитывать нелинейные зависимости между исходными данными и результатами классификации. В ряде случаев это существенно сказывается на точности классификации участков поверхности Земли.

В статье рассматривается модификация этого метода, в определенной мере лишенная отмеченного недостатка.

Постановка задачи

Рассмотрим задачу классификации исходя из следующих условий. Пусть задано отображение исходных ДЦЗ в векторы вещественных чисел одинаковой размерности. При этом векторы признаков участков земной поверхности составлены из числовых значений пикселей, обладающих одинаковыми координатами в рамках каждого растра мультиспектрального изображения (рис. 1). Под *образом* понимается n -мерный вектор-столбец $X = [x_1, \dots, x_i, \dots, x_n]^T$, где x_i — вещественные числа (*признаки* образа X) и "Т" — символ матричного транспонирования. Величины x_i отражают яркость определенных участков поверхности Земли в заданном спектральном диапазоне. Для каждого пикселя определена его окрестность путем предварительной обработки изображения по маске, применения морфологических операций и других стандартных методов предварительной обработки изображений.

Из векторов признаков может быть сформирована матрица вещественных чисел A размерностью $m \times n$. Известно, что любую матрицу, используя сингулярное разложение матрицы (англ. *Singular Value Decomposition* — SDV), можно представить в виде

$$A = \sum_{v=1}^h s_v L_v R_v^T,$$

где s_v — сингулярные числа; L_v — левые сингулярные векторы; R_v — правые сингулярные векторы; $h = \min(m, n)$. При этом слагаемые ряда упорядочены по убыванию сингулярных чисел и учтены дополнительные ограничения на элементы ряда [5]. Чем длиннее ряд, тем выше точность разложения.

Традиционный процесс SVD-классификации [4] состоит их этапов обучения и распознавания. Из

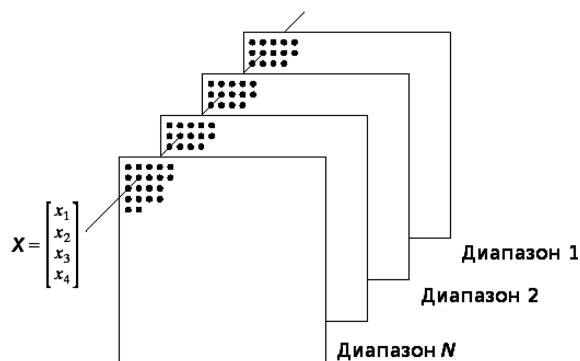


Рис. 1. Формирование вектора признаков для классификации средствами SVD-классификатора

векторов признаков формируется матрица A , для которой вычисляются первые g членов сингулярного разложения, $g \leq h$. Таким образом происходит существенное снижение размерности и объема данных. Полученные в результате наборы троек, состоящие из сингулярного числа и соответствующих ему левого, а также правого сингулярных векторов, есть результат обучения. Число таких троек определяет размерность нового пространства признаков. Процесс распознавания сводится к отображению классифицируемого вектора в новое пространство признаков и поиска наиболее близкого ему элемента.

Для учета нелинейных зависимостей между исходными данными и возможными классами предлагается учитывать веса признаков и классов. Требуется разработать модифицированный алгоритм SDV-классификации, позволяющий учитывать указанные веса.

Математическая формулировка задачи

Определим распознавание образов как отображение $\{X\} \rightarrow \{C\}$, где $\{X\}$ — множество всех образов, $\{C\}$ — множество всех классов, заданных их метками $c_1, \dots, c_u$.

Дано:

1. Набор из m обучающих образов $X_j, j = 1, \dots, m$.
2. Для каждого признака каждого обучающего образа задан интервал допустимых значений $x_i \in [a_i; b_i]$, $a_i \in R, b_i \in R, b_i > a_i, i = 1, \dots, n$. Этот интервал считается общим для всех признаков с одинаковым номером по всем обучающим образам X_j .
3. Для каждого обучающего образа X_j — соответствующий ему элемент множества классов $\{C_t\}_{t=1}^u$;
4. Для каждого класса обучающего образа — его вес $\beta_t \in [0; 1], t = 1, \dots, u$.
5. Для каждого признака всех обучающих образов — его вес $\gamma_i \in [0; 1], i = 1, \dots, n$.
6. Произвольный n -мерный вектор Z .

Необходимо найти: класс вектора Z с учетом заданных весов признаков и классов.

Алгоритм обучения

Алгоритм обучения системы классификации для решения указанной задачи может быть представлен в виде последовательности шагов.

1. Масштабирование интервала допустимых значений всех признаков $x_i, i = 1, \dots, n$, на общую для всех признаков шкалу $[0; b_i]$ с использованием следующей формулы¹:

$$x_i^* = \frac{(x_i - a_i)b_i^*}{b_i - a_i}; \quad x_i^*; \quad b_i^* = \max_i b_i,$$

где x^* — значение нормированного признака.

¹ Как правило, диапазон значений яркостей анализируемых растров одинаков; в таких случаях этот этап можно опустить.

2. Умножение каждого признака на его вес γ_i .
3. Формирование обучающей матрицы $A = [X_1 \dots X_m]^T$ размерности $m \times n$.

4. Вычисление максимального сингулярного числа s , а также левого и правого сингулярных векторов L и R обучающей матрицы по следующей итеративной схеме:

$$L_0 = [1, \dots, 1]^T;$$

$$R^T = L_{k-1}^T A; \quad R_k = \frac{R}{|R|};$$

$$L = AR_k; \quad L_k = \frac{L}{|L|}; \quad s_k = L_k^T AR_k,$$

где $k = 1, 2, \dots$,

$$|L| = \sqrt{l_1^2 + \dots + l_m^2};$$

$$|R| = \sqrt{r_1^2 + \dots + r_n^2}.$$

Критерием остановки служит выполнение условия $|s_k - s_{k-1}| < \varepsilon$, где ε — погрешность вычисления сингулярного разложения, задаваемая экспертом с учетом диапазона изменения входных данных.

5. Считать $s = s_k, L = L_k, R = R_k$.

6. Если заданная экспертом размерность нового пространства признаков $g \geq 1$, то слагаемые ряда сингулярного разложения, отличные от первого, могут быть вычислены, в частности, методом исчерпывания [4]. Векторы L_v и $R_v, v = 2, \dots, h$, вычисляются как первые левый и правый сингулярные векторы матрицы, полученной вычитанием из матрицы A ее приближения $v - 1$ членами вышеуказанного ряда.

7. Хранить первые g сингулярных чисел s и соответствующие им левые и правые сингулярные векторы R и L .

Алгоритм распознавания

Распознавание вектора признаков предлагается осуществлять в соответствии со следующим алгоритмом.

1. Масштабирование значения всех признаков образа Z и умножение их на вес согласно пп. 1 и 2 этапа обучения.

2. Для введенного вектора P — вычисление $w_v(P)$ (координат проекции вектора в новое пространство признаков):

$$w_v(P) = \frac{P^T R_v}{s_v}; \quad v = 1, \dots, g.$$

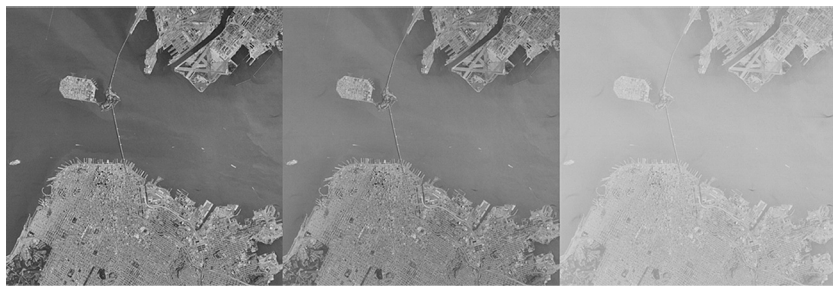


Рис. 2. Слева направо: три спектральных диапазона исходных данных — красный, зеленый и синий соответственно

3. Для каждого $j = 1, \dots, m$ вычисление² расстояния $P_j = \sqrt{\sum_{v=1}^g (w_v - [l_j]_v)^2}$, где $[l_j]_v$ — j -я компонента левого сингулярного вектора L_v .

4. Для всех $j = 1, \dots, m$ нормирование значений P_j :

$$d_j = \frac{P_j}{M};$$

$$M = \max_j P_j.$$

5. Поиск j^* среди всех $j = 1, \dots, m$ индекса, при котором величина $d_j \beta_t$ минимальна. Здесь β_t — вес класса, соответствующего обучающему образу X_j .

6. Считать класс образа X_j искомым классом образа Z .

Особенность предлагаемых алгоритмов в том, что они включают в себя новые правила преобразования исходных данных (этапы 1 и 2 алгоритма распознавания) и дополнительные правила учета весов классов (этапы 4 и 5 алгоритма распознавания). При этом за счет варьирования весов классов можно изменять соотношение ошибок первого и второго рода и повышать точность классификации, как показано ниже.

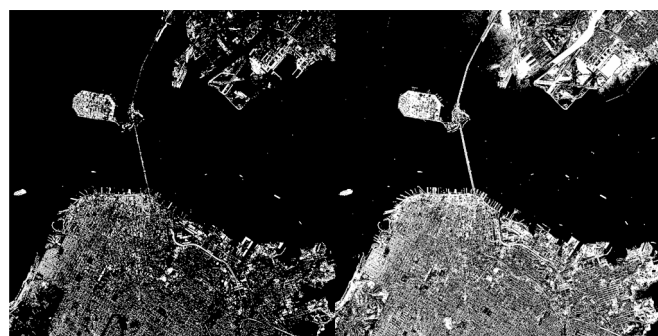


Рис. 3. Результаты классификации с единичными весами классов (слева) и с весами классов, равными 1 и 0,25 (справа). Метками классов выступают яркости пикселей (0 для водной поверхности и 255 для суши)

² В данном случае использование евклидова расстояния необязательно; допустимо использование других норм, если это приводит к улучшению результатов классификации.

Пример анализа данных дистанционного зондирования

В качестве тестовых данных использовались мультиспектральные изображения поверхности Земли³. Целью классификации являлось разделение пикселей на два класса: водной поверхности и суши. Исходное цветное изображение было преобразовано в три одноцветных изображения в цветовом пространстве RGB (рис. 2). Векторы признаков размерностью 3 составлялись непосредственно из яркостей пикселей с одинаковыми координатами. Результаты сравнения классификации с единичными весами классов β_t и с весами 1 и 0,25 для воды и суши соответственно приведены на рис 3. Анализ этих результатов показывает, что учет весов позволяет снизить число ошибок. Так, при традиционном алгоритме классификации процент неправильных решений составил 32,34 %, а при модифицированном методе всего лишь 20,79 %. Таким образом, полученные результаты позволяют сделать вывод о повышении и расширении возможностей метода классификации участков поверхности Земли на основе SVD за счет учета весовых коэффициентов.

Выводы

- ♦ Предложен усовершенствованный алгоритм SVD-классификации, позволяющий повысить точность классификации. Он позволяет компенсировать погрешности при формировании обучающих выборок и варьировать соотношение ошибок первого и второго рода при классификации.
- ♦ Полученные результаты могут найти применение при классификации данных большой размерности, характеризующихся нелинейными зависимостями между исходными данными и результатом классификации. Дальнейшее улучшение результатов может быть получено при классификации не только отдельных пикселей, но и более сложных по форме участков земной поверхности, для чего требуется разработка алгоритмов сегментации изображения и описания полученных совокупностей сегментов.

Список литературы

1. Хайкин С. Нейронные сети: полный курс. М.: Вильямс, 2008.
2. Chi-hau C. Image processing for remote sensing. Boca Raton, CRC Press, 2008.
3. Потапов А. С. Распознавание образов и машинное восприятие. СПб.: Политехника, 2007.
4. Tarakanov A. O., Skormin V. A., Sokolova S. P. Immunocomputing: Principles and applications. New York: Springer, 2003.
5. Голуб Дж., Ван Лоун Ч. Матричные вычисления. М.: Мир, 1999.

³ Источник: база изображений USC-SIPI, <http://sipi.usc.edu/database/index.html>.

УДК 004.89 + 004.021

П. В. Казаков, канд. техн. наук, доц.,
Брянский государственный
технический университет,
e-mail: pvk_mail@list.ru

Автоматизация синтеза оптимальных инвестиционных портфелей на основе кластерного генетического алгоритма

Рассматривается новый способ синтеза оптимальных инвестиционных портфелей с помощью кластерного генетического алгоритма. Его применение позволяет эффективно решать данную задачу как многоэкстремальную с учетом прогнозируемых доходности и риска инвестиционного планирования. Представлены результаты практического применения кластерного генетического алгоритма для определения оптимального инвестиционного портфеля на основе фактических данных.

Ключевые слова: инвестиционный портфель, эволюционное моделирование, кластерный генетический алгоритм, многоэкстремальная оптимизация

Введение

Важной задачей экономики и бизнеса, которая часто возникает на практике, является оптимизация инвестиционного планирования средств для их размещения в ценных бумагах. Частным случаем, с которым сталкиваются как организации, так и физические лица, занимающие активную позицию в стремлении увеличить свои капиталы, является формирование инвестиционного портфеля (ИП) [1–4]. Его целью является повышение качества инвестирования в виде надежного сбережения капитала или получения максимального дохода при допустимом риске. При этом предполагается наличие у инвестора определенной суммы денег, которую необходимо распределить по различным финансовым активам (акциям, облигациям, иностранной валюте и др.), вместе обладающими такими инвестиционными характеристиками, которые недостижимы с позиций отдельно взятой ценной бумаги.

При осуществлении портфельного инвестирования возникает проблема оптимального распределения средств, чтобы доходность портфеля была максимальной, а риск убытков минимальным. В зависимости от инвестиционных свойств ценных бумаг могут быть сформированы различные инвестиционные портфели, в каждом из которых будет собственный баланс между существующим риском, приемлемым для владельца портфеля, и ожидаемой доходностью в определенный период времени. Известно, что добавление в портфель различных, слабокоррелирующих ценных бумаг снижает риск инвестиций [2, 3]. Портфель, содержащий различные виды ценных бумаг, которые имеют низкий коэффициент корреляции, называется диверсифицированным. Важным является определение количества различных ценных бумаг в таком портфеле, а именно степень его диверсификации. Установлено, что максимальное снижение риска достигается, если в портфель отобраны от 10 до 15 различных ценных бумаг [2].

В общем случае модель оптимизации ИП предполагает наличие множества Парето-оптимальных решений при оценке соотношения "доходность—риск", расположенных на так называемой границе эффективности инвестиционных портфелей [3]. Данная модель предложена Г. Марковицем [1, 3, 4] и является основной в теории инвестиционного портфеля. Однако эта модель характеризуется высокой вычислительной трудоемкостью в плане применения численных методов оптимизации. Ввиду наличия комбинаторных свойств задача оптимизации ИП на основе модели Г. Марковица относится к *NP*-сложным: варьируя доли финансовых активов в ИП, можно сформировать их бесконечное множество с собственным балансом между ожидаемой доходностью и риском. В связи с этим данную модель обычно применяют с априорным заданием экспертом значений доходности или риска. В этом случае решением задачи является ИП, соответствующий максимальной доходности при заданном риске, либо ИП с минимальным риском при заданной доходности. Также используются другие модели оптимизации и выбора ИП, в частности модели У. Шарпа и Д. Тобина [3, 4]. Их особенностью является наличие определенных упрощений относительно модели Г. Марковица,

прежде всего для снижения объемов вычислений. Так, модель У. Шарпа предполагает линейный регрессионный анализ главных параметров и критериев задачи оптимизации ИП, а модель Д. Тобина — обязательное наличие в ИП доли из безрисковых активов. Таким образом, задача оптимизации ИП может быть решена с применением традиционных численных методов за ограниченное время. Однако потребность в дополнительных исходных данных от эксперта и определении значений ряда дополнительных параметров в моделях У. Шарпа и Д. Тобина усложняет их адаптацию к реальным инвестиционным условиям, а вынужденные упрощения в них снижают точность полученного результата. Поэтому развитие компьютерных методов моделирования оптимального размещения капитала в рискованные активы в условиях изменяющейся доходности остается актуальной задачей.

В настоящей статье рассматривается новый способ ее решения, основанный на модели Г. Марковица, но с некоторой модификацией, предполагающей многоэкстремальную оптимизацию по критерию "доходность", с последующим вычислением риска для найденных ИП (возможен и обратный вариант, когда многоэкстремальная оптимизация выполняется по риску, а затем вычисляется доходность ИП). Появление множества оптимальных решений связано не только с тем, что возможно более одного варианта состава ИП с максимальной доходностью. Также к набору оптимальных ИП могут быть добавлены ИП, которые имеют доходность несколько ниже (определяется экспертом), но в сочетании с меньшим риском, чем у лучших по доходности ИП. Здесь принимается во внимание тот факт, что окончательный выбор ИП будет проводиться экспертом одновременно по двум критериям: "доходность", "риск". Возможности существующих традиционных численных методов для многоэкстремальной комбинаторной оптимизации в многомерных поисковых пространствах являются очень ограниченными как по точности, так и по времени поиска. Поэтому для автоматизации поиска вариантов оптимальных ИП применялся кластерный генетический алгоритм [5], показавший высокую эффективность решения тестовых задач многоэкстремальной оптимизации.

1. Постановка и математическая модель задачи оптимизации инвестиционного портфеля

Наиболее общая математическая модель оптимизации инвестиционного портфеля предложена американским экономистом Г. Марковицем. Следуя целям автоматизации решения данной задачи,

а также учитывая доступные фактические данные по финансовым активам, предлагаемому на рынке ценных бумаг, эта модель была без потери общности скорректирована и имеет следующий вид:

$$f = \max \sum_{i=1}^n \bar{\gamma}_i \frac{\alpha_i x_i}{\sum_{i=1}^n \alpha_i x_i}, \quad \sum_{i=1}^n x_i \alpha_i \leq U,$$

где U — объем финансовых средств для формирования инвестиционного портфеля; α_i — начальная стоимость одной единицы ценных бумаг вида i ; x_i — объем приобретаемых финансовых активов вида i ; $x_i \in [x_i^-, x_i^+]$ — нижняя и верхняя границы объемов ценной бумаги вида i ; $\bar{\gamma}_i$ — ожидаемая

средняя доходность по i -му активу; $\frac{\alpha_i x_i}{\sum_{i=1}^n \alpha_i x_i}$ —

удельный вес стоимости i -го актива в ИП; n — число различных финансовых активов в ИП (далее $n = 15$); f — оптимизируемая функция доходности ИП с оптимальным значением f^* .

Значения $\bar{\gamma}_i$ рассчитываются на основе известного диапазона колебаний доходности финансового актива i за k периодов наблюдений (например, по факту проведения торгов на фондовой бирже) от минимального γ_i^- до максимального γ_i^+ . При этом экспертом учитываются вероятности получения граничных и промежуточных значений доходности

$$P(\gamma_{i,j}): \bar{\gamma}_i = \sum_{j=1}^k \gamma_{i,j} P(\gamma_{i,j}),$$

где $\gamma_i^- \leq \gamma_{i,j} \leq \gamma_i^+$ — доходность актива i за период j ; $\sum_{j=1}^k P(\gamma_{i,j}) = 1$.

При формировании ИП, кроме доходности, также учитывается уровень его риска как в целом, так и по отдельным активам.

Риск (дисперсия) портфеля из n активов рассчитывается по следующей зависимости [1]:

$$\sigma_R^2 = \sum_{i=1}^n \sum_{j=1}^n y_i y_j C_{i,j},$$

где $C_{i,j}$ — ковариация доходности активов, входящих в ИП; y_i, y_j — удельные веса активов в общей стоимости ИП.

В свою очередь, показатель ковариации между двумя активами A , B определяется по формуле [1]

$$C_{A,B} = \frac{\sum_{j=1}^k (\gamma_j^A - \bar{\gamma}^A)(\gamma_j^B - \bar{\gamma}^B)}{k},$$

где γ_j^A , γ_j^B — доходности активов A , B в j -м периоде; $\bar{\gamma}^A$, $\bar{\gamma}^B$ — средняя доходность активов A , B за k периодов.

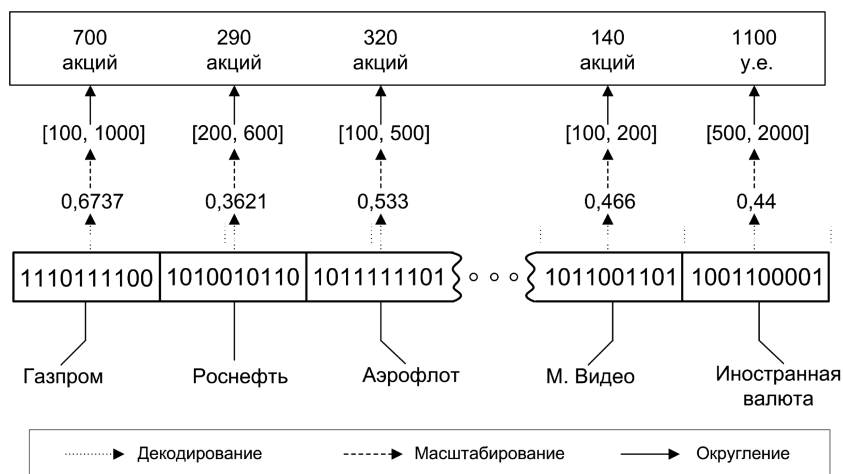
Особенностью данной математической модели является отсутствие необходимости изначального задания инвестором главных характеристик ИП — значений желаемой доходности и допустимого риска, что далеко не всегда можно сделать ввиду стохастически изменяющихся инвестиционных возможностей. В итоге это снижает психологическую нагрузку ответственности на эксперта, позволяя ему сделать окончательный выбор по соотношению "доходность—риск" уже после определения оптимальных ИП. Также важно снижение вычислительной сложности поиска решений благодаря разделению процесса поиска по критериям оптимизации.

2. Эволюционная модель задачи оптимизации инвестиционного портфеля

При создании эволюционной модели вначале необходимо выбрать форму кодирования значений переменных математической модели задачи, а также сконструировать *fitness*-функцию [6, 7].

Каждый вариант инвестиционного портфеля кодируется отдельной хромосомой из 15 двоичных генов фиксированной разрядности (см. рисунок). Декодированное значение гена представляет собой нормализованное в интервале $[0, 1]$ вещественное число. Поэтому при определении реального числа ценных бумаг актива i выполняется масштабирование в соответствующий этому активу интервал $[x_i^-, x_i^+]$ с последующим округлением до целого.

Для определения оптимальности хромосомы была сконструирована *fitness*-функция, учитывающая специфику решаемой задачи, а именно: получение максимальной итоговой доходности по ИП, желание лица, принимающего решение (ЛПР), вложить в ИП всю сумму или сохранить некоторый остаток денежных средств после формирования ИП, а также наличие или отсутствие контроля за превышением объема выделенных средств для формирования ИП. Последнее предполагает возможность увеличения начальных размеров капиталовложений для приобретения дополнитель-



Закодированный в виде хромосомы вариант инвестиционного портфеля

ного числа ценных бумаг по перспективному, с точки зрения доходности, активу. В результате *fitness*-функция получила следующий вид: *fitness-function*(FF) = $f + R$. Здесь первое слагаемое представляет собой исходную целевую функцию математической модели задачи, второе — функцию штрафа за нарушение ограничений математической модели. Для представления штрафной составляющей может быть использовано выражение следующего вида:

$$R = -\frac{2|V(C_i) - U| + (V(C_i) - U)}{2},$$

где $V(C_i)$ — объем финансовых средств, затраченных на формирование инвестиционного портфеля, соответствующего хромосоме C_i .

При $V(C_i) = U$ значение коэффициента штрафа $R = 0$, если $V(C_i) > U$, то назначается двойной штраф, так как нарушение этого ограничения является несовместимым с исходной математической моделью задачи. Если же $V(C_i) < U$, то R присваивается значение штрафа за неполное расходование средств на формирование ИП. Следствием этого может стать невозможность получения самого лучшего ИП, хотя и с сохранением для потребителя ИП некоторых исходных денежных средств.

В итоге окончательная форма *fitness*-функции записывается следующим образом:

$$FF = \max \left(\sum_{i=1}^n \bar{\gamma}_i \frac{\alpha_i x_i}{\sum_{i=1}^n \alpha_i x_i} - \frac{2|V(C_i) - U| + (V(C_i) - U)}{2} \right).$$

Оптимизация описанной математической модели предполагает получение множества оптимальных решений. Для их поиска применялось кластерное расширение генетического алгоритма (КГА) [5]. КГА реализует в себе принципы под-

держки разнообразия популяции в процессе генетического поиска за счет кластеризации хромосом с точки зрения близости кодируемых ими решений. Кластеры моделируются отдельной подпопуляцией, эволюционирующей параллельно с основной и содержащей "похожие" хромосомы. Кластеры определяются на основе меры близости фенотипов, что отличает КГА от существующих нишевых надстроек стандартного генетического алгоритма [6, 7], обычно использующих меру близости на основе приспособленности. Кластеры определяются для каждого поколения и по особым принципам копируются в следующее поколение [5].

3. Результаты применения кластерного генетического алгоритма для синтеза оптимальных инвестиционных портфелей

В качестве тестового примера далее будут использоваться следующие входные данные математической модели оптимизации ИП (табл. 1).

Предварительно, на основе анализа динамики за определенный период индексов и котировок ценных бумаг, выбирались активы-кандидаты в инвестиционный портфель. В данном случае про-

верялось выполнение условия наличия положительной доходности активов в течение одного месяца, а также учитывался коэффициент корреляции доходности активов. Для оценки риска брали результаты торгов ценными бумагами в течении трех месяцев. Представленные в табл. 1 значения параметров ИП являются фактическими на определенный период мониторинга рынка ценных бумаг [8] и при необходимости могут быть скорректированы под новые инвестиционные требования.

В целях получения наилучших результатов при поиске решения в сочетании с высоким быстродействием выполнялась настройка КГА. В табл. 2 представлены различные его тестовые конфигурации. Выбор наборов значений управляющих параметров КГА мотивируется случаями наиболее различного поведения алгоритма (динамика изменения лучшего, среднего, худшего решений) в тестовых запусках. Характеристики $\bar{f}^*$ и $\bar{N}_F$ соответствуют усредненным по 10 циклам стартов КГА на каждую конфигурацию соответственно значениям оптимального решения и числа вычислений *fitness*-функции, достаточного для локализации оптимальных решений. К ним относились хромосомы, у которых значение *fitness*-функции

Таблица 1

Параметры инвестиционного портфеля

№	Название актива	Начальная стоимость актива, руб.	Ожидаемая доходность, %	Минимальное число ценных бумаг в портфеле	Максимальное число ценных бумаг в портфеле
1	Газпром	104	7	100	1000
2	Роснефть	108	8	200	600
3	Аэрофлот	24	3	100	500
4	Объединенные машиностроительные заводы	150	12	100	500
5	Заволжский моторный завод	62	15	100	500
6	Новолипецкий металлургический комбинат	30	5	200	700
7	Приволжское морское пароходство	5	15	100	500
8	Брянская сбытовая компания	12	10	300	700
9	Калужская сбытовая компания	13	10	200	600
10	Мобильные ТелеСистемы	114	5	200	500
11	РБК информационные системы	13	12	300	800
12	ДИКСИ	56	11	100	500
13	Сбербанк	18	5	100	600
14	М.Видео	26	6	100	200
15	Иностранная валюта (€)	39	7	500	2000

Таблица 2

Результаты настройки кластерного генетического алгоритма

R_c	P_c , %	P_m , %	$\bar{f}^*$, %	Число оптимальных решений	Число вычислений <i>fitness</i> -функции
0,5	0,9	0,005	9,45	4	49375
		0,01	9,15	4	49732
		0,03	8,94	5	50243
	0,8	0,005	9,3	4	49454
		0,01	9,15	6	50115
		0,03	8,9	5	51326
0,3	0,7	0,005	8,65	6	51137
		0,01	8,46	7	51482
		0,03	8,3	5	51923
	0,9	0,005	9,45	9	50131
		0,01	9,45	6	50314
		0,03	8,9	7	51225
	0,8	0,005	9,35	6	52043
		0,01	9,12	6	52614
		0,03	8,87	5	53165
0,1	0,7	0,005	8,94	8	52076
		0,01	8,84	8	52622
		0,03	8,7	7	53132
	0,9	0,005	8,4	11	54125
		0,01	8,35	9	54511
		0,03	8,2	9	54822
	0,8	0,005	8,44	9	54133
		0,01	8,37	10	55014
		0,03	8,27	9	55422
0,7	0,9	0,005	8,4	8	55353
		0,01	8,27	7	56254
		0,03	8,2	7	56863

Варианты оптимальных инвестиционных портфелей

№ актива	ИП1	ИП2	ИП3	ИП4	ИП5	ИП6	ИП7	ИП8	ИП9
1	280	280	280	280	110	280	110	280	280
2	570	500	500	500	510	500	510	500	500
3	130	130	130	130	120	130	130	130	130
4	390	390	390	390	400	390	390	390	390
5	460	470	460	470	250	470	410	460	190
6	260	260	260	260	610	260	600	260	260
7	490	490	490	490	490	490	490	490	490
8	360	640	360	640	630	640	660	360	640
9	550	320	550	320	560	340	280	560	550
10	200	200	200	200	200	210	200	200	200
11	790	740	740	740	740	750	790	740	740
12	160	500	500	500	490	160	160	220	470
13	120	120	170	170	100	300	170	420	120
14	100	100	110	180	180	130	180	150	100
15	510	620	580	530	530	590	630	590	640
Доходность, %	9,55	9,54	9,53	9,52	9,45	9,44	9,44	9,44	9,43
Риск, %	12,39	12,2	11,65	11,58	11,74	10,37	11,62	11,27	12,1

было не более чем на 0,2 % хуже, чем у самой лучшей хромосомы в последней популяции. Для каждой конфигурации задавалось максимальное число поколений $N_g = 300$ и объем капиталовложений $U = 700\,000$ руб.

Анализ результатов настройки КГА позволяет сделать вывод о локализации различного числа оптимальных решений в зависимости от конфигурации КГА. Большое влияние на эффективность поиска оказало соотношение параметров N_p (размер популяции) и R_c (радиус кластера), важным является определение таких их значений, при которых сохранялся бы баланс между разнообразием популяции и направленным характером генетического поиска. Максимальное число найденных оптимальных ИП составило девять портфелей и было получено для следующих значений управляющих параметров: $N_g = 300$, $N_p = 200$, вероятность Кроссинговера $P_c = 0,9$, вероятность мутации $P_m = 0,005$, $R_c = 0,3$. Для полученных таким образом инвестиционных портфелей далее были определены возможные риски (табл. 3).

Анализ найденных вариантов ИП позволяет сделать вывод, что не один из них не является оптимальным одновременно по доходности и риску. Поэтому полученные результаты далее должны быть предъявлены эксперту, который, учтя вероятности рисков для каждого из найденных ИП, а также дополнительные условия инвестиционного планирования (пассивная или активная стратегия управления ценными бумагами), выработает окончательную рекомендацию для потребителя инвестиционного портфеля. Так, в табл. 3 выделен состав ИП, обеспечивающий приемлемую (относительно других вариантов) доходность при минимальном риске. В случае возникновения ситуации, когда эксперта не устроят уровни риска

по найденным ИП с оптимальной доходностью, предельные значения доходности могут быть уменьшены в пользу снижения итогового риска ИП.

Заключение

Текущее состояние финансового рынка требует быстрого и адекватного реагирования на его изменения. Поэтому особенно важно эффективное управление инвестиционным планированием. Применение кластерного генетического алгоритма для решения этой задачи позволило автоматизировать процесс определения различных оптимальных составов инвестиционных портфелей в сочетании с минимальным количеством дополнительных экспертной информации и данных для ее (задачи) математической модели. Это дает возможность эксперту быстро выбирать между различными наборами финансовых активов, достигая в итоге наилучшего сочетания между их доходностью и риском.

Список литературы

1. Серов В. М. Инвестиционный менеджмент. М.: ИНФРА-М, 2000. 272 с.
2. Мищенко А. В., Попов А. А. Некоторые подходы к оптимизации инвестиционного портфеля // Менеджмент в России и за рубежом. 2002. № 2.
3. Шапки А. С. Экономические и финансовые риски. Оценка, управление, портфель инвестиций. М.: Дашков и К°, 2003. 544 с.
4. Шарп У., Александер Г., Бейли Д. Инвестиционный менеджмент / пер. с англ. М.: ИНФРА-М, 2003. 257 с.
5. Казаков П. В. Кластерное расширение генетического алгоритма для решения многоэкстремальных задач оптимизации // Информационные технологии. 2009. № 8. С. 33–38.
6. Гладков Л. А., Курейчик В. В., Курейчик В. М. Генетические алгоритмы. М.: ФИЗМАТЛИТ, 2006. 320 с.
7. Аверченков В. И., Казаков П. В. Эволюционное моделирование и его применение. Брянск: Изд. БГТУ, 2009. 200 с.
8. Социальная сеть инвесторов [Электронный ресурс]. — URL: <http://tikr.ru>.

И. А. Чучуева, аспирант,
МГТУ им Н. Э. Баумана,
e-mail: chuchueva@yandex.ru

Модель экстраполяции временных рядов по выборке максимального подобия

Предложена новая модель экстраполяции временных рядов по выборке максимального подобия. Приводится описание модели, а также алгоритм решения задачи ее идентификации. Оценка эффективности модели экстраполяции проведена на временных рядах энергопотребления и цен на электроэнергию.

Ключевые слова: прогнозирование, экстраполяция, потребление, цена РСВ, временные ряды, оптовый рынок электроэнергии и мощности

Введение

Задача прогнозирования будущих значений временного ряда на основании его текущих и прошлых значений является основой для планирования в экономике и торговле, планирования выпуска продукции, складского контроля, управления и оптимизации промышленных процессов и т. д. [1]. В настоящее время методы прогнозирования продолжают интенсивно развиваться параллельно с развитием систем хранения данных и вычислительных мощностей.

Методы прогнозирования можно разделить на интуитивные и формализованные [2]. Согласно обзору [3] формализованные методы делятся на две категории: статистические и методы искусственного интеллекта. В статистических методах для вычисления прогнозируемого значения временного ряда определяется некоторое уравнение, полученное на основании изучения истории временного ряда и отражающее зависимость будущих значений ряда от его истории и/или внешних факторов. Методы искусственного интеллекта имитируют процесс мышления человека в изучении истории и предсказании будущих значений ряда. В статье рассматривается статистический метод прогнозирования временных рядов.

Математическая зависимость значений временного ряда за пределами его фактических значений от этих фактических значений называется *моделью экстраполяции* [4]. В настоящей работе предлагается новая модель экстраполяции временных рядов на основе *выборки максимального подобия*.

Исследование эффективности предложенной модели экстраполяции выполнено на временных рядах цен рынка на сутки вперед оптового рынка электроэнергии и мощности (далее — цен РСВ) и энергопотребления того же рынка (далее — энергопотребления). Экстраполированные значения энергопотребления и цен РСВ необходимы организациям, регулирующим энергосистему, для управления ее техническими и экономическими параметрами. Кроме того, экстраполированные значения энергопотребления и цен РСВ помогают участникам энергорынка повышать свои финансовые результаты [3, 5–7].

Заметим, что задача прогнозирования цен РСВ для России является новой. Особенность ее состоит в том, что по мере реформирования энергорынка России алгоритм расчета цен РСВ подвергается изменениям. Цены РСВ до 01.01.2008 рассчитывались по одному алгоритму, затем алгоритм был изменен. Важность прогнозирования временного ряда цен РСВ для участников рынков электроэнергии отмечается, например, в работе [3].

Задача прогнозирования энергопотребления имеет множество специализированных моделей, созданных за последние годы [8–10], однако исследования в данной области продолжаются.

Статья содержит описание предлагаемой модели экстраполяции (раздел 1), а также алгоритм решения задачи идентификации модели (раздел 2). Далее в статье приведены результаты, подтверждающие эффективность предложенной модели (раздел 3). В заключении указаны особенности предложенной модели.

1. Модель экстраполяции

В работе речь идет о дискретных временных рядах, значения которых получены в моменты времени $t_1, t_2, t_3, \dots, t_N$. Указанные моменты времени могут быть неравноотстоящими.

Временной ряд $Z = z(t_1), z(t_2), z(t_3), \dots, z(t_N)$ обозначим $Z_1^N = z_1, z_2, z_3, \dots, z_N$. Набор последовательных значений $Z_t^M = z_t, z_{t+1}, z_{t+2}, \dots, z_{t+M-1}$, лежащих внутри временного ряда Z_1^N , назовем *выборкой* из этого ряда длины M с моментом начала отсчета t ; $M \in [1, N-1]$; $t \in [1, N-M]$. Разность начал отсчетов выборок Z_t^M, Z_{t-k}^M назовем *задержкой* k , $k \in [1, t-1]$.

Коэффициент корреляции выборок Z_t^M, Z_{t-k}^M определим как

$$\rho_k = \frac{\text{cov}(Z_t^M, Z_{t-k}^M)}{\sqrt{D[Z_t^M]} \sqrt{D[Z_{t-k}^M]}}, \quad (1)$$

где $\text{cov}(Z_t^M, Z_{t-k}^M)$ — ковариация исходных выборок, а $D[Z_t^M]$ и $D[Z_{t-k}^M]$ — их дисперсии [1].

На основе коэффициента корреляции (1) введем меру подобия выборок Z_t^M, Z_{t-k}^M

$$\rho_k^M = \frac{\text{cov}(Z_t^M, Z_{t-k}^M)}{\sqrt{D[Z_t^M]} \sqrt{D[Z_{t-k}^M]}} \in [0, 1].$$

Величина ρ_k^M зависит от длины M выборок Z_t^M, Z_{t-k}^M , а также от задержки k . Мера подобия ρ_k^M характеризует степень линейной зависимости выборок Z_t^M и Z_{t-k}^M : чем ближе к единице значение ρ_k^M , тем выше их линейная зависимость.

Определим для выборки Z_t^M величины $\rho_1^M, \rho_2^M, \dots, \rho_{t-1}^M$ и найдем их максимальное значение

$$\rho_{k\max}^M = \max(\rho_1^M, \rho_2^M, \dots, \rho_{t-1}^M).$$

Выборку, соответствующую задержке $k_{\max}$, обозначим $Z_{t-k\max}^M$ и назовем *выборкой максимального подобия* для выборки Z_t^M . Легко видеть, что для выборки $Z_{t-k\max}^M$ справедливо равенство

$$\rho_{k\max}^M = \frac{\text{cov}(Z_t^M, Z_{t-k\max}^M)}{\sqrt{D[Z_t^M]} \sqrt{D[Z_{t-k\max}^M]}}.$$

Гипотеза подобия. Если выборки $Z_t^M, Z_{t-k\max}^M$ имеют значение величины $\rho_{k\max}^M$, близкое к единице, то для некоторых значений P и выборок $Z_t^{M+P}, Z_{t-k\max}^{M+P}$ значение величины $\rho_{k\max}^{M+P}$ также близко к единице.

Представленные в разделе 3 результаты прогнозирования подтверждают справедливость гипотезы для временных рядов энергопотребления и цен РСВ. Для временных рядов из других пред-

метных областей справедливость гипотезы необходимо проверять.

Перейдем к векторному обозначению выборки $Z_t^M = (z_t, z_{t+1}, \dots, z_{t+M-1})^T$ и временного ряда $Z_1^N = (z_1, z_2, \dots, z_N)^T$.

Аппроксимируем выборку Z_t^M с помощью выборки $Z_{t-k\max}^M$:

$$\begin{aligned} Z_t^M &= a_1 Z_{t-k\max}^M + a_0 + E^M; \\ \check{Z}_t^M &= a_1 \check{Z}_{t-k\max}^M + a_0. \end{aligned} \quad (2)$$

Здесь a_1 и a_0 — вещественные константы; E^M — M -мерный вектор значений ошибок аппроксимации; $\check{Z}_t^M$ — аппроксимированные значения Z_t^M . Тогда выборку Z_{N+1}^P можно выразить через некоторую выборку Z_τ^P , лежащую внутри исходного временного ряда Z_1^N , в виде

$$\check{Z}_{N+1}^P = a_1 \cdot Z_\tau^P + a_0. \quad (3)$$

Алгоритм определения выборки Z_τ^P имеет следующий вид:

- для выборки Z_{N-M+1}^M находим выборку максимального подобия $Z_{k\max}^M$, где $k\max^* = N - M + 1 - k\max$;
- согласно (2) аппроксимируем выборку Z_{N-M+1}^M с помощью выборки $Z_{k\max}^M$:

$$\check{Z}_{N-M+1}^M = a_1 \cdot Z_{k\max}^M + a_0; \quad (4)$$

- в соответствии с гипотезой подобия в качестве выборки Z_τ^P используем выборку $Z_\tau^P = Z_{k\max^*+M}^P$, т. е. выборку, расположенную на оси времени сразу за выборкой максимального подобия (рис. 1).

В формуле (4) вещественные константы a_1 и a_0 определяются путем решения методом наименьших квадратов задачи

$$S = \sum_{i=1}^M (\check{Z}_{N-M+1-i}^M - Z_{N-M+1-i}^P)^2 \rightarrow \min,$$

подробно рассмотренной в работе [11].

Таким образом, экстраполированные значения $\check{Z}_{N+1}^P$ временного ряда Z_1^N определяются формулой

$$\begin{aligned}\check{Z}_{N+1}^P &= a_1 \cdot Z_{k_{\max}^* + M}^P + a_0 = \\ &= \text{EMMLS}(M),\end{aligned}\quad (5)$$

которая представляет собой модель экстраполяции временных рядов по выборке максимального подобия (*extrapolation model based on maximum likeness set*, EMMLS).

2. Идентификация модели экстраполяции

Задачу идентификации модели (5), т. е. задачу нахождения параметра M , предлагается решать по следующей схеме.

- Исходный временной ряд Z_1^N разделим на три части в пропорции 2:2:1. Полученные части назовем *базовый период* (40 %), *тестовый период* (40 %) и *контрольный период* (20 %) временного ряда соответственно.
- Исходя из решаемой задачи прогнозирования определяем значение P , а также диапазон возможных значений параметра M . Первоначально рекомендуется брать широкий диапазон возможных значений M , например $M \in [\bar{P}, 30 \cdot \bar{P}]$, а затем его последовательно уточнять.
- Для каждого значения параметра M из установленного диапазона прогнозируем значения временного ряда внутри *тестового периода*.
- По результатам прогнозирования определяем среднюю абсолютную ошибку прогноза (*mean absolute error*, MAE)

$$\text{MAE} = \frac{1}{K} \sum_{i=1}^K |\check{Z}_i - Z_i|, \quad (6)$$

где K — число значений временного ряда, попавших в *тестовый период*.

- Строим график зависимости MAE от M для тестового периода и определяем диапазон значений M , соответствующий устойчивому минимуму MAE.
- На последнем шаге экспертом выбирается параметр модели M внутри диапазона устойчивого минимума.

Рассмотрим в качестве примера зависимость MAE от M для временного ряда энергопотребления европейской территории РФ, представленного на рис. 2. Первоначально был выбран диапазон $M \in [24, 360]$, внутри которого выделены значения $M \in [168, 240]$. На последнем шаге экспертно было выбрано значение $M = 216$.

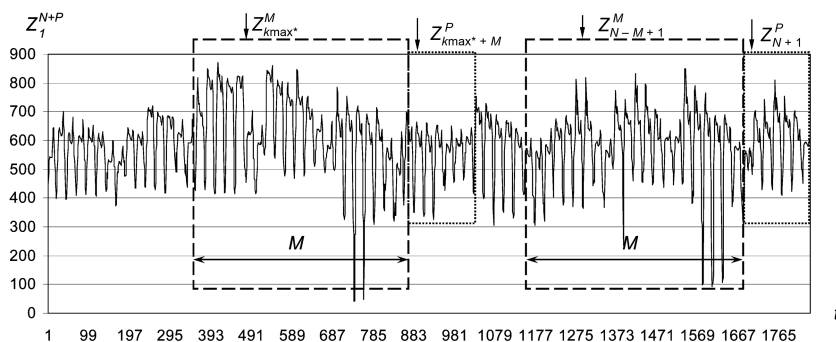


Рис. 1. Положения выборок $Z_{k_{\max}^*}^M$, $Z_{k_{\max}^* + M}^P$, Z_{N-M+1}^M и Z_{N+1}^P на оси времени

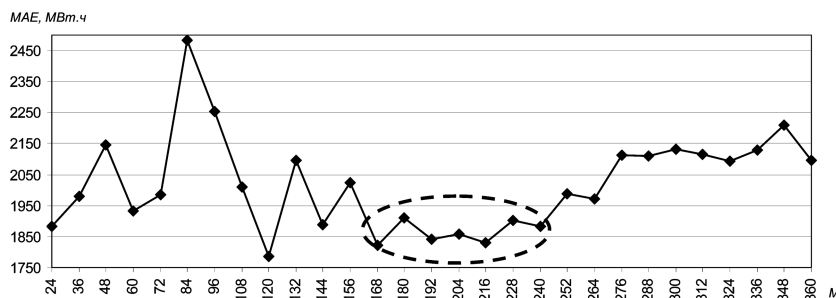


Рис. 2. Зависимость MAE от параметра модели M , $M \in [24, 360]$

3. Эффективность модели экстраполяции

Исследования эффективности модели экстраполяции (4) выполнены на следующих временных рядах, предоставленных ОАО "АТЭС" [12]:

- энергопотребление европейской ценовой зоны (ЕЦЗ);
- энергопотребление сибирской ценовой зоны (СЦЗ);
- цена РСВ европейской ценовой зоны (ЕЦЗ);
- цена РСВ сибирской ценовой зоны (СЦЗ).

Каждый временной ряд содержит почасовые равноотстоящие значения за период с 01.09.2006 по 30.09.2009. Кроме экстраполяции почасовых значений указанных временных рядов на следующие сутки ставится задача экстраполяции агрегированных суточных значений — суммарного энергопотребления за сутки и среднего значения цены РСВ за сутки. Параметры временных рядов представлены в табл. 1.

Задача экстраполяции временных рядов цен РСВ и энергопотребления, имеющих почасовое разрешение, заключается в нахождении 24 почасовых значений временного ряда на будущие сутки. Для рядов, имеющих суточное разрешение, задача экстраполяции состоит в нахождении одного значения на будущие сутки. Согласно работе [3] обе задачи экстраполяции относятся к категории краткосрочного прогнозирования.

В качестве *контрольного периода* прогнозирования использован период с 01.03.2009 по 30.09.2009 (семь месяцев, около 5000 значений для времен-

Параметры временных рядов

Разрешение	Временной ряд	Длина ряда	Среднее значение	Стандартное отклонение	Минимум	Максимум
Почасовое	Энергопотребление ЕЦЗ (МВт·ч)	27 023	81 441	10 640	57 847	110 586
	Энергопотребление СЦЗ (МВт·ч)	27 023	22 338	3 023	15 328	30 666
	Цена РСВ ЕЦЗ (руб/МВт·ч)	27 023	610	194	0	1 559
	Цена РСВ СЦЗ (руб/МВт·ч)	27 023	368	185	0	1 029
Суточное	Энергопотребление ЕЦЗ (МВт·ч)	1 126	1 954 502	218 255	1 519 197	2 438 008
	Энергопотребление СЦЗ (МВт·ч)	1 126	536 101	68 930	408 408	695 606
	Цена РСВ ЕЦЗ (руб/МВт·ч)	1 126	610	144	221	1 204
	Цена РСВ СЦЗ (руб/МВт·ч)	1 126	368	176	0	734

ных рядов, имеющих часовое разрешение; около 200 значений для рядов, имеющих суточное разрешение).

В табл. 2, 3 приведены численные результаты точности прогнозирования, а также указаны параметры моделей M каждого временного ряда, полученные путем решения задачи идентификации по схеме раздела 2. Приведенные значения параметра M могут быть взяты за основу при работе с временными рядами энергопотребления и цен РСВ. Для временных рядов другой предметной области необходимо решать задачу идентификации заново.

Для сравнения точности экстраполяции кроме величины MAE оценивалась также средняя абсолютная ошибка в процентах (*mean absolute percentage error*, MAPE)

$$MAPE = \frac{1}{Q} \sum_{i=1}^Q \frac{|\tilde{z}_i - z_i|}{z_i} 100 \%,$$

где Q — число значений временного ряда, попавших в контрольный период.

3.1. Прогнозирование энергопотребления. Результаты прогнозирования энергопотребления представлены в табл. 2.

Полученные значения величины MAPE краткосрочного прогнозирования лежат в диапазоне от 1,04 до 1,86 %.

Проведем сравнение достигнутой точности с точностью прогнозирования энергопотребления, указанной в других публикациях. В работе [3] описаны результаты краткосрочного прогнозирования энергопотребления Шанхайской энергосистемы (*Shanghai Power Grid*) и приведены значения MAPE различных моделей экстраполяции в диапазоне от 2,8 до 3,4 % в зависимости от их особенностей. Разработчики специализированного программного пакета "Энергостат" в статье [7] приводят значения MAPE краткосрочного прогнозирования энергопотребления для регионов РФ за 2000 г., которые лежат в диапазоне от 1,21 % до 3,32 % (в зависимости от региона). Австрийская

компания iRM (<http://www.irm.at>), производитель программного продукта ЮРТ PRO для прогнозирования энергопотребления, указывает значения MAPE краткосрочного прогнозирования энергопотребления 1—2 %.

Приведенные результаты краткосрочного прогнозирования энергопотребления показывают эффективность применения предложенной в работе модели экстраполяции временных рядов по выборке максимального подобию для решения задачи прогнозирования временных рядов энергопотребления.

Таблица 2

Результаты прогнозирования энергопотребления

Временной ряд	Разрешение	M	MAE, МВт·ч	MAPE, %	Время экстраполяции, ч
Энергопотребление европейской ценовой зоны	Почасовое	216	1 347	1,04	1,80
	Суточное	6	19 736	1,10	0,15
Энергопотребление сибирской ценовой зоны	Почасовое	24	378	1,86	4,20
	Суточное	8	7 727	1,44	0,05

Таблица 3

Результаты прогнозирования цен РСВ

Временной ряд	Разрешение	M	MAE, руб/МВт·ч	MAPE, %	Время экстраполяции, ч
Цена РСВ европейской ценовой зоны	Почасовое	360	49,54	7,00	2,10
	Суточное	16	34,50	4,81	0,04
Цена РСВ сибирской ценовой зоны	Почасовое	84	65,90	39,78	4,93
	Суточное	30	60,39	32,93	0,04

3.2. Прогнозирование цен РСВ. Результаты прогнозирования цены РСВ представлены в табл. 3.

Из таблицы видно, что значения MAPE для цен РСВ сибирской ценовой зоны велики в связи с тем, что в контрольном периоде имеются нулевые значения цен РСВ, которые при анализе ошибок замещались на 0,01 (т. е. 1 коп/МВт·ч).

Выше отмечалось, что задача экстраполяции цен РСВ для энергорынка России является новой, потому нет возможности выполнить сравнение точности краткосрочного прогнозирования. В единственном доступном источнике [13] приведено значение MAE для краткосрочного прогнозирования цен РСВ, равное 55,09 руб/МВт·ч.

Для аналогичной задачи прогнозирования цен энергорынков Испании и Калифорнии в работе [5] указаны значения MAPE, равные 5–10 % (в зависимости от месяца). Кроме того, в работе [5] указано, что данный уровень точности прогнозирования является достаточным для эффективного финансового планирования на указанных рынках. Предложенная в данной работе модель экстраполяции позволила получить для европейской ценовой зоны значения MAPE, равные 4,8 и 7,0 % (табл. 3). На этом основании можно утверждать, что разработанная модель экстраполяции эффективна для краткосрочного прогнозирования цен РСВ европейской ценовой зоны. В то же время из табл. 3 видно, что задача экстраполяции цен РСВ сибирской ценовой зоны требует дополнительных исследований и доработок модели.

Заключение

В статье предложена модель экстраполяции временных рядов по выборке максимального подобию. Основные особенности модели экстраполяции заключаются в следующем:

- модель экстраполяции по выборке максимального подобию решает задачи краткосрочного прогнозирования (от 1 до 50 значений вперед);
- модель не накладывает ограничений ни на непрерывность, ни на стационарность временного ряда;
- для работы с моделью желательно, чтобы временной ряд содержал не менее 1000 отсчетов;

- модель может быть расширена в целях решения задачи экстраполяции временного ряда с учетом влияния внешних факторов, представленных в виде других временных рядов (например, учет температуры при прогнозировании энергопотребления).

Представленные в работе результаты подтверждают эффективность использования модели для прогнозирования исследуемых временных рядов цен РСВ и энергопотребления. В развитие работы планируется оценить эффективность модели экстраполяции для задач прогнозирования временных рядов из других предметных областей.

Автор выражает благодарность своему научному руководителю д-ру техн. наук, проф. Ю. Н. Павлову за помощь в обсуждении модели, постановке задачи для экспериментов, а также подготовке статьи.

Список литературы

1. Бокс Дж., Дженкинс Г. М. Анализ временных рядов, прогноз и управление. М.: Мир, 1974. С. 406.
2. Тихонов Э. Е. Прогнозирование в условиях рынка. Новосибирск: Изд. СевКавГТУ, 2006. С. 8–15.
3. M. Sc. Jingfei Yang. Power System Short-term Load Forecasting. Elektrotechnik und Informationstechnik der Technischen Universität Darmstadt. 2006. С. 81–84.
4. Extrapolation [Электронный ресурс] // The free encyclopedia "Wikipedia" 2001. URL: <http://en.wikipedia.org/wiki/Extrapolation>.
5. Reinaldo C., Garcia A. GARCH Forecasting Model to Predict Day-Ahead Electricity Prices. German Institute of Economic Research, Germany. 2003. С. 9–12.
6. Макоклюев Б. И., Еч В. Ф. Взаимосвязь точности прогнозирования и неравномерности графиков электропотребления // Энергетик. 2005. № 5. С. 20.
7. Макоклюев Б. И., Владимирова А. И., Фёфелова Г. И. Прогнозирование потребления электроэнергии АО Мосэнерго // Энергетик. 2001. № 4. С. 12.
8. Badran Saeed M. Short term Electrical Load Forecasting // Australian Journal of Basic and Applied Sciences. 2009. N 1. С. 9.
9. Kumar M. Short-term load forecasting using artificial neural network techniques. Department of Electrical Engineering National Institute of Technology Rourkela. 2009. С. 48.
10. Al-Shareef A. J., Mohamed E. A., Al-Judaibi E. One hour ahead load forecasting using artificial neural network for the Western Area of Saudi Arabia. World Academy of Science, Engineering and Technology. 2008. С. 6.
11. Draper N. R., Smith H. Applied regression analysis. New York: Wiley, In press, 1981. С. 693.
12. Открытое акционерное общество "Администратор торговой системы оптового рынка электроэнергии и мощности". URL: <http://www.atsenergo.ru/>.
13. BI EnergoPrice. Прогнозирование цен на электроэнергию. URL: http://www.bi-grouplabs.ru/Rech/electricity/BI_EnergoPrice.html.

Р. Г. Гюльмамедов, канд. техн. наук,
 зав. Департаментом развития
 информационного общества,
 Министерство связи
 и информационных технологий,
 Республика Азербайджан, г. Баку,
 e-mail: gulmamedovr@rambler.ru

Выбор стратегии информатизации инновационного региона. Когнитивный подход

В условиях перехода к экономике знаний информатизация становится одним из ведущих факторов инновационного развития регионов. Рассматривается задача выбора стратегии информатизации инновационного региона. Отмечаются трудности решения задачи, обусловленные ее сложным слабоструктурированным характером, ограничивающим возможности применения традиционных методов экономико-математического моделирования и принятия решений. Предлагается методология когнитивной поддержки задачи, открывающая новые возможности для преодоления этих трудностей. Инструменты когнитивной поддержки могут быть полезными для менеджеров ИКТ-проектов и администраторов, принимающих решения.

Ключевые слова: экономика знаний, инновационный регион, стратегия информатизации, когнитивный подход

Введение

Современная концепция информационной экономики ставит фактор информатизации в число важнейших условий экономического роста как общества в целом, так и отдельных его субъектов и, в первую очередь, субъектов регионального уровня [1]. Однако на практике реализация этого теоретического положения сталкивается с многочисленными трудностями, обусловленными реальными условиями и реальными возможностями конкретных регионов [2], такими, например, как отсутствие в регионе развитой ИКТ¹-инфраструктуры, отсутствие навыков ИКТ-менеджмента у руководства предприятий, неблагоприятная инвестиционная политика региональных властей в отношении ИКТ-сектора, недостаточная надежность и безопасность ИКТ-инфраструктуры, неблагоприятные условия ведения бизнеса в регионе, не-

монотонная динамика макроэкономических процессов и т. д.

Важное практическое значение в связи с этим приобретает вопрос правильного выбора стратегии информатизации (СИ), учитывающий реальные условия и реальные возможности региона в контексте стратегических приоритетов и целей экономического развития региона (увеличение ВВП, конкурентоспособности национальной экономики, уровня доходов населения, обеспечение национальной безопасности и др.).

Научная поддержка вопроса выбора в столь сложных условиях повсеместно сталкивается с проблемой неопределенности, практически исключающей применимость традиционных методов экономико-математического моделирования и принятия решений.

Новые возможности для решения вопроса открывает методология когнитивного моделирования [3, 4], активно развиваемая в последние годы в теории управления сложными слабоструктурированными объектами.

Методология когнитивного моделирования

Методология когнитивного моделирования (КМ) основана на моделировании субъективных представлений экспертов об управляемой ситуации и включает:

- методы когнитивной (познавательной-целевой) структуризации знаний экспертов о ситуации;
- формализованное представление знаний экспертов в виде когнитивной карты;
- методы анализа когнитивной карты, ориентированные на решение практических задач управления ситуацией.

Когнитивная карта (КК) — это ориентированный граф (F, W) , где F — множество вершин, отражающих существенные (базисные) факторы управляемой ситуации (целевые факторы, управляющие факторы, факторы внешней среды); W — множество дуг, отражающих причинно-следственные (каузальные) отношения между факторами ситуации. Разнообразие КК определяется различными способами задания значений факторов и характера межфакторных отношений. По этим двум признакам выделяют: знаковые когнитивные карты, нечеткие когнитивные карты, нечеткие когнитивные карты на правилах (RBFCM-Rule Based Fuzzy Cognitive Maps) [3, с. 50].

Когнитивные карты служат как средством структурирования и формализации управляемых ситуаций, так и средством их анализа. Различные интерпретации вершин, дуг и межфакторных отношений приводят к различным методам анализа КК.

¹ В соответствии с [2] под ИКТ (информационные и коммуникационные технологии) понимается "оборудование, программное обеспечение и услуги в области информационных технологий, а также телекоммуникационное оборудование и услуги".

В настоящее время базовыми методами анализа являются:

- методы анализа влияний, направленные на определение знака и силы влияния между любой парой факторов;
- методы прогноза состояния ситуации в режиме саморазвития и в режиме управляемого развития [3, с. 51—55].

В последние годы методология КМ развивается в направлении совершенствования аппарата анализа КК [3, 4]. Методы анализа позволяют осуществлять модельные эксперименты над КК в целях решения прикладных задач управления и, в частности, решения приоритетной задачи сегодняшней управленческой практики — задачи формирования стратегии когнитивного управления. Подробно особенности этой задачи рассматриваются в работе [5] сотрудников Института проблем управления — ведущего НИИ РАН в области когнитивных технологий. В работе предлагается универсальный *метод формирования стратегии управления* слабоструктурированными объектами.

Методология когнитивного моделирования стратегий информатизации

Основные положения. Уже первые попытки применения метода в проектах информатизации выявили его существенные ограничения. Ограничения обусловлены тем обстоятельством, что динамика управляемого развития ситуаций в работе [5] моделируется с помощью аппарата линейных динамических систем Ф. Робертса [6]. К сожалению, возможности этого аппарата не позволяют моделировать все то многообразие нелинейных (немонотонных) межфакторных отношений, которое возникает в реальных проектах информатизации (мультипликативные связи, обратные связи, временное запаздывание, пороговые эффекты и др.).

Это потребовало модификации метода в части разработки: 1) предметно-ориентированной КК для задачи выбора СИ и 2) аппарата моделирования нелинейной (немонотонной) динамики этой карты.

Когнитивная карта СИ. Идеологическую основу разработанной карты составила структурная модель EIU² [2]. Модель содержит:

- эмпирическую зависимость экономического эффекта от использования ИКТ в развитых и развивающихся странах;
- перечень факторов, которые могут способствовать внедрению ИКТ и которые будут играть решающую роль в усилиях Европы по ускоре-

нию экономического роста и повышению производительности труда;

- немонотонную логику межфакторных отношений, возникающих в процессе "встраивания" ИКТ в экономику регионов.

Для учета этих отношений была разработана RBFCM-карта специального вида. База правил карты описывает весь набор межфакторных отношений, определенных моделью EIU. Для описания временных зависимостей между факторами использованы элементы темпоральной логики [7, с. 111].

Аппарат моделирования динамики КК. Разработан механизм моделирования динамики КК, основанный на элементах нечеткого продукционного вывода и темпоральной логики. Теоретическую основу механизма составила модель продукционной динамики, предложенная чл.-корр. РАН Г. С. Осиновым (Институт программных систем РАН) [8].

Механизм позволяет моделировать динамику СИ с учетом таких важных феноменологических особенностей модели EIU, как:

- нелинейный (U-образный) характер зависимости экономического роста от индекса развития ИКТ-инфраструктуры региона;
- наличие *порогового значения* этого индекса, после которого возможно положительное влияние информационного фактора на экономический рост;
- влияние на темпы этого роста ряда дополнительных условий, таких как *качество бизнес-среды, уровень ИКТ-менеджмента, качество ИКТ-инфраструктуры* региона (установленных "эталонной моделью" инновационной экономики, предложенной Всемирным экономическим форумом (Давос, 2008));
- наличие *периода адаптации* региона к информационному воздействию;
- *временное запаздывание* воздействия некоторых управляющих факторов, таких как обретение руководством предприятий региона навыков ИКТ-менеджмента, разработка и внедрение нормативно-правовой базы электронной коммерции, внедрение единых технических стандартов сетевой инфраструктуры и т. д. [2].

В настоящее время разработан демонстрационный прототип (общепринятый при создании knowledge-based моделей) КМ СИ. Опытная эксплуатация показала, что модель не дает тривиальных или абсурдных результатов, позволяет формировать неочевидные альтернативные СИ, которые, как правило, оказывались лучше СИ предлагаемых менеджерами ИКТ-проектов, а также осуществлять дивергентный (многокритериальный, многофакторный, многовариантный, сценарный) анализ обсуждаемых СИ, крайне важный в условиях сегодняшней нестабильной экономики.

² Модель EIU разработана по результатам обширного обследования в 60 странах, опросах более 100 руководителей компаний из 18 отраслей, подробных интервью с ведущими европейскими политиками, руководителями бизнеса и видными учеными.

Пример

В качестве примера приводятся результаты решения демонстрационной задачи выбора СИ.

Основные характеристики КК СИ

Целевой фактор: ВВП на душу населения (ВП).

Управляющие факторы (по [2]):

М. Навыки ИКТ-менеджмента на предприятиях региона.

Н. Эффективность НИР (стимулирование эффективных прикладных НИР, создание НИИ-центров и технопарков в сфере ИКТ).

И. Культура инноваций (государственная поддержка инноваций и предпринимательства в сфере ИКТ, свободный доступ к венчурному (рисковому) капиталу).

Р. Открытые и свободные рынки ИКТ (наличие открытых и конкурентных рынков телекоммуникаций, либерализация рынков телекоммуникаций, предотвращение образования виртуальных монополий на рынках широкополосного доступа).

Г. Уровень внедрения ИКТ в госорганах (е-правительство, е-образование и т. д.).

Б. Надежность ИКТ-инфраструктуры (единые сетевые стандарты, правовая защищенность и техническая безопасность индивидуальных сетевых ресурсов физических и юридических лиц).

Лингвистическая шкала для оценки текущих значений факторов КК и степени взаимовлияния факторов:

0,1 — ОЧЕНЬ_НИЗКИЙ (ОЧЕНЬ_СЛАБОЕ)

0,3 — НИЗКИЙ (СЛАБОЕ)

0,5 — УМЕРЕННЫЙ (УМЕРЕННОЕ)

0,7 — ВЫСОКИЙ (СИЛЬНОЕ)

0,9 — ОЧЕНЬ_ВЫСОКИЙ (ОЧЕНЬ_СИЛЬНОЕ)

Исходное состояние ситуации S_0 : М = 0,3; Н = 0,3; И = 0,1; Р = 0,3; Г = 0,5; Б = 0,3; ВП = 0,5.

Постановка задачи. Сценарий: в качестве ведущих управляющих факторов региона приняты факторы Б, М, И. Задача поиска оптимальной СИ заключается в нахождении такой последовательности воздействий на управляющие факторы, которая позволила бы увеличить темпы роста *ВВП на душу населения* от УМЕРЕННЫХ (исходное значение целевого фактора, равное 0,5) до ВЫСОКИХ (желаемое значение целевого фактора, равное 0,7).

Результаты моделирования. Такое управление, полученное в результате решения задачи, представлено на рисунке (см. четвертую сторону обложки).

В левой части рисунка отражена динамика активности в узлах *"Надежность ИКТ-инфраструктуры"* (Б), *"Навыки ИКТ-менеджмента"* (М) и *"Культура инноваций"* (И), которые последовательно с течением времени улучшаются (активность в этих узлах растет). Результатом такого управления становится увеличение темпов роста *"ВВП на душу*

населения" — от УМЕРЕННЫХ (0,5) до ВЫСОКИХ (0,7) (правая часть рисунка). Из рисунка видно, что для такого увеличения темпов роста ВВП необходимо повышать интенсивность всех трех управляющих факторов, но в разной последовательности и в разной степени. Прежде всего, надо значительно увеличивать *"Надежность ИКТ-инфраструктуры"* (например, за счет внедрения единых стандартов сетевой безопасности, надежных юридических и технических мер защиты интеллектуальной собственности — платежных Интернет-систем, электронных удостоверений личности (digital identities), электронных контрактов и т. д.). Основные усилия должны быть направлены на опережающий рост этого фактора; темпы его роста должны расти быстрее и более всего (до уровня выше 0,9) к концу периода имитации.

Остальные управляющие факторы по степени их влияния на результат распределяются следующим образом. Вторым по значимости является фактор *"Навыки ИКТ-менеджмента"*. В самом начале имитации темп активности этого фактора несколько уступает темпу фактора *"Надежность ИКТ-инфраструктуры"*, но существенно опережает темп фактора *"Культура инноваций"*. Для достижения заданного результата (увеличения темпов роста ВВП) на *"Культуру инноваций"* нужно обратить внимание только на третьем шаге имитации, когда будет достигнут первоначальный рост ВВП. После четвертого шага имитации требования к активности фактора *"Культура инноваций"* вновь можно ослабить (после этого момента темпы роста активности факторов *"Навыки ИКТ-менеджмента"* и *"Культура инноваций"* уравниваются).

Заключение

По оценкам Организации Экономического Сотрудничества и Развития (ОЭСР), сегодня в условиях информационного бума регионы могут инвестировать чрезмерные средства в ИКТ "либо в стремлении компенсировать недостаток квалификации, либо из-за отсутствия четкой рыночной стратегии информатизации" [9]. Поэтому власти регионов и руководители бизнеса должны сосредоточить внимание в первую очередь на правильном адекватном выборе СИ, которая с наибольшей вероятностью может "запустить в действие" информационный фактор экономического роста. Вопрос создания инструментов научной поддержки такого выбора и, в частности, инструментов когнитивной поддержки приобретает в связи с этим особую актуальность.

КМ могут быть использованы при разработке и рабочей корректировке "Концепций", "Приоритетных направлений" и "Программ" информатизации на федеральном, региональном и субрегиональном уровнях.

Список литературы

1. Макаров В. Л., Варшавский А. Е., Сутягин В. С. и др. Инновационный менеджмент в России: вопросы стратегического управления и научно-технологической безопасности. М.: Наука, 2004. 880 с.
2. Reaping the benefits of ICT: Europe's productivity challenge. 2004. (русс. перев.: "Реализация преимуществ ИКТ и экономический рост в Европе" — информационный документ, подготовленный Economist Intelligence Unit (EIU) — аналитической службой журнала "The Economist", по заказу компании Microsoft) URL: <http://www.eiu.com>
3. Труды междунар. конф. "Когнитивный анализ и управление развитием ситуаций". CASC'2007. М.: Ин-т проблем управления РАН. 2007. 247 с. URL: <http://www.ipu-conf.ru>, <http://CASC'2007>
4. Proceedings of the Intern. Conf. on Cognitive Modelling, ICCM 2004, Pittsburgh, Pennsylvania, USA, July 30 — August 1. 2004. 564 p.
5. Авдеева З. К., Коврига С. В., Макаренко Д. И., Максимов В. И. Когнитивный подход в управлении // Проблемы управления. 2007. № 3. С. 2—8.
6. Робертс Ф. С. Дискретные математические модели с приложениями к социальным, биологическим и экологическим задачам: Пер. с англ. М.: Наука, 1986. 312 с.
7. Поспелов Д. А. Ситуационное управление: Теория и практика. М.: Наука, 1986. 288 с.
8. Осипов Г. С. Динамика в системах, основанных на знаниях // Известия Академии наук. Теория и системы управления. 1998. № 5. С. 24—28.
9. Организация Экономического Сотрудничества и Развития (ОЭСР). Наука, технологии и промышленность: Перспективы 2008. URL: <http://www.oecd.org/html>

ИНФОРМАЦИОННЫЕ ТЕХНОЛОГИИ В МЕДИЦИНЕ

УДК 004.326; 681.178; 621.391; 519.21

В. М. Малыхин, канд. техн. наук, доц.,
А. В. Меркушева, канд. техн. наук,
инж.-исследователь,
Санкт-Петербургский государственный
политехнический университет,
e-mail: v_m_malykhin@mail.ru

Нейронные сети с вейвлет-преобразованием для классификации биосигналов

Рассмотрены методы классификации биомедицинских сигналов, обеспечивающих неинвазивную форму диагностики состояния отдельных органов. Основой служит искусственная нейронная сеть (НС), реализующая вейвлет-преобразование и принципы распознавания образов. Приведена структура НС, алгоритм ее обучения и элементы применения в кардиологии.

Ключевые слова: биологические сигналы, классификация, распознавание образов, вейвлет-преобразование, нейронная сеть, обучение

Сигналы, регистрируемые биомедицинскими датчиками (биосигналы) с поверхности тела человека, отражают как состояние отдельных органов, так и общий статус организма или его частей. Поэтому биосигналы (БС) обеспечивают получение существенной информации без каких-либо инвазивных измерений. Однако для диагностических решений достаточно сложно интерпретировать длительно регистрируемые БС, поскольку необхо-

димо выделить существенные признаки и определить, какой вид информации скрыт в БС. Ответы на эти вопросы получают методами распознавания образов, которые могут быть реализованы на основе нейронных сетей с вейвлет-преобразованием.

Распознавание образов (РО) включает последовательность этапов:

- преобразование образа в признаки, которые дают концентрированное представление образа (и в идеале содержат всю важную информацию);
- отбор среди этих признаков наиболее значительных, хорошо представляющих образ и не избыточных;
- классификация, которая в соответствии с отобранными признаками относит анализируемый образ к определенному классу.

Модель РО имеет разнообразные приложения, а в процессе биомедицинской диагностики образ идентифицируется как формализованные симптомы, регистрируемые сигналы или набор изображений пациента. Полученные классы представляют набор возможных диагнозов или диагностических утверждений. В зависимости от вида и типа образов и их формального описания для интерпретации их делается выбор среди статистических, синтаксических методов или методов искусственного интеллекта.

Статистический переход к РО использует измерения и преобразование структуры образа в вектор признаков (ВП). Отбор признаков выполняют последовательно, а иногда интуитивно, полагаясь на опыт экспертов. Рассматриваемые классификации зависят от распределений вероятностей при-

знаков. Наиболее часто (особенно в медицинских приложениях) используется линейный классификатор Байеса, предполагающий гауссово распределение признаков. Применимы и непараметрические методы (k -ближайших ВП) или полиномиальный классификатор для негауссова распределения ВП.

Синтаксический подход к РО основан на классификации описаний или графах структуры образа. Так называемые примитивы (линии, дуги или углы) рассматриваются как признаки, которые изменяются совместно. Классы образов представляются различающимися грамматиками строк примитивов. Процесс классификации выполняется группировкой сходных строк. Этот метод приспособлен к характеристике графических свойств измеряемых БС, но в клинической диагностике используется редко.

РО с использованием методов искусственного интеллекта [1, 2] ограничено теорией многосвязных систем, которые, в частности, реализуются в виде искусственных нейронных сетей (НС) [3–5]. Структура НС зависит от решаемой задачи, а параметры НС (ее синаптические веса) получаются в процессе обучения. НС сохраняет информацию о стратегии решаемой задачи, поэтому правило обучения соответствует приложению — выбору медицинского решения для диагностики.

Комбинация преимуществ НС в задачах РО с эффективным методом извлечения признаков включает две фазы:

- получение малого числа признаков из общего вектора признаков¹⁾, поступающего на вход НС. Это достигается путем время-частотного преобразования (ВЧП) вектора признаков [6–8], причем малое число легко интерпретируемых коэффициентов ВЧП-разложения составляют извлеченные признаки;
- вычисленные таким образом признаки рассматриваются как входные для НС простой структуры, которая служит для классификации.

Кроме синаптических весов (СВ) функционирование НС-системы РО зависит от параметров ВЧП, поэтому фаза обучения НС включает не только настройку СВ, но и дополнительную настройку время-частотных параметров преобразования, реализуемого на входных узлах НС. Такая процедура позволяет выполнять оптимальную автоматическую настройку входных параметров НС в зависимости от желаемых корректных классификаций на обучающей выборке.

¹⁾ Вектор признаков (ВП) составляют выборочные отсчеты биосигнала (БС).

Биосигналы, используемые для диагностики, часто являются нестационарными, поэтому для извлечения признаков целесообразно использовать ВЧП БС. Такое представление БС обеспечивает две точки зрения на образ в форме метода сигнала. Это его частотная характеристика и ее временное поведение²⁾.

Из всех ВЧП в НС-структуре для РО выбрано вейвлет-преобразование $w(a, \tau)$. Основным элементом такой НС-структуры является внутренний слой НС с функцией преобразования нейронов в виде вейвлет-преобразования вектора признаков. Математически вейвлет-преобразование выражается "внутренним произведением" (а фактически сверткой) БС $s(t)$ с вейвлет-функцией $h_a(t) = (1/\sqrt{a})h(t/a)$ [10–12]. Другая форма представления возможна через преобразования Фурье сигнала $S(f)$ и вейвлет-функции $H_a(f)$. Общие взаимосвязи определяются соотношениями

$$\begin{aligned} H_a(f) &= (1/\sqrt{a})aH(af) = \sqrt{a}H(af); \\ w(a, \tau) &= \langle s; h_a \rangle = (1/\sqrt{a}) \int_{-\infty}^{\infty} s(t)h_a^*\left(\frac{t-\tau}{a}\right)dt = \\ &= \sqrt{a} \int_{-\infty}^{\infty} S(f)H^*(af)\exp(j2\pi ft)df, \end{aligned} \quad (1)$$

где a — параметр масштаба; τ — параметр смещения ("сдвига") вейвлет-функции; $\langle \cdot; \cdot \rangle$ — символ свертки; * означает комплексное сопряжение; взвешивающий коэффициент $1/\sqrt{a}$ приводит вейвлет-функции для всех значений a к постоянной энергии.

Вейвлет-преобразование $w(a, \tau)$ может интерпретироваться как корреляционная функция между сигналом $s(t)$ и вейвлет-функцией $h_a^*(t)$. Свертка временных сигналов соответствует произведению их преобразований Фурье (ПФ):

$$\begin{aligned} \langle s(\tau); h_a(r) \rangle &= (1/\sqrt{a})\langle s(\tau); h(\tau/a) \rangle \Leftrightarrow \\ &\Leftrightarrow (1/\sqrt{a})S(f)aH^*(af) = \sqrt{a}S(f)H^*(af). \end{aligned} \quad (2)$$

Свертка, так же как ее дуальная операция в области ПФ, предполагает интерпретацию прохождения сигнала $s(t)$ через фильтр, представленный его импульсным откликом $h_a^*(\tau)$ или его передаточной функцией $aH^*(af)$.

Часто базовые вейвлет-функции модулируются окном, т. е. они составляются как произведение двух функций, таких как осциллирующий множитель, модулированный функцией Гаусса. Кроме того,

²⁾ Одновременная мгновенная оценка частоты в конкретный текущий момент времени возможна только с конечной точностью в силу принципа неопределенности: $\Delta\omega \cdot \Delta t \geq 0,5$ [9].

они должны удовлетворять некоторым дополнительным условиям. Модулированной является, в частности, комплексная вейвлет-функция Морлета [7]. Ее временное и частотное представления имеют вид

$$h_{\text{Мор}}(t) = (1/\sqrt{a})\exp[-0,5(t/a)^2 + 2\pi j f_0 t];$$

$$H_{\text{Мор}}(f) = \sqrt{2\pi a} \exp[-2\pi^2(a f - f_0)]. \quad (3)$$

При изменении параметра масштаба a вейвлет изменяет форму своей зависимости от времени: при большом значении a вейвлет становится растянутой низкочастотной функцией, при малом значении — сжимается и соответствует высокочастотной функции.

Анализ с использованием вейвлет-преобразования может рассматриваться в рамках системы многоканальных согласованных фильтров с шириной полосы, пропорциональной частоте [13]. Такая система характеризуется конкретной вейвлет-функцией и параметром f_0 . Спектр сигнала $S(f)$ умножается на функцию пропускания каждого фильтра $a_k H(a_k f)$, после чего преобразуется обратно во временную область (путем обратного ПФ). Такая процедура обеспечивает очень хорошее спектральное разрешение (за счет ухудшения временного разрешения) в области низких частот и очень хорошее временное разрешение (за счет ухудшения частотного разрешения) в области высоких частот.

При достаточно большом разнообразии топологий и структур НС и стратегий решения нелинейных задач [5, 14, 15] для области РО БС достаточно использовать хорошо известный многослойный персептрон (МСП)³⁾. В зависимости от входного вектора признаков $\mathbf{x}_i$ (индекс i — номер этого вектора в наборе N образов) выход МСП y_i оценивается в соответствии с правилом прямого распространения сигнала в НС. При РО нейроны выходного слоя МСП показывают, к какой категории или классу, $j = 1, \dots, M$, принадлежит вектор признаков $\mathbf{x}_j$. Выход нейрона o_j определяется взвешенной суммой выходов нейронов предшествующего слоя o_k , порогом нейрона θ_j и функцией преобразования нейронов f :

$$o_j = f(\sum_k w_{kj} o_k + \theta_j) = 1 + \exp[-(\sum_k w_{kj} o_k + \theta_j)] \quad (4)$$

(Выражение приведено для сигмоидной активационной функции преобразования нейронов:

³⁾ В составе МСП несколько слоев, в узлах которых находятся искусственные нейроны (ИН). Число их определяется размерностью решаемой задачи (в частности, размерностью вектора признаков при РО). Все ИН взаимно связаны направленными связями, сила которых определяется весовыми коэффициентами. Взвешенная связь вместе с функцией преобразования нейрона образует передаточную функцию ИН для сигнала, поступающего на следующий элемент МСП, т. е. взвешенные выходы нейронов поступают на все ИН следующего слоя, а на его входы поступают сигналы с узлов предшествующего слоя [5].

$f(x) = [1 + \exp(-x)]^{-1}$), которая наиболее часто используется в неспецифичных задачах, решаемых с МСП [3, 4].

Во время фазы обучения веса НС w_{kj} подстраиваются, чтобы минимизировать среднюю квадратичную ошибку E между желаемым выходным вектором $\mathbf{d}_i$ и действительным выходом $\mathbf{y}_i$ для всех входных векторов $\mathbf{x}_i$:

$$E = \sum_{i=1}^N \sum_{j=1}^N (d_{ij} - y_{ij})^2 = \min. \quad (5)$$

Поскольку эта задача является нелинейной, она не имеет решения в замкнутой форме. При решении ее итеративным градиентным методом частные производные по весам НС могут вычисляться с использованием обобщенного дельта-правила [3, 4], и результат достигается методом обратного распространения ошибки. Для весов связей на выходные нейроны МСП справедливо выражение

$$\partial E / \partial w_{\text{вых} kj} = - \sum_{i=1}^N \delta_{ij} o_{\text{скр.слик}} =$$

$$= - \sum_{i=1}^N (d_{ij} - y_{ij}) y_{ij} (1 - y_{ij}) o_{\text{скр.слик}}, \quad (6)$$

где $o_{\text{скр.сл}}$ — выход скрытого слоя нейрона.

Для двухслойного персептрона частные производные по весам нейронов в скрытом слое определяются выражением

$$\partial E / \partial w_{\text{скр.сл} kj} = - \sum_{i=1}^N \delta_{ij} x_{ik} =$$

$$= - \sum_{i=1}^N o_{\text{скр.сли} j} (1 - o_{\text{скр.сли} j}) (\sum_m \delta_{im} w_{\text{вых} jm}) x_{ik}. \quad (7)$$

Выражения (6) и (7) справедливы для сигмоидной функции преобразования нейрона⁴⁾. При обучении НС для начальных параметров (синаптических весов) принимаются случайные значения, а затем они обновляются на каждой итерации, пока ошибка НС не достигнет приемлемого минимума.

Расширение МСП, используемое в анализируемом варианте нейросетевого метода РО для БС, состоит во введении дополнительных узлов, которые обучаются представлению сигнала в виде его ВЧП. Такая гибридная структура НС является достаточно эффективным средством для РО при сигналах с изменяющимися во времени частотными характеристиками.

⁴⁾ Основа представлений о методе обратного распространения ошибки и аналитические выражения для компонент вектора градиента среднеквадратичной ошибки МСП разработаны Румельхартом (Rumelhart et al.) [16]. Обучение НС рекомендуется проводить несколько раз, используя различные случайные значения ее начальных параметров, а потом выбрать веса НС, дающие наименьшую ошибку.

Нейронные сети с вейвлет-преобразованием

Нейронные сети с вейвлет-преобразованием (НС_ВП) для классификации объединяют аспекты вейвлет-преобразования, реализующего фазу извлечения признаков, и возможность принимать решение исходя из характеристики полученных признаков. На стадии обучения НС_ВП не только приобретает адекватные решающие функции на достаточно сложной области задания БС, но и отслеживает ту часть пространства параметров (синаптических весов), которая подходит для надежного РО входных БС. Таким образом, НС_ВП является расширением МСП путем введения узлов с функцией преобразования нейронов в виде вейвлет-преобразования, которые выполняют предобработку с извлечением признаков из входного БС. Базовый вариант НС_ВП показан на рис. 1 (см. четвертую сторону обложки).

Узлы вейвлет-преобразования, которые настраиваются во время обучения НС, имеют функцию преобразования нейронов в виде $h((t - \tau_k)/a_k)$ (т. е. в виде модификации базовой вейвлет-функции $h(t)$). Эти узлы описываются параметром сдвига τ_k и параметром масштаба a_k , который обратно пропорционален частоте.

Формально выход узлов φ_{ik} является сверткой (или внутренним произведением) функции преобразования нейронов узла h_k и сигнала s_i , который является входом НС_ВП (индекс $i = 1, \dots, N$ означает номер сигнала):

$$\varphi_{ik} = \langle h_k; s_i \rangle = \int_t h^*\left(\frac{t - \tau_k}{a_k}\right) s_i(t) dt. \quad (8)$$

Выход узлов внутреннего слоя НС можно трактовать как величины непрерывного вейвлет-преобразования $w(a_k, \tau_k)$, т. е. они представляют корреляцию между модификацией вейвлет-функции h_k и сигналом s_i . В терминах системы многоканальных согласованных фильтров эти же величины трактуются как величины $s_k(\tau, a_k)$ на выходе полосовых фильтров в момент τ_k .

В верхней части рис. 1 представлена классифицирующая часть НС_ВП, она основывает свое решение на выходах вейвлет-узлов. Активация узлов y_{ij} классифицирующей части дает информацию, к какой категории или классу $j = 1, \dots, M$ принадлежит сигнал $s_i(t)$.

Во время обучения минимальная квадратичная ошибка $E = \sum_{i=1}^N \sum_{j=1}^M (d_{ij} - y_{ij})^2 = \min$ между желаемым вектором d_i на выходе НС и его фактическим значением y_i достигается путем использования градиентного метода. Частные производные весов w_{jk} в классификационной части НС вычисляются со-

гласно обобщенному дельта-правилу, описанному выше.

Кроме весов при обучении настраиваются параметры вейвлет-узлов. Частные производные по параметру масштаба a_k и по параметру сдвига τ_k зависят от вида базовой вейвлет-функции и определяются в соответствии с алгоритмом обратного распространения ошибки E :

$$\begin{aligned} \partial E / \partial \tau_k &= - \sum_{i=1}^N \sum_{j=1}^{H:M} \delta_{ij} w_{ij} (\partial \varphi_{ik} / \partial \tau_k), \\ \partial E / \partial a_k &= - \sum_{i=1}^N \sum_{j=1}^{H:M} \delta_{ij} w_{ij} (\partial \varphi_{ik} / \partial a_k), \end{aligned} \quad (9)$$

где δ_{ij} определено в алгоритме обратного распространения ошибки (6), (7); N — число БС в обучающей выборке; H и M — число весов, ведущих к узлам в скрытом слое или к выходному слою соответственно⁵⁾.

Таким образом, на каждой итерации фазы обучения веса и вейвлет-параметры изменяются так, чтобы снизить ошибку E . Когда достигается сходимость, вейвлет-узлы представляют ту часть время-частотной плоскости, которая наиболее подходит для РО (т. е. для решения о категории анализируемого БС). Схема обучающего алгоритма НС_ВП для классификации (РО БС) показана на рис. 2.

- **Выполнение алгоритма обучения НС_ВП с использованием вейвлет-функции Морлета.** Если в качестве базовой выбрать комплексную вейвлет-функцию Морлета [8], то согласно непрерывному вейвлет-преобразованию (где так называемая скейлограмма описывает распределение энергии БС во время-частотной плоскости) абсолютная величина внутреннего произведения $\langle h_k; s_i \rangle$ для каждого вейвлет-узла определяется выражением:

$$\begin{aligned} \varphi_{ik} &= \left| \int_t h\left(\frac{t - \tau_k}{a_k}\right) s_i(t) dt \right| = \\ &= \sqrt{\left[\int_t s_i(t) \cos\left(\omega_k \frac{t - \tau_k}{a_k}\right) \exp\left(-0,5 \left(\frac{t - \tau_k}{a_k}\right)^2\right) dt \right]^2 + \dots} \rightarrow \end{aligned}$$

⁵⁾ Число вейвлет-узлов в НС_ВП зависит от размерности решаемой задачи (т. е. от дискретного представления БС), но оно должно быть возможно небольшим, чтобы улучшить способность НС к обобщению — реализации РО на реальных БС, не участвовавших в обучении НС. Кроме того, как отмечалось, полезно проводить обучение при различных начальных значениях параметров НС (выбираемых случайным образом) и выбирать лучшее решение для НС_ВП, обеспечивающее минимальную ошибку. Для начальных значений вейвлет-параметров используются элементы априорной информации: если различие в группе БС ожидается для высоких частот в начальном временном интервале БС, то, по крайней мере, один вейвлет-узел должен быть инициализирован в этой части время-частотной плоскости.

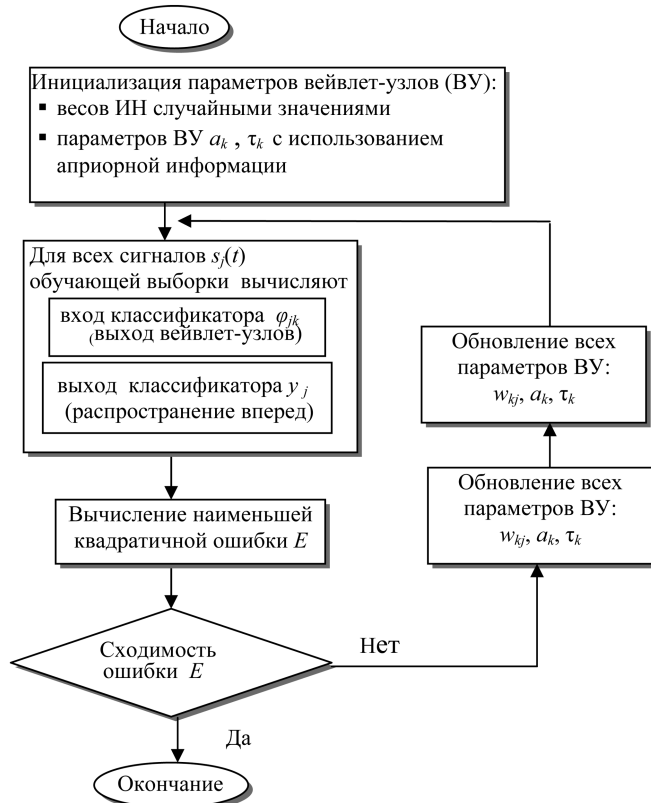


Рис. 2. Схема обучающего алгоритма нейронной сети с вейвлет-преобразованием для классификации (распознавания образов биосигналов)

$$\rightarrow + \left[\int_t s_i(t) \sin\left(\omega_k \frac{t-\tau_k}{a_k}\right) \exp\left(-0,5\left(\frac{t-\tau_k}{a_k}\right)^2\right) \right]^2 = \sqrt{o_{\cos ik}^2 + o_{\sin ik}^2}. \quad (10)$$

В правой части соотношения (10) введено простое обозначение для слагаемых, стоящих под корнем. Вместо фиксированной величины f_0 в вейвлет-преобразовании здесь использован частотный параметр $\omega_k = 2\pi f_k$, что придает большую общность полученному выражению.

Замена интегрирования суммированием (в силу дискретной обработки) приводит к выражениям, определяющим частные производные для "энергии E вейвлет-узла" с базовой функцией Морлета:

$$\frac{\partial E}{\partial \tau_k} = - \sum_{i=1}^N \sum_{j=1}^{H;M} \delta_{ij} w_{kj} \frac{1}{\varphi_{ik}} \sum_{m=1}^{\tau_m} s_i(t_m) \times \exp\left(-0,5\left(\frac{t_m-\tau_k}{a_k}\right)^2\right) \frac{1}{a_k} \left(\sigma_{\cos ik} \left[\omega_k \sin\left(\omega_k \frac{t_m-\tau_k}{a_k}\right) + \frac{t_m-\tau_k}{a_k} \cos\left(\omega_k \frac{t_m-\tau_k}{a_k}\right) \right] + \right.$$

$$\left. + \sigma_{\sin ik} \left[-\omega_k \cos\left(\omega_k \frac{t_m-\tau_k}{a_k}\right) + \frac{t_m-\tau_k}{a_k} \sin\left(\omega_k \frac{t_m-\tau_k}{a_k}\right) \right] \right); \quad (11a)$$

$$\frac{\partial E}{\partial a_k} = \frac{t-\tau_k}{a_k} \frac{\partial E}{\partial \tau_k}; \quad (11b)$$

$$\frac{\partial E}{\partial \omega_k} = - \sum_{i=1}^N \sum_{j=1}^{H;M} \delta_{ij} w_{kj} \frac{1}{\varphi_{ik}} \sum_{m=1}^{\tau_m} s_i(t_m) \times \exp\left(-0,5\left(\frac{t_m-\tau_k}{a_k}\right)^2\right) \frac{t_m-\tau_k}{a_k} \left[-\sigma_{\cos ik} \sin\left(\omega_k \frac{t_m-\tau_k}{a_k}\right) + \sigma_{\sin ik} \cos\left(\omega_k \frac{t_m-\tau_k}{a_k}\right) + \frac{t_m-\tau_k}{a_k} \sin\left(\omega_k \frac{t_m-\tau_k}{a_k}\right) \right], \quad (11c)$$

где δ_{ij} — параметр алгоритма обратного распространения ошибки, аналогичный (6) и (7); N — число сигналов в обучающей выборке; H — число весов к узлам скрытого слоя НС; M — число весов, идущих к узлам слоя классификации.

Эффективная схема вычислений этих частных производных может уменьшить общее время выполнения обучающего алгоритма для НС_ВП.

Область применения НС_ВП для классификации БС

Наиболее широкой областью применения НС_ВП является кардиология, где классификация БС трактуется как определение наиболее вероятной категории электрокардиографического сигнала (ЭКС, ЭКГ). Мотивацией применения наиболее совершенных средств анализа БС на основе нейронных сетей связано с тем, что подверженность вентрикулярной тахикардии (ВТ) рассматривается как одна из причин внезапной смерти, обусловленной кардиологическим заболеванием. Именно поэтому существует повышенный интерес к идентификации пациентов с высоким риском, которые склонны к прогрессирующему развитию ВТ.

В последнее десятилетие так называемые запаздывающие потенциалы (ЗП), появляющиеся на поверхностной регистрации ЭКГ, считаются своеобразными маркерами, указывающими пациента с ВТ. Однако ЗП появляются как малые сигналы с амплитудой в несколько микровольт и с частотным диапазоном 40—300 Гц. Обнаружение и оценка этих БС является нелегкой задачей, учитывая существенный уровень шума, который всегда присутствует при их регистрации. В качестве меры

некоторого снижения шума проводится предобработка: усреднение нескольких циклов ЭКС и последующая высокочастотная фильтрация [17]. Несколько позже в качестве более совершенной обработки ЗП предложено использовать ВЧП, поскольку высокочастотная составляющая ЭКС, включая ЗП, дает зависящие от времени характеристики, которые служат простейшей формой анализа спектра мощности [18, 19].

Развитие метода получения количественной оценки и минимизации числа признаков ЭКС и ЗП, включающей классификацию (элементы РО), связано с построением описанных выше структур НС_ВП. Сообщается об исследовании двух групп пациентов (по 20 и 25 пациентов из различных больниц и при наличии двух контрольных групп по 30 человек здоровых людей), которые исследовались для установления ВТ посредством электрической стимуляции. По клиническим критериям у 25 участников второй группы по крайней мере один раз регистрировался ранее инфаркт [17].

Для всех пациентов и контрольных групп проведено усреднение трех ортогональных отведений при дискретизации БС 1000 Гц. Каждая регистрация, состоящая из 300 точек отсчетов, включает полный QRS-комплекс. После прохождения высокочастотного фильтра Баттерворта 1-го порядка рассчитывается ортогональный вектор сигнала. Затем этот сигнал проходит через высокочастотный фильтр, чтобы снизить установленные низкочастотные компоненты. Фильтр работает, вычисляя разницу между сглаженным векторным сигналом и его несглаженным вариантом, поэтому дополнительные искажения практически незаметны. Сглаживание выполнено с интерполяцией методом сплайнов. Разностный сигнал поступает на НС_ВП в качестве входного образа ЭКС.

Процедуру классификации можно описать следующим образом. Для определения минимального числа вейвлет-узлов (ВУ), необходимого для извлечения информативных признаков из различных образов БС, сначала сравнивается НС_ВП с одним ВУ и с двумя ВУ. В обоих случаях использовалась базовая вейвлет-функция Морлета. Результаты время-частотного анализа [20] и опыт клиницистов показывают, что ЗП возникает примерно через 100 мс после начала QRS-комплекса⁶⁾. Сопоставление двух простейших вариантов НС_ВП определено по чувствительности и правильности классификации между здоровыми людьми (в контроле) и пациентами. Эти две характеристики служат критерием качества выбранного метода. В резуль-

тате показано, что классификация пациентов с ВТ относительно контроля составила 95 % практически независимо от числа использованных признаков.

Заключение

- ♦ Рассмотрены методы классификации биомедицинских сигналов (БС), обеспечивающие неинвазивный способ диагностики состояния отдельных органов и статуса организма в целом. Основой служит искусственная НС, у которой функция преобразования нейронов, расположенных во внутреннем слое, реализует вейвлет-преобразование БС, поступающего на вход сети. Даны элементы принципов распознавания образов, которые использованы при построении НС для классификации БС. Приведена функциональная схема нейронной сети с вейвлет-преобразованием (НС_ВП) и алгоритм ее обучения.
- ♦ Кратко описаны этапы преобразований сигнала с выделением информативных признаков и основные подходы к распознаванию образов.
- ♦ Даны элементы теории время-частотного преобразования сигнала и его разновидности — вейвлет-преобразования.
- ♦ Показан расширенный набор параметров НС_ВП, которые подлежат настройке на стадии обучения НС, и правило выбора начальных значений для этих параметров.
- ♦ Представлена основная схема градиентного метода минимизации нелинейного функционала. В качестве функционала использована квадратичная ошибка НС_ВП, которая включает разности компонент желаемого и фактического выхода обучаемой сети.
- ♦ Показан метод вычисления составляющих градиента функционала ошибки.
- ♦ Описан алгоритм обучения НС_ВП с использованием вейвлет-функции Морлета.
- ♦ Обсуждена одна из значимых областей применения НС_ВП — распознавание и классификация ЭКГ у пациентов из группы с повышенным риском развития вентрикулярной тахикардии.

Список литературы

1. Рассел С. Искусственный интеллект: Современный подход. М.: Изд. Вильямс, 2007. 1408 с.
2. Тим Дж. М. Программирование искусственного интеллекта в приложениях. М.: Изд. ДМК, 2006. 311 с.
3. Haykin S. Neural Networks, Prentice Hall, 1999. 836 p.
4. Меркушева А. В. Применение нейронной сети для анализа нестационарного сигнала. I. Основные принципы; II. Исследование и оптимизация нейронной сети // Научное приборостроение. 2003. Т. 13, № 1. С. 64—71. С. 72—84.
5. Меркушева А. В. Нейросетевые методы обработки сигналов в информационных системах: Элементы структуры, принципы обучения и мера многообразия отображений нейронной сети // Информационные технологии. 2005. № 3. С. 9—20.

⁶⁾ Это можно объяснить электрофизиологическим фоном порождения потенциала во время деполяризации волокон сердечной мышцы. Кроме того, частотный спектр ЗП находится в диапазоне 40—300 Гц.

6. **Consalves P., Baraniuk R., Bertran R. G.** Bilinear time-frequency transformation and its modifications // *Signal Processing Letters*. 1996. V. 3. N 3. P. 82.
7. **Меркушева А. В.** Классы преобразований нестационарного сигнала в ИИС: I. Элементы теории, II. Время-частотные преобразования // *Научное приборостроение*. 2002. Т. 12. № 2. С. 50—58, С. 59—70.
8. **Меркушева А. В.** Время-масштабные преобразования для анализа динамики нестационарного сигнала в ИИС // *Информационные технологии*. 2005. № 2. С. 41—48.
9. **Malychina G. F., Mercusheva A. V.** Metrological aspects of non-stationary signal transformation for dynamic spectrum analysis // 10th IMEKO TC7 International Symposium on Advances of Measurement Science. Saint-Petersburg, Russia. 2004. V. 1. P. 212—216.
10. **Cohen A., Covacevic R.** Wavelets: Mathematical background // *Proceedings of IEEE*. 1996. V. 84. N 4.
11. **Мала С.** Вейвлеты в обработке сигналов. М.: Мир, 2005. 670 с.
12. **Меркушева А. В.** Время-масштабные (вейвлет-)преобразования для спектрально-временного анализа // *Научное приборостроение*. 2002. Т. 12. № 3. С. 68—82.
13. **Малыхина Г. Ф., Меркушева А. В.** Согласованное многоканальное разделение сигнала с вейвлет-преобразованием и модифицированной полифазной структурой для спектрального вычитания фона речевого сигнала // *Научное приборостроение*. 2008. Т. 18. № 1. С. 110—117.

14. **Малыхина Г. Ф., Меркушева А. В.** Элементы статистической концепции обучения нейронной сети и прогнозирования точности ее функционирования // *Научное приборостроение*. 2005. Т. 15. № 1. С. 29—45.
15. **Малыхина Г. Ф., Меркушева А. В.** Концепции информации и энтропии в нейросетевых методах анализа сигналов в ИИС // *Информационные технологии*. 2005. № 12. С. 26—36.
16. **Haykin S.** *Neural Networks, A comprehensive foundation*. N-Y: Prentice Hall, 1984. 380 p.
17. **Breithardt G., Cain M. E., El-Sherif N., Flowers N. C., Hombach V.** Standards for analysis of ventricular late potentials using high resolution or signal averaged electrocardiography // *Europ. Heart Journal*. 1991. N 12. P. 473—480.
18. **Morlet D., Peyrin F., Desseigne P., Touboul P., Rubel P.** Wavelet analysis of high resolution signal-averaged ECGs in postinfarction patients // *J. of Electrocardiology*. 1993. Vol. 23 (suppl. 4). P. 311—320.
19. **Meste P., Rix H., Caminal P., Thacor N. V.** Ventricular late potentials characterization in time-frequency domain by means of wavelet transform // *IEEE Transactions on Biomedical Engineering*. 1994. Vol. 41. P. 625—634.
20. **Schramm J., Frumento C., Dickhaus H., Schmitt C., Braehmarm J.** Ventricular late potentials in patients with and without inducible sustained ventricular tachycardia: a combined analysis in the time and frequency domain // *Europ. Heart Journal*. 1991. Vol. 12. P. 563—572.

УДК 519.72

С. В. Зыкин, д-р техн. наук, зав. лаб.,
Омский филиал института математики СО РАН,
А. Н. Полуянов, канд. техн. наук, науч. сотр.,
e-mail: andrey.poluyanov@gmail.com
Омский филиал института математики СО РАН,
А. К. Чернышев, д-р мед. наук, проф.,
Омская государственная медицинская академия,
А. И. Ревзин, канд. мед. наук,
зав. отделением анестезиологии,
Клинический медико-хирургический центр
Министерства здравоохранения Омской области

Технология подготовки и анализа данных для построения медицинских оценочных шкал¹

Рассматривается технология подготовки и анализа данных на примере построения шкалы оценки состояния пациентов. Предложен итерационный подход с контролем корректности к формированию многомерного представления данных из исходного реляционного представления. Рассмотрена технология проведения анализа данных и формирования оценочных шкал.

Ключевые слова: межмодельные преобразования, гиперкубическое представление данных, шкалы оценки состояния пациентов

1. Постановка задачи

Анализ накопленной информации является актуальной проблемой для многих исследовательских и прикладных задач. Традиционным способом ее решения в настоящее время является технология оперативной аналитической обработки данных OLAP (*online analytical processing*) [1—8]. Основой OLAP-технологии является построение гиперкубического (многомерного) представления данных.

На основе статистических данных по историям болезни пациентов с печеночной недостаточностью была поставлена задача аналитической обработки данных для получения выводов об эффективности применения новых методов лечения и реабилитации пациентов с заболеваниями печени.

Целью анализа данных была выбрана шкала оценки печеночной недостаточности. Шкалы оценки тяжести и прогноза состояний пациентов и их компьютеризированные варианты продолжают разрабатывать и применять в современной медицине критических состояний [9]. Однако основным недостатком большинства шкал является "одномоментность" их использования, т. е. программное обеспечение не позволяет отследить состояние пациента в динамике. Кроме того, отсутствие визуального компонента представления о динамике критического состояния у индивидуального больного затрудняет максимальное продвижение в практическую сферу подобного рода шкал.

¹ Работа выполнена по проекту № 2.2 ОМН РАН за 2009 г., грант РФФИ № 09-07-00059-а.

В последние годы отмечено, что среди основных патологических синдромов, отягощающих течение послеоперационного периода, особенно после операций на печени, наиболее тяжело проявляется синдром острой печеночной недостаточности. Имеются попытки разработки компьютеризированных шкал для объективной диагностики степени тяжести печеночной недостаточности [10, 11], однако эти исследования еще малочисленны и не получили достаточно широкого применения.

В данной работе использована технология итерационной подготовки многомерного представления данных и реализовано программное обеспечение визуальной интерпретации состояния пациентов.

Технология итерационной подготовки многомерных представлений предполагает хранение исходных данных в нормализованном реляционном виде. Рассмотрим схему соответствующей базы данных (подчеркнуты ключевые атрибуты отношений).

Пациенты

<u>ID пациента</u>	ФИО	<u>Дата рождения</u>	<u>Дата поступления</u>	№ группы	Адрес
--------------------	-----	----------------------	-------------------------	----------	-------

Группы

<u>№ группы</u>	Группа (описание)
-----------------	-------------------

Виды показателей

<u>ID анализа</u>	Наименование анализа	Единицы измерения	Комментарий
-------------------	----------------------	-------------------	-------------

Исходные показатели анализов

<u>ID пациента</u>	<u>ID анализа</u>	<u>Дата взятия</u>	Значение анализа
--------------------	-------------------	--------------------	------------------

В отношении "Пациенты" хранятся данные по пациентам. Отношение "Группы" предназначено для разделения пациентов на группы, определяющие вид применяемого метода лечения. Всего было представлено три группы пациентов. К группам 1 и 2 относятся пациенты, находящиеся в тяжелом состоянии. Состояние пациентов группы 1 является наиболее тяжелым. При лечении пациентов групп 1 и 2 применяли новые методы лечения. Группа 3 — контрольная, при лечении пациентов этой группы использовались обычные методы лечения. Отношение "Виды показателей" содержит описание анализов, применяемых для контроля состояния пациентов. В отношении "Исходные показатели анализов" хранятся значения показателей анализов пациентов за определенную дату.

В работах [1—4] предполагается постоянное хранение и обновление промежуточного представления данных в виде гиперкубов (MOLAP). В данной работе предлагается следующая последовательность работы с данными.

1. Исходные данные должны быть представлены в реляционном нормализованном виде, и к ним обеспечивается доступ по технологии OLTP (*online transaction processing*).

2. Пользовательское представление данных в виде гиперкубов, реализующее технологию OLAP, обеспечивается инструментарием, преобразующим исходные данные в гиперкуб.

3. Гиперкубическое представление данных далее используется для выявления закономерностей, предсказания неизвестных значений и т. д.

Существуют инструментальные средства, частично или полностью реализующие перечисленные технологии, например, Microsoft Analysis Services, ORACLE Analytic Workspace Manager. Однако использование этих средств не позволяет динамически формировать промежуточные представления данных.

В данной работе предполагается, что основой аналитической работы пользователя является необходимость формирования новых гиперкубов, а не многократное формирование реализации одного и того же гиперкуба. Следовательно, основное внимание необходимо акцентировать на сокращении времени формирования схемы нового гиперкуба, а формирование представления гиперкуба должно быть выполнено автоматически соответствующими алгоритмами.

В работе [8] для автоматизации формирования гиперкуба предлагается использовать следующую последовательность преобразований:

$$RRD \Rightarrow TJ \Rightarrow ST,$$

где *RRD* — реляционное представление данных; *TJ* — таблица соединений; *ST* — гиперкуб "семантическая трансформация". В данном случае *RRD* — представление исходной модели данных, *ST* — целевой. Представление *TJ* является промежуточным.

Выделим основные преимущества предлагаемой в данной работе технологии перед упомянутыми инструментами:

- гиперкубическое представление, сформированное в разработанной системе, удовлетворяет свойству соединения без потерь информации, что предотвращает появление лишних значений в рабочей области гиперкуба;
- использование алгоритма автоматического формирования контекста приложения упрощает процесс формирования схемы гиперкуба, существенно снижая требования к квалификации пользователя в области информационных технологий;
- в предлагаемом подходе есть возможность наложения на данные контекстных ограничений, позволяющих на содержательном уровне (сделав выбор между всеми или только связанными значениями) контролировать набор кортежей, попадающих в результирующее представление. Рассмотрим подробнее предлагаемую технологию.

2. Формирование таблицы соединений

Пусть R — отношение, определенное на множестве атрибутов $U = \{A_1, A_2, \dots, A_n\}$. Результатом нормализации R является декомпозиция $\rho(R_1, R_2, \dots, R_k)$, где R_i — отношения, определенные на подмножествах атрибутов $[R_i] \subseteq U$. Кортеж $t_i[X]$ — совокупность значений атрибутов $A_j \in X \subseteq U$, заданных в кортеже $t_i \in R_i$. Определим пустое значение атрибута A_j в кортеже t : $t[A_j] = emp$. При этом значение emp не равно любому другому значению, в том числе неопределенному и пустому.

Определение 1. Выражение $R_1 \bowtie R_2 \bowtie \dots \bowtie R_k$, где $\bowtie$ — операция естественного соединения, будем называть существующим соединением, если для совокупности отношений R_i , $i = 1, \dots, k$, существует хотя бы одна перестановка $V_1, V_2, \dots, V_k$ отношений $R_1, R_2, \dots, R_k$, такая что $([V_1] \cup [V_2] \cup \dots \cup [V_j]) \cap [V_{j+1}] \neq \emptyset$, $j = 1, \dots, k-1$.

Таблица соединений (S, I) , где S — отношение, определенное на множестве атрибутов U , формируется следующим образом: пусть $t \in S$ — некоторый кортеж. Для каждого кортежа u каждого существующего соединения формируется кортеж t по следующим правилам: $t[A_j] = u[A_j]$, если атрибут A_j принадлежит соединению, и $t[A_j] = emp$ в противном случае. Каждому кортежу поставим в соответствие битовый вектор $l(t) = (l_1(t), l_2(t), \dots, l_k(t))$, где $l_j(t) = 1$, если отношение R_j участвует в текущем соединении, и $l_j(t) = 0$ в противном случае.

Рассмотрим отношение частичного порядка над кортежами $t \in S$.

Определение 2. Кортеж $t \in S$ является менее определенным или равным кортежу $t' \in S$, когда для любого атрибута A_j выполнено: если $t[A_j] \neq t'[A_j]$, то $t[A_j] = emp$ и $l_j(t') \geq l_j(t)$, $j = 1, \dots, k$. В этом случае будем писать: $t < t'$ и назовем кортеж t подчиненным кортежу t' .

Замечание. Рассмотренное определение частичного порядка означает, что кортеж t' содержит в себе все менее определенные либо равные кортежи. Следовательно, в представлении S достаточно хранить только кортеж t' .

Определение 3. Представление данных S со схемой $(A_1, A_2, \dots, A_n, I)$, в котором удалены все подчиненные кортежи, будем называть S -таблицей.

Таким образом, в S -таблице из промежуточных соединений и исходных отношений остаются кортежи, которые не являются подчиненными кортежам естественных соединений для большего числа отношений. Будем называть эти кортежи остатками.

Для реализации ограничений на данные предлагается ввести понятие контекстного ограничения.

Определение 4. Совокупность отношений $\{R_{m(1)}, R_{m(2)}, \dots, R_{m(p)}\}$ будем называть естественным контекстным ограничением, если при формировании S -таблицы участвует результат естественного соединения отношений $\{R_{m(1)}, R_{m(2)}, \dots, R_{m(p)}\}$ и не участвуют их остатки.

Дополнительное ограничение в виде логического выражения F накладывается на все кортежи S -таблицы. Для задания F удобно воспользоваться представлением в виде конъюнктивной нормальной формы (КНФ): $F = F_1 \wedge F_2 \wedge \dots \wedge F_q$, где $F_i = F_{i1} \vee F_{i2} \vee \dots \vee F_{is(i)}$, а F_{ij} — атомарные условия: $A_l \Theta const$ либо $A_l \Theta A_m$, где Θ — операция ($=$, $\neq$, $\leq$, $\geq$, $<$, $>$). Неопределенные атомарные условия на промежуточных естественных соединениях заменяются на значение "ИСТИНА" или "ЛОЖЬ" так, чтобы не влиять на промежуточные результаты.

В работе [8] рассматривается алгоритм построения таблицы соединения "сверху вниз". То есть сначала формируется естественное соединение всех отношений, входящих в декомпозицию $\rho(R_1, R_2, \dots, R_k)$, и результат дополняется к изначально пустой таблице соединения. Затем формируются естественные соединения для подмножеств отношений, являющиеся сочетаниями без повторов из k элементов по j , где $j = k-1, k-2, \dots, 1$, и удовлетворяющих определению 1. Возможна другая стратегия формирования S -таблицы: "снизу вверх". Сначала формируются кортежи в S -таблице для единичных отношений, затем для парных соединений и т. д., с удалением подчиненных кортежей из S -таблицы. С точки зрения конечного результата обе эти стратегии неразличимы. Однако вторая стратегия требует значительного объема памяти для хранения промежуточных результатов при одинаковом числе операций ввода/вывода.

Выделим основные свойства таблицы соединений.

Свойство 1. Если декомпозиция $\rho(R_1, R_2, \dots, R_k)$ обладает свойством соединения без потерь информации и отсутствуют многозначные зависимости $\emptyset \twoheadrightarrow X(Y)$, то естественное соединение $R_1 \bowtie R_2 \bowtie \dots \bowtie R_k$ будет существующим.

Свойство 2. Таблица соединений S для произвольной схемы реляционной базы данных всегда существует и единственна.

Свойство 3. (основное). Любое существующее соединение подмножества отношений базы данных равно проекции на S по этому подмножеству отношений.

Таким образом, существование таблицы соединений для любых реляционных баз данных позволяет считать ее обобщением понятия "универсальное реляционное отношение".

3. Формирование гиперкуба

Сформированная S -таблица является источником данных для построения гиперкуба. Пусть X — множество атрибутов из U , входящих без изменения в результирующее представление ST ; Y — множество атрибутов из U , значения которых становятся именами атрибутов в результирующем представлении; Z — множество атрибутов из U , значения которых становятся значениями вновь введенных атрибутов, $W = U \setminus (X \cup Y \cup Z)$. На множества X , Y , Z накладываются следующие ограничения: $X \cap Y = \emptyset$, $X \cap Z = \emptyset$, $Y \cap Z = \emptyset$. Множества атрибутов X и Y являются измерениями, а значения Z — мерами.

Схема ST имеет следующий вид: $Sch(ST) = \{X, Dom(Y) \times \{Z\}\}$, где Sch — схема отношения; Dom — множество допустимых значений атрибутов, $\times$ — декартово произведение множеств, формирующее новые имена измерений. $Dom(Y) = Dom(Y_1) \times Dom(Y_2) \times \dots \times Dom(Y_m)$, $Y_j \in Y$. Вновь введенные атрибуты имеют следующие имена: $y_1^i \dots y_m^j Z_p$, где y_l^j — j -е значение атрибута Y_l ; Z_p — атрибут из множества Z .

Представление ST со схемой $Sch(ST)$ является плоской таблицей. Оно преобразуется в гиперкуб следующим способом: $Dom(X)$ — значения вдоль первой координаты; $Dom(Y_1) \times \{Z\}$ — значения вдоль второй координаты; $Dom(Y_2) \times \{Z\}$ — значения вдоль третьей координаты и т. д. Достаточным условием существования представления ST является наличие функциональной зависимости $XY \rightarrow Z$.

Описание алгоритмов, используемых для формирования таблицы соединений и гиперкуба, приведены в работе [12].

4. Анализ данных и построение диагностической шкалы

Для анализа эффективности применяемых методов лечения и реабилитации было построено следующее гиперкубическое представление данных.

Выбранные отношения: Пациенты, Группы, Виды показателей, Исходные показатели анализов;

Атрибуты множества X : Дата взятия, ФИО пациента;

Атрибуты множества Y : Наименование анализа, № группы;

Атрибуты множества Z : Значение анализа;

Ограничения: $F = \text{"Наименование анализа"} = ?$, $C = \{\{\text{Виды показателей, Исходные показатели анализов}\}\}$.

Затем данные были проанализированы по непараметрическому критерию Уилкоксона [13]. Получены следующие результаты.

При сравнении групп 1 и 2 с группой 3 (контрольная группа) по значениям параметров АСТ, АЛТ, ЛИИ получено, что число инверсий параметра в группе пациентов 1 и 2 по сравнению с группой 3 на третий день лечения лежит в интервале значений с уровнем значимости 5 %. Это позволяет принять гипотезу о том, что значения параметров принадлежат одной генеральной совокупности. Сравнение тех же больных по тому же параметру на первый день лечения показывает, что эта гипотеза должна быть отвергнута.

Результаты обработки данных по критерию Уилкоксона приведены в таблице.

На основании данного исследования можно сделать вывод об эффективности применяемого метода лечения для групп 1 и 2.

Следующим этапом исследования является шкала диагностики степени тяжести заболевания пациентов. За основу были взяты применяемые в медицине шкалы для оценки состояния больного APACHE 2 и SAPS 2 [9].

По используемым в шкале параметрам интервальные границы были уточнены при обработке статистических данных с использованием меры Кульбака. Границы интервалов выбирались исходя из максимизации информационной меры Кульбака значимости параметров при сравнении выборок по группам 1, 2 и 3.

Расчет баллов для оценки состояния пациента осуществляется по следующей формуле:

$$M = \left(\frac{\sum \text{баллов по параметрам}}{\sum \text{максимальных баллов по параметрам}} \right) \cdot 100,$$

Результаты обработки данных по критерию Уилкоксона
Интервал: 995,3—1596,7

№ дня	Азот	Мочевина	Креатинин	Белок	Билирубин	ЛИИ	ПТИ	Глюкоза	АСТ	АЛТ	Калий	Натрий
1	1553,0	1639,0	1644,0	753,0	1979,0	2105,0	558,0	1680,0	1774,0	1708,0	1639,0	939,0
2	1678,0	1884,0	1682,0	1342,0	1918,0	1162,0	1208,0	1010,0	892,0	665,0	1224,0	1455,0
3	1910,0	1916,0	1692,0	1099,0	1941,0	1100,0	976,0	1151,0	1105,0	732,0	1059,0	1142,0
4	1662,0	1753,0	1463,0	1379,0	1941,0	1160,0	1311,0	834,0	1035,0	681,0	936,0	1987,0
5	1643,0	1762,0	1451,0	1203,0	1990,0	1420,0	1015,0	999,0	1244,0	858,0	1026,0	1804,0
8	1416,0	1625,0	1626,0	1062,0	1948,0	1479,0	1289,0	1266,0	1573,0	871,0	552,0	1412,0
10	1435,0	1731,0	1427,0	906,0	1810,0	1543,0	1114,0	1234,0	1661,0	1238,0	1127,0	1420,0

где M — баллы оценки состояния пациента (от 0 до 100); $\Sigma_{\text{баллов по параметрам}}$ — сумма баллов по параметрам,

значения по которым заданы пользователем исходя из балльной оценки интервала, в который попало значение параметра (или исходя из номера интервала для строковых параметров); $\Sigma_{\text{максимальных баллов по параметрам}}$ —

сумма максимально возможных баллов по параметрам, значения по которым заданы пользователем.

Исходя из данной формулы расчета баллов оценки состояния пациента получаем возможные значения, принадлежащие интервалу от 0 до 100. Значению 100 соответствует наиболее тяжелое состояние пациента.

При расчете показаний по шкале учитываются только те параметры, значение которых задан пользователем, остальные параметры исключаются из расчета — это позволяет оценивать состояние пациента по ограниченному набору параметров. Данное свойство шкалы выгодно отличает ее от других подобных шкал (например, APACHE 2 и SAPS 2), где в формуле для расчета используются баллы по всем показателям, даже если данные по ним не заданы, что приводит к искажению оценки состояния пациента в случае, если заданы не все значения показателей шкалы.

5. Описание работы программного обеспечения

Для визуализации состояния пациентов и хранения данных по пациентам реализовано программное обеспечение "Шкала оценки печеночной недостаточности".

Все данные, необходимые для работы программного обеспечения, хранятся в базе данных. В схему исходной реляционной базы введены дополнительные отношения "Диапазоны анализов пациентов", "Шкала" и "Баллы".

Диапазоны анализов пациентов

ID пациента	ID анализа	Дата взятия	№ диапазона
-------------	------------	-------------	-------------

Шкала

ID анализа	№ диапазона	Описание диапазона	Балл
------------	-------------	--------------------	------

Баллы

ID пациента	Дата взятия	Расчетный показатель по шкале
-------------	-------------	-------------------------------

Отношение "Шкала" содержит данные для оценки состояния пациентов в баллах (от 0 до 100). В отношении "Шкала" хранятся разбиение на диапазоны по всем показателям и баллы для каждого из диапазонов. Данное отношение заполняется после выполнения анализа данных с использованием меры Кульбака на обучающих выборках. При формировании диапазона в шкале учитываются

экспериментальные данные и экспертная оценка специалиста, поэтому поле "Описание диапазона" заполняется вручную при задании шкалы. Поскольку ввод числового значения каждого анализа пациента трудоемок и для расчета показателя по шкале достаточно номера диапазона, в который попадает значение анализа, то в схему добавляется отношение "Диапазоны анализов пациентов". Это отношение предназначено для хранения диапазонов, в которые попадают значения анализов пациентов на определенную дату. Для того чтобы многократно не пересчитывать значение расчетного показателя по шкале, в схему добавляется отношение "Баллы", содержащее вторичные расчетные данные. В данном отношении хранятся расчетные показатели состояния пациентов на определенную дату.

Функционально работа с программой разделена на четыре экрана.

Экран "Выбор пациента" предназначен для ввода в программу данных нового пациента, изменения данных о пациенте, удаления данных о пациенте и выбора пациента из списка пациентов, ранее введенных в программу.

После того как пациент выбран, возможен переход на экраны "Ввод данных" и "График".

На экране "Ввод данных" осуществляется ввод новых данных по пациенту на определенные дату и время по 27 показателям, используемым для оценки состояния пациента. При вводе данных осуществляется расчет баллов по шкале оценки печеночной недостаточности по выбранному пациенту в соответствии с алгоритмами, заложенными в логику работы программы, и баллами по применяемым для оценки параметрам, хранящимся в базе данных. Также на экране "Ввод данных" осуществляется просмотр и редактирование архивных данных пациента.

На экране "График" осуществляется построение графика на основании всех введенных по пациенту данных. На данном экране можно в динамике отследить изменение состояния пациента.

Заключение

Разработанное программное обеспечение "Шкала оценки печеночной недостаточности" позволяет в динамике оценивать состояние пациентов с заболеваниями печени и визуализировать это состояние, получить заключение об эффективности применяемого метода лечения.

Данное программное обеспечение предполагает возможность дальнейшего развития и модификации шкалы оценочных параметров для расширения области применения программы и приведения шкалы в соответствие потребностям медицинско-

го учреждения, в котором будет использоваться программное обеспечение.

Разработанный программный продукт успешно применяется в Государственном учреждении здравоохранения Омской области "Клинический медико-хирургический центр Министерства здравоохранения Омской области" и Федеральном государственном учреждении "Западно-Сибирский медицинский центр Российского здравоохранения". На разработанный программный продукт получено свидетельство об отраслевой регистрации разработки [14].

Список литературы

1. **Harinarayan V., Rajaraman A., Ullman J. D.** Implementing Data Cubes Efficiently // SIGMOD Conference. Montreal, CA. 1996. P. 205—216.
2. **Gray J., Chaudhuri S., Bosworth A.** etc. Data Cube: A Relational Aggregation Operator Generalizing Group-By, Cross-Tab, and Sub-Totals // Data Mining and Knowledge Discovery. 1997. N 1. P. 29—53.
3. **Педерсен Т. Б., Йенсен К. С.** Технология многомерных баз данных // Открытые системы. 2002. № 1. С. 45—50.
4. **Parsaye K.** Surveying Decision Support: New Realms of Analysis // Database Programming and Design. 1996. — N 4. — P. 26—33.
5. **Parsaye K.** OLAP and Data Mining: Bridging the Gap // Database Programming and Design. 1997. N 2. P. 30—37.
6. **Перевозчикова О. Л., Тульчинский В. Г.** Структурный анализ и синтез маршрутных схем для задач Masgo Mining // Кибернетика и системный анализ. 2003. N 1. С. 82—95.
7. **Armstrong R.** Seven Steps to Optimizing Data Warehouse Performance // Computer. 2001. V. 34, N 12. P. 76—79.
8. **Зыкин С. В.** Формирование гиперкубического представления реляционной базы данных // Программирование. 2006. № 6. С. 348—354.
9. **Александрович Ю. С., Гордеев В. И.** Оценочные и прогностические шкалы в медицине критических состояний: Справочник. СПб.: Сотис, 2007. 137 с.
10. **Battista S., Bar F., Pollet C., Bucchi M. C., Torchio M., Marzuoli M., Pagnir R., Molino G.** HEPAScore: un sistema di supporto alla decisione per l'identificazione, la stadiazione clinica e la valutazione funzionale delle epatopatie // Ann Ital Med Int. 1999. N 14. P. 20—28.
11. **Adams L. A., Bulsara M., Rossi E., DeBoer B., Speers D., George J., Kench J., Farrell G., McCaughan G. W., Jeffrey G. P.** Hepascore: an accurate validated predictor of liver fibrosis in chronic hepatitis C infection // Clin Chem. 2005. N 51. P. 1867—1873.
12. **Полуянов А. Н.** Автоматизация формирования гиперкубического представления данных // Системы управления и информационные технологии. 2008. 2.2 (32). С. 289—293.
13. **Бронштейн И. Н., Семендяев К. А.** Справочник по математике для инженеров и учащихся вузов. — М.: Наука, 1981. 720 с.
14. **Ревзин А. И., Полуянов А. Н., Чернышев А. К., Зыкин С. В., Шутов В. Ю.** Шкала оценки печеночной недостаточности // Отраслевой фонд алгоритмов и программ. Свидетельство об отраслевой регистрации разработки № 12149 от 12.01.2009 "Национальный информационный фонд документов" ГКЦИТ Федерального агентства по образованию. Номер государственной регистрации: 50200900155 от 16.01.2009.

НОВАЯ КНИГА ПО ФИЛОСОФСКИМ ПРОБЛЕМАМ ИНФОРМАТИКИ

Колин К. К. Философские проблемы информатики. — М.: БИНОМ, 2010. — 270 с.

В книге рассматриваются история зарождения и развития информатики как науки, эволюция ее предметной области, современные философские и научно-методологические проблемы, а также перспективы развития информатики как самостоятельной отрасли науки. Книга предназначена для научных работников, преподавателей, аспирантов и студентов высших учебных заведений, а также для тех категорий читателей, которые интересуются состоянием, проблемами и перспективами развития современной науки.

Она может быть также использована аспирантами и соискателями ученой степени кандидата наук при их подготовке к сдаче кандидатского экзамена по курсу "История и философия науки".

Основные разделы книги:

- Глава 1. Краткая история развития информатики.
- Глава 2. Информатика в современной системе наук.
- Глава 3. Философские проблемы информатики.
- Глава 4. Новая информационная реальность и социальная информатика.
- Глава 5. Информатика и образование.
- Глава 6. Перспективные направления развития информатики.

Книга содержит достаточно большой список литературы для самостоятельного изучения, а в приложениях приведены программы учебных курсов "Философские проблемы информатики" и "Социальная информатика" для аспирантов и магистров вузов.

Более подробная информация о книге приведена на сайте издательства "БИНОМ. Лаборатория знаний" (<http://www.lbz.ru>), а также на сайтах Института проблем информатики РАН ([www/ipiran.ru](http://www.ipiran.ru)) и Научно-образовательного центра "Информационное общество" Челябинской государственной академии культуры и искусств (<http://noc.chagi.ru>).

УДК 546:025.4.03

Н. Н. Киселева, д-р хим. наук, зав. лаб.,
В. А. Дударев, канд. техн. наук, ст. науч. сотр.
Учреждение Российской академии наук
Институт металлургии и материаловедения
им. А. А. Байкова РАН
e-mail: kis@imet.ac.ru

База данных "Информационные ресурсы неорганической химии и материаловедения"

Разработана база данных IRIC — *Information Resources of Inorganic Chemistry*, содержащая информацию о компьютерных информационных ресурсах в области неорганической химии и материаловедения, доступная пользователям глобальной сети Интернет (<http://iric.imet-db.ru>). БД содержит информацию об организациях-владельцах (название, страна, почтовый и электронный адреса, телефон и т. д.), ключевые слова, по которым возможен поиск, а также ссылки на публикации. В настоящее время БД содержит сведения о более 100 информационных системах. БД IRIC предназначена для навигации специалистов в море информационных ресурсов в вышеуказанных предметных областях. БД IRIC имеет две версии — русскоязычную и англоязычную.

Ключевые слова: база данных, информационные ресурсы в области неорганической химии и материаловедения, Интернет

Введение

Перед любым исследователем рано или поздно возникает проблема получения информации, необходимой для проведения научных работ. Для ее решения в настоящее время в мире созданы и эксплуатируются множество баз данных (БД) по свойствам неорганических веществ и материалов (БД СНВМ) [1–8]. Только за пять последних лет число БД в этих предметных областях возросло более чем в 2 раза. Большинство разработанных систем являются фактографическими, хотя объем информации в области неорганических веществ и материалов, содержащейся в документальных БД (*Science-direct* (<http://www.sciencedirect.com>), *Wiley* (www.wiley.com), *Springer* (www.springer.com), *American*

Chemical Society (www.acs.org), *Chemical Abstract Service* (www.cas.org), *STN* (www.stn-international.de), *ВИНИТИ* (www2.viniti.ru), *e-library* (www.elibrary.ru) и т. д.), значительно превосходит суммарный объем данных фактографических систем.

Очевидно, что ни одна из созданных в мире БД СНВМ не содержит полного описания свойств веществ и материалов. На современном этапе развития науки и промышленности зачастую только исчерпывающая характеристика материала позволяет принять решение об его использовании, поскольку специалистам в области материаловедения требуется знание самых разных параметров веществ. Однако всесторонний анализ свойств того или иного вещества нередко является затруднительным по причине разбросанности информации по различным источникам. Таким образом, для получения полной информационной картины нередко приходится проводить поиск сразу в нескольких БД СНВМ, содержащих разрозненные данные. Традиционной проблемой последних лет стала проблема поиска БД, содержащих нужные для специалистов данные. Для решения этой проблемы в Институте металлургии и материаловедения им. А. А. Байкова РАН (ИМЕТ РАН) разработана справочная БД "Информационные ресурсы неорганической химии и материаловедения" — *Information Resources of Inorganic Chemistry* (IRIC).

1. Структура базы данных IRIC

При разработке концептуальной структуры БД (рис. 1) учитывались возможные запросы потенциальных пользователей. Основное внимание было уделено формированию двуязычного (на русском и английском языках) постоянно пополняемого списка ключевых слов, а также уточнению информации об организациях-владельцах БД (название, почтовый и электронный адреса, телефон, факс, *e-mail*). Большинство данных было извлечено из публикаций и Интернета. В результате нормализации структуры БД были сформированы 13 реляционных таблиц (рис. 1).

Основная таблица БД IRIC — **Databases** — содержит идентификатор БД, полное и сокращенное названия БД СНВМ и комментарии на русском и английском языках, номера телефонов и факсов, *e-mail* организации-владельца БД, а также ее адрес в сети Интернет. Таблица **KeywordsInfo**

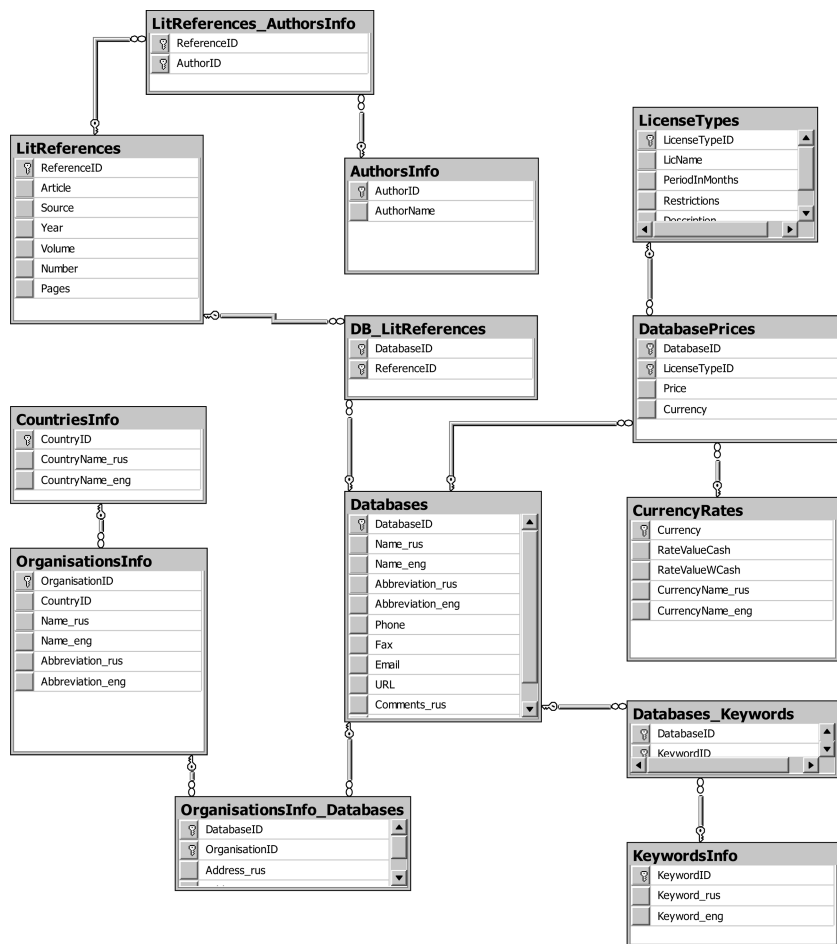


Рис. 1. Структура БД IRIC

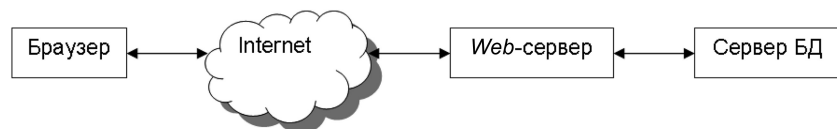


Рис. 2. Трехзвенная архитектура построения Web-приложений

включает список ключевых слов, отражающих содержание БД СНВМ, на русском и английском языках. В таблице **CountriesInfo** дан список стран, в которых разработаны БД СНВМ, на русском и английском языках. Таблица **OrganisationsInfo** включает данные об организациях—владельцах БД СНВМ: идентификатор организации-владельца; соответствующий идентификатор страны, в которой расположена организация; полное и сокращенное названия организации на русском (для российских организаций) и английском языках. В таблице **OrganisationsInfo_Databases** дано соответствие между идентификаторами БД СНВМ и организации-владельца, а также адрес организации на английском и русском (для российских организаций) языках. Соответствие между идентификаторами БД и ключевыми словами устанавливается

таблицей **Databases_Keywords**. Таблицы **LicenseTypes**, **DatabasePrices** и **CurrencyRates** включают информацию о типе лицензий и стоимости доступа к БД СНВМ.

Таблица **LitReferences** содержит основную информацию о публикациях, из которых почерпнута информация о БД СНВМ: номер статьи, ее название; название журнала (сборника, трудов и т. п.); год публикации, том, номер и страницы. В таблице **AuthorsInfo** дан пронумерованный список авторов публикаций. Соответствие между номерами статей и идентификаторами авторов устанавливается таблицей **References_AuthorsInfo**, а между номерами статей и идентификаторами БД СНВМ — таблицей **DB_LitReferences**.

Разработанная структура (рис. 1) позволяет обработать все возможные запросы пользователей, упрощает взаимодействие с БД и гарантирует непротиворечивость хранящихся данных. Ссылочная целостность обеспечивается использованием первичных и внешних ключей.

2. Пользовательский интерфейс БД IRIC

Для обеспечения наиболее простого доступа к БД IRIC со стороны конечных пользователей был разработан динамический Web-сайт, предоставляющий удобный пользовательский интерфейс для работы с базой данных. Сайт выполнен в качестве Web-приложения по классической трехзвенной архитектуре (на основе концепции

"клиент-сервер"), включающей в себя три звена: браузер (клиент); Web-сервер (сервер приложений); сервер БД (рис. 2). БД IRIC доступна пользователям через сеть Интернет: <http://iric.imet-db.ru>. Для поиска информации в БД пригоден любой браузер. Информация предоставляется в открытом доступе на русском и английском языках, что открывает возможность использования БД широким кругом специалистов не только в нашей стране, но и за рубежом.

Просмотр и поиск информации о материологических ресурсах в БД IRIC осуществляется с помощью трех классов запросов:

- общий запрос — просмотр всей информации по выбранному разделу меню;
- простой запрос — поиск данных по одному критерию;

Google™ This page is in Russian. Translate it using Google Toolbar? [Learn more](#) Translate

IRIC CHEMISTRY DATABASES
INFORMATION RESOURCES ON INORGANIC CHEMISTRY

Базы данных

Разработчики БД

Ключевые слова

Библиография

Сложные запросы

Карта сайта

Базы данных

Организация-разработчик: Показать все БД

Страна:

Название:

Аббревиатура:

Список баз данных
Фильтр по организации-разработчику: Институт металлургии и материаловедения им.А.А.Байкова РАН (ИМЕТ РАН).
Фильтр по стране: Россия.

№	Название (аббревиатура)	Контакты	Комментарий
1	БД по свойствам неорганических соединений (Фазы) [подробнее...]	Сайт: http://phases.imet-db.ru Тел: +74991352591 Факс: +74991358680 e-Mail: kis@ultra.imet.ac.ru	Термодинамические, кристаллографические и сверхпроводящие свойства неорганических соединений, образованных тремя и четырьмя элементами.
2	БД по фазовым диаграммам систем с полупроводниковыми фазами (Диаграмма)	Сайт: http://diag.imet-db.ru Тел: +74991352591	Фазовые диаграммы систем с полупроводниковыми соединениями, кристаллографические и термодинамические свойства фаз.

Рис. 3. Пример сложного запроса к БД IRIC

- сложный запрос — поиск данных по нескольким критериям.

Примером сложного запроса является поиск материаловедческих БД, разработанных в определенной стране и организации. На рис. 3 показана часть результатов запроса для России и организации-разработчика ИМЕТ РАН.

3. БД IRIC как инструмент анализа основных тенденций в разработке информационных ресурсов в области неорганического материаловедения

Основная задача создания и эксплуатации баз данных — информационное обслуживание специалистов. Однако БД предоставляют значительно более широкие возможности оперирования информацией. Одной из таких возможностей является статистический анализ и выявление основных тенденций развития предметной области (в данном случае — информационные ресурсы неорганической химии и материаловедения).

Анализ тематики ключевых слов позволил выделить наиболее распространенные параметры неорганических веществ и материалов, информация о которых содержится в БД СНВМ (рис. 4, см. третью сторону обложки). Традиционное первое место по числу БД занимают фактографические информационные системы, аккумулирующие информацию по термодинамическим и теплофизическим свойствам. Интенсивно разрабатываются и широко используются в фундаментальных и прикладных исследованиях и промышленности БД с кристаллографической и кристаллохимической

информацией. В последние годы увеличилось число БД, содержащих сведения о механических свойствах (прочности, усталости, ползучести и т. д.) и электрических свойствах неорганических веществ и материалов.

Число компьютерных информационных ресурсов и их объем могут служить показателем научного и промышленного потенциала страны. Анализ информации БД IRIC позволил выявить страны-лидеры в области разработки материаловедческих информационных ресурсов. Мировой лидер — США — занимает первое место в списке разработчиков БД СНВМ (рис. 5, см. третью сторону обложки), главным образом за счет развития информационных систем *National Institute of Standards and Technology*. В последнее десятилетие Япония неуклонно наращивает объем информационных ресурсов (следует отметить вклад *National Institute of Materials Science* — основного японского провайдера в области материаловедческих БД). В тройку лидеров входит пока и Россия, главным образом за счет БД СНВМ академических институтов (ИМЕТ РАН, ИВТАН и т. д.).

Анализ информации БД IRIC, в том числе и предоставляемых ею полных текстов статей, позволил выявить и основные тенденции в области разработки и эксплуатации современных БД по свойствам неорганических веществ и материалов:

- доступ к информации из сети Интернет, который позволяет "доставить" необходимую и самую "свежую" информацию непосредственно на рабочее место химика или материаловеда;
- экспертная оценка хранящейся информации, для которой привлекаются высококвалифици-

- рованные специалисты, что дает в руки пользователя не просто "сырую" информацию, а рекомендуемые значения;
- оснащение БД СНВМ средствами анализа информации — от традиционных термодинамических расчетов и статистических процедур до современных средств поиска закономерностей в данных, позволяющих прогнозировать поведение объектов и обеспечивающих принятие решений;
 - интеграция БД по веществам и материалам в целях предоставления пользователю наиболее полной информации о свойствах конкретного вещества, а также для последующего анализа совокупной информации о веществах и материалах.

Заключение

Базы данных по свойствам неорганических веществ и материалов являются высокотехнологичными продуктами современной науки. Неслучайно наибольшее число БД СНВМ разработано и эксплуатируется в наиболее развитых странах: США, Японии и других передовых государствах. Разработка БД СНВМ стимулируется следующими факторами: стремлением сократить затраты на необоснованное дублирование исследований и поиск информации; необходимостью получения достоверных и полных данных о свойствах веществ и материалов, а также развитием современных информационных технологий и, в первую очередь, Интернета.

В последние 10 лет резко возросло число БД для химиков и материаловедов, и этот процесс ускоряется с каждым годом. В связи с этим разработанная БД IRIC является удобным навигатором в море информационных ресурсов в области неорганической химии и материаловедения.

Созданная БД IRIC позволяет материаловедам не только получать информацию о существующих в мире БД по свойствам неорганических веществ

и технологиям их получения, но и проводить поиск таких БД по заданным ключевым словам. В последние годы наблюдается тенденция к кооперации в разработке БД и к интеграции уже созданных БД как на национальном, так и на международном уровнях. Это вызвано стремлением устранить дублирование работ и уменьшить затраты на разработку и поддержание БД. Создание БД IRIC как информационного ресурса, содержащего информацию о материаловедческих БД, является важным шагом на пути к созданию единой интегрированной материаловедческой информационной системы.

Работа выполнена при частичной финансовой поддержке РФФИ, гранты № 09-07-00194, 08-07-00437 и 09-01-12060.

Список литературы

1. Киселева Н. Н. Компьютерное конструирование неорганических соединений. Использование баз данных и методов искусственного интеллекта. М.: Наука, 2005. 288 с.
2. World Databases in Chemistry (World Databases Series.) / Ed. C. J. Armstrong. New Jersey: Bowker-Saur, 1996. 1200 p.
3. Baba T., Yamashita Y., Nagashima A. Function Sharing and Systematic Collaboration between a Networking Database System and Printed Media on Thermophysical Properties Data // J. Chem. and Eng. Data. 2009. V. 54, N 9. P. 2745—2757.
4. Земсков В. С., Кузнецов Ф. А., Уфимцев В. Б. Банки данных по полупроводниковым и другим материалам электронной техники и процессам их получения // Изв. вузов. Материалы электронной техники. 1998. № 3. С. 13—16.
5. Spencer P. J. Development of thermodynamic databases and their relevance for the solution of technical problems // Z. Metallk. 1996. Bd. 87, H. 7. S. 535—539.
6. Eriguchi K., Shimura K. Factual Databases for Materials Design and Manufacturing // ISIJ Int. 1990. V. 30. N 6. P. 409—416.
7. Kim B. K., Rhyim Y. M. Materials database project in Korea // Proc. Int. Symp. on Materials Database (MITS 2008). July 17—18, 2008. Tsukuba, NIMS. 2008. P. 68—96.
8. Begley E. F., Dapkunas S. J. A Guide to Locating and Accessing Computerized Numeric Materials Databases // J. Materials Engineering and Performance. 1993. V. 2. N 6. P. 881—886.
9. Дударев В. А. База данных по информационным ресурсам в области неорганического материаловедения // VI Российская ежегодная конференция молодых научных сотрудников и аспирантов: сборник статей. М.: Интерконтакт-Наука, 2009. С. 127—129.

УДК 519.72

В. Б. Вяткин, канд. техн. наук,
г. Екатеринбург,
e-mail: vbvzbv@yandex.ru

Введение в синергетическую теорию информации

С помощью синергетического подхода к определению количества информации рассмотрены особенности отражения дискретных систем через свои части и их взаимоотношения с окружающей средой. Получены новые информационные функции и показана их взаимосвязь с традиционными мерами информации. Установлены структурный закон сохранения суммы хаоса и порядка, информационный закон отражения и закон сохранения информации.

Ключевые слова: информация, негэнтропия, энтропия, отражение, система, часть, структура, хаос, порядок

Введение

Ранее [1] автором был представлен новый — синергетический — подход к определению количества информации, в котором за информацию принимаются сведения о конечном множестве как едином целом, а мерой информации является средняя длина интегративного кода элементов. При этом были получены формула количества информации (I_A), самоотражаемой произвольным конечным множеством A , и формула негэнтропии отражения (I_{AB}), т. е. количества информации, которую отражают (воспроизводят) друг о друге как о целостном образовании два пересекающихся конечных множества A и B :

$$I_A = \log_2 M_A; \quad (1)$$

$$I_{AB} = \frac{M_K^2}{M_A M_B} \log_2 M_K, \quad (2)$$

где $K = A \cap B$; M_A , M_B , M_K — число элементов в составе множеств A , B , K .

В настоящей статье дается дальнейшее развитие синергетического подхода к определению количества информации, которое позволяет говорить о том, что нами разрабатывается новое направление в исследовании количественных аспектов фе-

номена информации, именуемое как *синергетическая теория информации*. В нижеследующем изложении с помощью формул (1) и (2) рассматриваются информационные особенности отражения множества A через совокупность пересекающихся

с ним множеств $B_1, B_2, \dots, B_N$, таких что $\bigcup_{i=1}^N B_i \supseteq A$

и $\bigcap_{i=1}^N B_i = \emptyset$. При этом указанные множества выступают в качестве автономных систем, выделенных в составе некоторой системы D по единичному значению признака P_A и совокупности значений признака $P_B = P_{B_1}, P_{B_2}, \dots, P_{B_N}$ соответственно.

Также предварительно отметим, что отражаемая система A по отношению к окружающей среде (элементам дополняющего множества $D \setminus A$) может быть как закрытой, так и открытой в информационном отношении. То есть система A информаци-

онно закрыта, если $\bigcup_{i=1}^N B_i = A$ (рис. 1, а), и информационно открыта, если $\bigcup_{i=1}^N B_i \supset A$ (рис. 1, б).

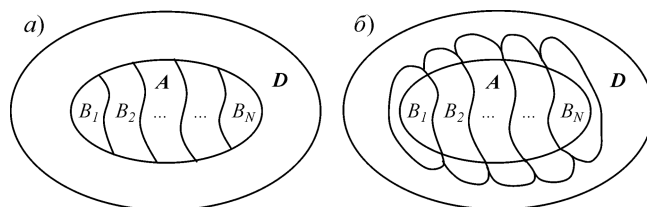


Рис. 1. Отражение системы A через совокупность множеств $B_1, B_2, \dots, B_N$:

а — система A информационно закрыта; б — система A информационно открыта

С учетом сделанных замечаний перейдем к непосредственному рассмотрению особенностей отражения системы A через совокупность множеств (систем) $B_1, B_2, \dots, B_N$ и начнем со случая, когда система A информационно закрыта.

Отражение информационно закрытых систем

Первое, что выступает перед нами при анализе информационных аспектов отражения системы A , — это информация I_A , отражаемая системой о самой себе как о целостном образовании. Посредством множеств $B_1, B_2, \dots, B_N$ система делится на N частей

с числом элементов $M_{B_1}, M_{B_2}, \dots, M_{B_N}$. Очевидно, что каждая часть отражает о системе как о целом определенную информацию, численно равную негэнтропии отражения (2). Соответственно, общее количество информации о системе A , отражаемое (отраженное) через совокупность ее частей, равно аддитивной негэнтропии отражения (I_Σ), формула которой с учетом того, что в условиях информационной закрытости системы $M_{B_i} = M_{K_i}$, имеет вид

$$I_\Sigma = \sum_{i=1}^N I_{AB_i} = \sum_{i=1}^N \frac{M_{K_i}}{M_A} \log_2 M_{K_i}. \quad (3)$$

Сравним между собой отражаемую (I_A) и отраженную (I_Σ) информации, для чего правые части выражений (1) и (3) умножим на M_A и разделим на $\log_2 M_A$. В результате этой операции для $N > 1$ получаем очевидное неравенство

$$M_A > \sum_{i=1}^N M_{K_i} \frac{\log M_{K_i}}{\log_2 M_A},$$

из которого следует, что $I_A > I_\Sigma$. Иначе говоря, *не вся информация о системе A как едином целом отражается через совокупность ее частей, и всегда существует некоторая часть отражаемой информации I_A , которая остается неотраженной*. Эта неотраженная информация характеризует неопределенность, неадекватность отражения системы через свои части, что в соответствии с распространенной интерпретацией термина *энтропия* как меры неопределенности чего-либо позволяет назвать ее *энтропией отражения*. Обозначим энтропию отражения символом S и определим ее величину как разность между отражаемой и отраженной информациями:

$$S = I_A - I_\Sigma. \quad (4)$$

В соответствии с формулами (1) и (3) запишем:

$$S = \log_2 M_A - \sum_{i=1}^N \frac{M_{K_i}}{M_A} \log_2 M_{K_i}. \quad (5)$$

Умножив и разделив аргумент второго логарифма в выражении (5) на M_A и заменяя при этом логарифм произведения суммой логарифмов, получаем:

$$S = \log_2 M_A - \sum_{i=1}^N \frac{M_{K_i}}{M_A} \log_2 \frac{M_{K_i}}{M_A} - \log_2 M_A \sum_{i=1}^N \frac{M_{K_i}}{M_A}.$$

Так как $\sum_{i=1}^N M_{K_i} = M_A$, то из последнего выражения следует:

$$S = - \sum_{i=1}^N \frac{M_{K_i}}{M_A} \log_2 \frac{M_{K_i}}{M_A}. \quad (6)$$

Отношение M_{K_i}/M_A с позиций теории вероятностей представляет собой вероятность p_i встречи элементов, обладающих i -м значением признака P_B среди общего числа элементов системы A . Поэтому формула энтропии отражения информационно закрытых систем (6) может быть представлена также в следующем виде:

$$S = - \sum_{i=1}^N p_i \log_2 p_i. \quad (7)$$

Полученная формула (7) широко известна в науке как энтропия множества вероятностей Шеннона [2] и занимает в традиционной теории информации фундаментальное положение, являясь мерой количества информации, возможности выбора и неопределенности. Данный факт говорит о том, что *вероятностная и синергетическая теории информации, имея предметом своего познания различные виды информации (вероятностный, связанный с управлением, и синергетический, существующий независимо от управления), в то же самое время непосредственно взаимосвязаны между собой отношением взаимного проникновения друг в друга*. При этом необходимо отметить следующее. В теории информации Шеннона энтропия (7) вводится в рассмотрение эмпирическим путем как функция, удовлетворяющая априорным требованиям к мере неопределенности выбора одной из N различных возможностей. В синергетической теории информации, в свою очередь, энтропия (7) получена аналитически как разность (4) между отражаемой и отраженной информациями. Иначе говоря, в синергетической теории информации энтропия (7) представляет собой меру неотраженной информации и в силу этого является вторичной, т. е. выводимой через негэнтропию отражения функцией. Сказанное, по-видимому, позволяет утверждать, что *с информационно-генетических позиций синергетическая теория информации является первичной по отношению к вероятностной теории информации Шеннона*. Примечательно, что данное утверждение полностью согласуется с выводом академика А. Н. Колмогорова о том, что "теория информации должна предшествовать теории вероятностей, а не опираться на нее" [3, с. 249].

Анализ выражений (3) и (6) показывает, что аддитивная негэнтропия I_Σ и энтропия отражения S зависят от числа частей системы и их соотношения между собой по числу элементов. При этом чем более хаотичной является структура системы, т. е. чем больше частей выделяется в ее составе и чем меньше эти части отличаются друг от друга по числу элементов, тем больше S и меньше I_Σ . И, наоборот, чем больше порядка в структуре системы, т. е. чем меньше число ее частей и чем больше части отличаются друг от друга по числу элементов, тем

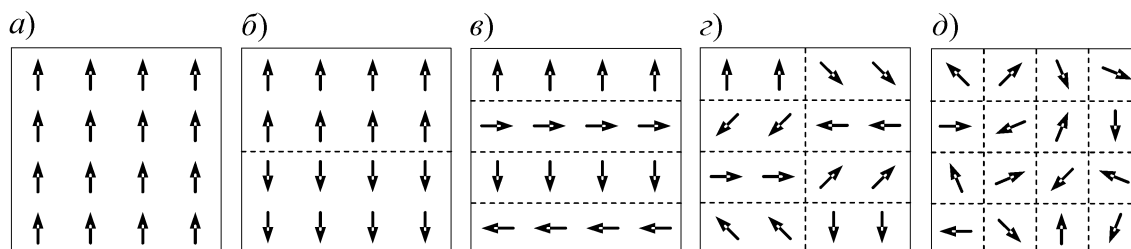


Рис. 2. Деление системы на части по признаку "направление движения элементов"

больше I_Σ и меньше S . Это означает, что I_Σ и S зависят от степени структурной упорядоченности и хаотичности отражаемой системы и, соответственно, могут рассматриваться как меры порядка и хаоса.

Проиллюстрируем сказанное примером. На рис. 2 приведены различные состояния системы, состоящей из 16 элементов, которые характеризуются признаком "направление движения". По значениям этого признака система последовательно принимает пять состояний, которым соответствует ее деление на 1, 2, 4, 8, 16 равновеликих (для простоты) частей.

В состоянии, соответствующем рис. 2, а, все элементы движутся в одном направлении, и в структуре системы наблюдается идеальный порядок. На рис. 2, д имеем полярную противоположность, т. е. каждый элемент системы обладает строго индивидуальным направлением движения и структура системы является максимально хаотичной. Состояния системы, соответствующие рис. 2, б, в, г, занимают промежуточное положение по отношению к состояниям на рис. 2, а, д и характеризуются тем, что в их структуре присутствует как хаотичность, так и упорядоченность. Рассчитывая для каждого состояния системы по формулам (3) и (6) значения аддитивной негэнтропии и энтропии отражения, последовательно получаем: $I_\Sigma = 4; 3; 2; 1; 0$; $S = 0; 1; 2; 3; 4$. То есть нарастанию хаоса и уменьшению порядка в структуре системы на рис. 2 соответствует уменьшение значений аддитивной негэнтропии и увеличение значений энтропии отражения.

Аддитивная негэнтропия I_Σ как отраженная информация и энтропия отражения S как неотраженная информация являются составными частями отражаемой информации I_A и

$$I_A = I_\Sigma + S. \quad (8)$$

В то же самое время, как было показано, I_Σ и S количественно характеризуют структуру отражаемой системы со стороны ее упорядоченности и хаотичности. Поэтому равенство (8) интерпретируется также в виде следующего выражения:

$$\text{порядок} + \text{хаос} = \text{const.}$$

То есть можно констатировать, что в лице равенства (8) мы имеем закон сохранения суммы хаоса и

порядка, который справедлив при любых структурных преобразованиях системы. Иначе говоря, чтобы мы ни делали с системой без изменения общего числа элементов, на сколько бы частей ни разбивали ее по значениям какого-либо признака и в каком бы соотношении по числу элементов ни находились между собой части, сумма хаоса и порядка в структуре системы всегда будет оставаться неизменной. Например, для вышеописанной системы на рис. 2 по всем состояниям имеем:

$$I_\Sigma + S = 4.$$

Хаотичность и упорядоченность структуры в своей совокупности определяют в целом структурную организацию системы, и, соответственно, для ее количественной характеристики может использоваться та или иная функция, аргументами которой являются меры хаоса и порядка. В качестве такой функции структурной организации может использоваться отношение аддитивной негэнтропии к энтропии отражения, которое будем называть R -функцией¹:

$$R = \frac{I_\Sigma}{S}. \quad (9)$$

То есть значения R -функции говорят о том, что и в какой мере преобладает в структуре системы: хаос или порядок. Так, если $R > 1$, то в структуре системы преобладает порядок, в противном случае, когда $R < 1$, — хаос. При $R = 1$ хаос и порядок уравниваются друг друга, и структурная организация системы является равновесной. Например, для состояний системы на рис. 2 последовательно имеем: $R = \infty, 3, 1, 0,33, 0$.

Рассмотрим теперь особенности взаимоотношений аддитивной негэнтропии I_Σ и энтропии отражения S при постоянном числе элементов системы M_A и различном числе ее частей N , соотносящихся между собой по числу элементов M_{B_i} произвольным образом. При этом сначала определим максимальные и минимальные значения I_Σ и S при данном N , основываясь на формулах (3) и (6) и выражении (8).

¹ Название функции дано по первой букве англ. слова *reflection*, что в переводе на русский язык означает *отражение*.

Энтропия отражения является максимальной, когда все части системы представлены одинаковым числом элементов, т. е.:

$$S^{\max} = -N \frac{1}{N} \log_2 \frac{1}{N} = \log_2 N. \quad (10)$$

Соответственно, минимальное значение аддитивной негэнтропии отражения имеет вид

$$I_{\Sigma}^{\min} = \log_2 M_A - \log_2 N = \log_2 \frac{M_A}{N}. \quad (11)$$

Максимальное значение, в свою очередь, аддитивная негэнтропия будет принимать тогда, когда число элементов в одной части равно $M_A - N + 1$, а каждая из остальных $N - 1$ частей включает в себя только один элемент, т. е.

$$I_{\Sigma}^{\max} = \frac{M_A - N + 1}{M_A} \log_2 (M_A - N + 1). \quad (12)$$

Соответственно, минимальное значение энтропии отражения

$$S^{\min} = \log_2 M_A - \frac{M_A - N + 1}{M_A} \log_2 (M_A - N + 1). \quad (13)$$

Построим графики $I_{\Sigma}^{\min}$, $I_{\Sigma}^{\max}$, $S^{\min}$, $S^{\max}$ как функций от N при фиксированном M_A и проведем анализ полученной диаграммы, которую будем называть *информационным полем отражения дискретных систем* (рис. 3).

Приведенные графики изначально образуют два контура: энтропийный — $abdefha$ и негэнтропийный, или информационный, — $cdfghbc$, которые локализуют соответствующие области всех возможных значений энтропии и аддитивной негэнтропии отражения информационно закрытых систем. Пересечение данных контуров по точкам b и f , где наблюдаются равенства $I_{\Sigma}^{\min} = S^{\max}$ и

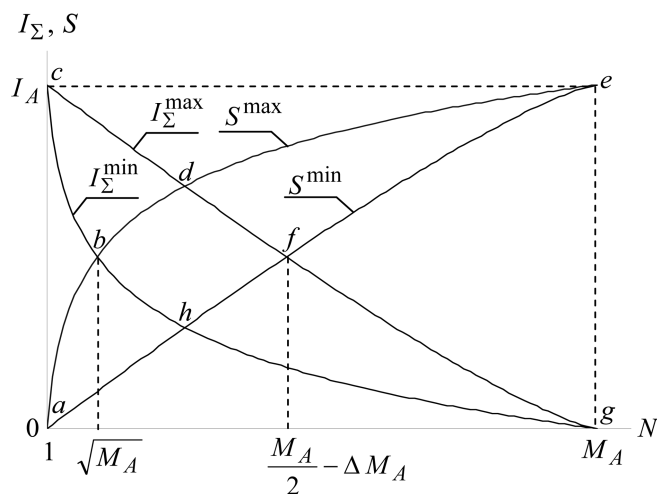


Рис. 3. Информационное поле отражения дискретных систем ($M_A = 100$)

$I_{\Sigma}^{\max} = S^{\min}$, позволяет по проекциям этих точек на горизонтальную ось выделить три интервала значений N (левый, центральный, правый) с присущими каждому интервалу особенностями взаимоотношений I_{Σ} и S . Рассмотрим эти особенности, предварительно определив значения N , соответствующие точкам b и f .

В точке b согласно выражениям (10) и (11) имеем уравнение $\log_2 N_b = \log_2 \frac{M_A}{N_b}$, решение которого

дает $N_b = \sqrt{M_A}$ (с округлением до ближнего большего целого). Аналогичное уравнение для точки f , образованное правыми частями выражений (12) и (13), не имеет аналитического решения. То есть значение N_f может быть определено только численным путем. (Для членов последовательности $N = 1, 2, \dots, M_A$ определяются $I_{\Sigma}^{\max}$ и $S^{\min}$, и в качестве значения N , соответствующего точке f , принимается то значение, после которого $S^{\min} > I_{\Sigma}^{\max}$.)

Но при этом, как показывают расчеты, значение N_f остается меньше $\frac{M_A}{2}$ на некоторую величину ΔM_A , которая увеличивается по мере роста M_A , т. е.

$$N_f = \frac{M_A}{2} - \Delta M_A.$$

Таким образом, при заданном M_A имеем три формализованных интервала значений N :

- левый: $1 \leq N < \sqrt{M_A}$;
- центральный: $\sqrt{M_A} \leq N \leq \frac{M_A}{2} - \Delta M_A$;
- правый: $\frac{M_A}{2} - \Delta M_A < N \leq M_A$.

Отражение систем, попадающих в левый интервал, характеризуется тем, что при любых соотношениях их частей между собой по числу элементов сохраняется неравенство $I_{\Sigma} > S$. То есть в левом интервале значений N структурная упорядоченность системы всегда больше ее хаотичности и $R > 1$.

В правом интервале наблюдается противоположная ситуация, когда при любых структурных преобразованиях системы выполняется неравенство $S > I_{\Sigma}$, что соответствует преобладанию хаоса над порядком. При этом $R < 1$.

В центральном интервале происходит пересечение областей возможных значений I_{Σ} и S , вследствие чего между ними здесь наблюдаются различные соотношения и проявляется тенденция к взаимному уравновешиванию структурного порядка и хаоса. R -функция при этом может быть как

больше, так и меньше единицы, а в ряде случаев (по линии bf) имеет место и равенство $R = 1$.

Иначе говоря, в левом интервале значений N наблюдается необратимое преобладание в структуре системы порядка над хаосом, а в правом интервале, наоборот, — необратимое преобладание хаоса над порядком. В центральном интервале, в свою очередь, преобладание как порядка над хаосом, так и хаоса над порядком является обратимым.

Указанные особенности интервалов значений N позволяют любые дискретные системы с конечным множеством элементов в зависимости от числа частей, на которые они разделяются в плоскости какого-либо признака, формализованно классифицировать на три типа:

- упорядоченные (левый интервал);
- синергетичные (центральный интервал);
- хаотичные (правый интервал).

Например, система на рис. 2 в различных своих состояниях будет классифицироваться как: a , b — упорядоченная; v — синергетичная; z , d — хаотичная.

Отражение информационно открытых систем

При информационном открытии системы (см. рис. 1, b) осуществляется ее непосредственная взаимосвязь (взаимодействие) с окружающей средой, в результате чего часть отражаемой информации I_A переходит в окружающую среду и распределяется по элементам дополняющих множеств $B_i \setminus K_i \neq \emptyset$. Соответственно, по отношению к информационно закрытому состоянию системы уменьшается аддитивная негэнтропия и увеличивается энтропия отражения. То есть собственно отражение системы становится более неопределенным, рассеянным, хаотичным. Так как отношение M_{K_i}/M_A не зависит от того, открыта или закрыта система, то энтропия S (6), являясь функцией внутреннего строения системы, при информационном открытии последней не изменяется, т. е. является инвариантной величиной относительно любых взаимоотношений системы с окружающей средой. Это говорит о том, что при открытии системы в общей энтропии ее отражения появляется новая составляющая (ΔS) как функция, характеризующая взаимосвязь системы с внешней средой. В соответствии со сказанным, для того чтобы отличать друг от друга энтропии S и ΔS , будем пользоваться терминами *внутренняя* (S) и *внешняя* (ΔS) *энтропии отражения*.

Определим величину ΔS исходя из соблюдения баланса между отражаемой, отраженной и неотраженной информацией. Так как отражаемая информация I_A и внутренняя энтропия S при информационном открытии системы не изменяются, то внешняя энтропия ΔS равна соответствующему

умению уменьшению аддитивной негэнтропии I_Σ информационно закрытого состояния системы:

$$\Delta S = I_\Sigma - I_\Sigma^*, \quad (14)$$

где I_Σ^* — аддитивная негэнтропия отражения информационно открытой системы.

Представляя выражение (14) в развернутом виде и проводя простые преобразования, получаем:

$$\Delta S = \sum_{i=1}^N \frac{M_{B_i} - M_{K_i}}{M_{K_i}} \frac{M_{K_i}^2}{M_A M_{B_i}} \log_2 M_{K_i}. \quad (15)$$

Первый сомножитель в выражении (15) представляет собой отношение числа элементов дополняющего множества $B_i \setminus K_i$ к числу элементов связующего множества $K_i = A \cap B_i$. Иначе говоря, этот сомножитель есть отношение области непосредственной взаимосвязи множества B_i с окружающей систему A средой к области непосредственной взаимосвязи этого же множества с самой системой A . Будем называть это отношение *коэффициентом информационной открытости системы* (μ_{B_i}) *по множеству* B_i :

$$\mu_{B_i} = \frac{M_{B_i}}{M_{K_i}} - 1, \quad 0 \leq \mu_{B_i} < \infty. \quad (16)$$

Совокупность второго и третьего сомножителей, в свою очередь, представляет собой негэнтропию отражения $I_{AB_i}^*$ открытой системы. Поэтому, с учетом (16), формула внешней энтропии ΔS имеет вид

$$\Delta S = \sum_{i=1}^N \mu_{B_i} I_{AB_i}^*. \quad (17)$$

Вернемся к рассмотрению информационного поля отражения дискретных систем, представленного на рис. 3. Очевидно, что прямоугольный контур $asega$ ограничивает в целом площадь информационного поля, включающего в себя, при данном M_A , все возможные значения аддитивной негэнтропии и энтропии отражения как открытых, так и закрытых в информационном отношении систем. В предыдущем разделе было показано, что если система является информационно закрытой, то области возможных значений аддитивной негэнтропии и энтропии отражения ограничены контурами $cdfghbc$ и $abdefha$ соответственно. Когда система информационно открывается, ее отражение становится более неопределенным и области возможных значений как аддитивной негэнтропии, так и энтропии отражения начинают увеличиваться. При этом новые возможные значения аддитивной негэнтропии располагаются ниже графика $I_\Sigma^{\min}$, а энтропии отражения — выше гра-

фика $S^{\max}$. В случае неограниченного возрастания общего информационного открытия системы, т. е. когда $\forall \mu_{B_i} \rightarrow \infty$, минимум аддитивной негэнтропии и максимум общей энтропии отражения выходят на свои асимптотические значения, лежащие на линиях ag и ce соответственно. Таким образом, комплексные области возможных значений отраженной и неотраженной информации ограничены контурами: аддитивная негэнтропия отражения — $acdfga$, энтропия отражения — $acefha$. Заканчивая описание распределения отражаемой информации по информационному полю отражения $asega$, отметим также, что в его пределах существует довольно обширная область $efge$, в которую никогда не попадают значения как отраженной, так и неотраженной информации.

Информационный закон отражения и закон сохранения информации

Составим информационный баланс отражения дискретных систем для общего случая:

$$I_A = I_{\Sigma}^* + S + \Delta S. \quad (18)$$

Выражение (18) является устойчивым однозначным соотношением между отражаемой, отраженной и неотраженной информацией, справедливым для любых дискретных систем, представленных конечным множеством элементов, вне зависимости от степени их информационной открытости, числа отражающих объектов и их соотношения между собой по числу элементов, что позволяет придать ему статус закона. Дадим этому закону название *информационный закон отражения* и сформулируем его следующим образом: *при отражении системы A через совокупность систем $B_1, B_2, \dots, B_N$,*

таких что $\bigcup_{i=1}^N B_i \supseteq A$ и $\bigcap_{i=1}^N B_i = \emptyset$, происходит разделение отражаемой информации на отраженную и неотраженную части, первая из которых равна аддитивной негэнтропии, а вторая — сумме внутренней и внешней энтропий отражения.

Информационный закон отражения по своей форме (18) подобен известным физическим законам сохранения (энергии, электрического заряда и т. д.). О необходимости установления подобного рода законов в теории информации указывалось математиками и философами еще в 50—60 гг. прошлого столетия [4, 5]. Отмечалось, что законы сохранения в теории информации должны учитывать взаимные переходы друг в друга различных видов информации. Такие переходы можно видеть при анализе выражения (18), если, в соответствии с формулой (7), рассматривать внутреннюю энтропию отражения S в качестве вероятностной

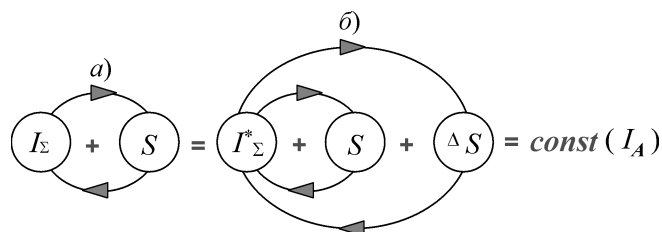


Рис. 4. Графическое выражение закона сохранения информации: а — информационно закрытая система; б — информационно открытая система

информации, а I_{Σ} , I_{Σ}^* , ΔS считать, естественно, разновидностями синергетической информации. Короче говоря, выражение (18) на межвидовом информационном уровне представляет собой *закон сохранения информации*, который можно сформулировать следующим образом: *суммарное количество синергетической и вероятностной информации, характеризующей структуру дискретной системы и ее взаимоотношения с окружающей средой, при фиксированном числе элементов системы является постоянной величиной.*

Раскроем содержание установленного закона сохранения информации, для чего представим его в иллюстративной форме (рис. 4).

Информационно закрытые системы. Когда система информационно закрыта, соотношения между различными видами информации и их взаимные переходы друг в друга (рис. 4, а) определяются тем, в каком направлении осуществляются структурные преобразования системы. Так, если число частей системы уменьшается, а тем более, если при этом какая-то часть системы начинает доминировать по числу элементов, то возрастает количество синергетической информации I_{Σ} и уменьшается количество вероятностной S . В пределе, когда все элементы системы в плоскости какого-либо признака становятся неотличимыми друг от друга и, соответственно, система не разделяется на части ($N = 1, p = 1$), ее вероятностная информация полностью переходит в синергетическую информацию и как вид перестает существовать в системе, т. е. $I_{\Sigma} = I_A, S = 0$. Разнообразия элементов и возможности выбора при этом нет. При обратных структурных преобразованиях, когда система все более делится на части, т. е. увеличивается разнообразие ее элементов и возрастают возможности выбора, идет противоположный информационный процесс и синергетическая информация переходит в вероятностную. Этот процесс исчерпывает себя и прекращается, когда число частей системы становится равным числу ее элементов ($N = M_A, p_i = 1/M_A$) и, соответственно, $I_{\Sigma} = 0, S = I_A$.

Таким образом, мы видим, что каждый из отмеченных противоположных процессов в информационно-количественном отношении заканчивается одним и тем же математическим результатом,

выражаемым чисто комбинаторной формулой количества информации. При этом, с содержательной точки зрения, в первом случае, когда уменьшается число частей и в структуре системы возрастает порядок, мы приходим к формуле количества информации, самоотражаемой конечным множеством [1], а во втором случае, при максимальной хаотичности структуры, имеем также равенство двоичному логарифму Хартли [6], который лежит в основе традиционного комбинаторного подхода к определению количества информации [7]. По-видимому, можно сказать, что *комбинаторное количество информации имеет двойственную природу и в зависимости от того, с какой стороны его рассматривать, может относиться как к синергетическому, так и к вероятностному виду информации.*

Обращаясь теперь к равенству (8), в свете сказанного можно также утверждать, что это равенство, помимо прочего, представляет собой уравнение взаимосвязи различных подходов к количественному определению информации. То есть *комбинаторный, вероятностный и синергетический подходы к определению количества информации неразрывно взаимосвязаны между собой и в своей совокупности составляют единую количественную основу общей теории информации.*

Информационно открытые системы. Когда система информационно открывается, начинается ее взаимодействие с окружающей средой, в результате чего часть информации, отражаемой системой, переходит в окружающую среду в виде внешней энтропии отражения ΔS , равной уменьшению аддитивной негэнтропии отражения (рис. 4, б). Если в последующем система начинает обратно закрываться ($\forall \mu_B \rightarrow 0$), то эта информация возвращается в систему в виде соответствующего увеличения значений I_Σ^* . Так как по определению (14) $\Delta S < I_\Sigma$, то независимо от степени увеличения открытости системы ($\forall \mu_B \rightarrow \infty$) во внешнюю среду может перейти только ограниченное количество информации. То есть можно сделать следующий вывод: *система не может отдать во внешнюю среду, а также получить обратно из внешней среды количество информации, превосходящее аддитивную негэнтропию отражения ее информационно закрытого состояния.*

Так же, как в случае информационной закрытости, рассмотрим два полярных друг другу структурных состояния системы. В первом случае, когда система предельно упорядочена ($N = 1$), в ней актуализирована только синергетическая информация и до открытия системы $I_\Sigma = I_A$, $S = 0$. После информационного открытия системы и его неограниченного возрастания по окружающей среде распространяется практически вся отражаемая системой информация и, соответственно, $I_\Sigma^* \rightarrow 0$, $\Delta S \rightarrow I_A$.

Во втором случае, при максимальной хаотичности структуры системы ($N = M_A$), вся ее информация представлена только вероятностным видом, который инвариантен относительно взаимоотношений системы с внешней средой. Соответственно, как до открытия системы, так и после ее открытия сохраняется равенство $S = I_A$, говорящее о том, что при любой степени открытости системы обмена информации с окружающей средой не происходит и, следовательно, $I_\Sigma^* = 0$, $\Delta S = 0$. То есть, достигая максимально возможной структурной хаотичности, система становится "невидимой" для внешней среды или, выражаясь по-другому, множество A в плоскости признака $P_B = P_{B_1}, P_{B_2}, \dots, P_{B_N}$ при $N = M_A$ перестает восприниматься окружающей средой как система.

Заключение

В двух статьях (настоящей и [1]) на основе простого аксиоматического базиса, без опоры на теорию вероятностей и традиционную теорию информации пройден путь от рассмотрения информационных аспектов элементарного конечного множества до установления на межвидовом информационно-количественном уровне закона сохранения информации. Получены неизвестные ранее информационно-синергетические функции и показана их непосредственная взаимосвязь с традиционными мерами информации. Все это говорит о том, что в лице синергетической теории информации мы имеем новую научную теорию. Дальнейшее развитие этой теории и внедрение ее в практику научных и прикладных исследований, по нашему мнению, будет иметь значение не только для общей теории информации, но и для тех предметных областей, где объекты познания представимы в виде дискретных систем с конечным множеством элементов.

Список литературы

1. Вяткин В. Б. Синергетический подход к определению количества информации // Информационные технологии. 2009. № 12. С. 68–73.
2. Шеннон К. Работы по теории информации и кибернетике. М.: Изд-во иностр. лит., 1963. 830 с.
3. Колмогоров А. Н. Теория информации и теория алгоритмов. М.: Наука, 1987. 304 с.
4. Харкевич А. А. Очерки общей теории связи. М.: Гостехиздат, 1955. 268 с.
5. Новик И. Б. Негэнтропия и количество информации // Вопросы философии. 1962. № 6. С. 118–128.
6. Хартли Р. В. Л. Передача информации // Сб.: Теория информации и ее приложения. М.: Физматгиз, 1959. С. 5–35.
7. Колмогоров А. Н. Три подхода к определению понятия "количество информации" // Проблемы передачи информации. 1965. Т. 1. № 1. С. 3–11.

Указатель статей, опубликованных в журнале «Информационные технологии» в 2010 г.

ОБЩИЕ ВОПРОСЫ

Гуревич И. М. Информационные методы исследования физических систем: обзор первых достижений. № 12.

ВЫЧИСЛИТЕЛЬНЫЕ СИСТЕМЫ, СЕТИ И СИСТЕМЫ СВЯЗИ

Аббасова Т. С. Повышение эффективности эксплуатации высокоскоростной кабельной системы с помощью виртуальных технологий. № 12.

Алиев И. М. Управление трафиком в транзитных BGP/MPLS-сетях. № 3.

Белянина Н. В., Прокопенко М. Н., Серовиков С. А. Выбор программно-аппаратной концепции организации распределенной системы экологического мониторинга. № 6.

Беневоленский С. Б., Нагорных Д. Н., Смирнова А. А. Особенности создания резервируемого хранилища центра обработки данных распределенных вычислительных систем. № 11.

Васенин В. А., Инюхин А. В., Шевелев М. В. Фрагмент территориально-распределенной информационно-вычислительной среды на основе методологии Grid. № 1.

Грушин А. И., Ростовцев А. В., Чинь Куанг Киен. Упаковка и распаковка потоков данных для системы ГЛОНАСС. № 3.

Давыдов А. И., Шахов В. Г., Ядрышников И. Б. Анализ абонентской нагрузки в сетях сотовой связи. № 5.

Давыдов А. И., Шахов В. Г., Ядрышников И. Б. Анализ протокола IPv6 в современных операционных системах. № 10.

Жуков О. Д. Информационные технологии на основе полиномиальной алгебры. № 7.

Захаров В. М., Столов Е. Л., Шалагин С. В. Алгоритм вычисления остатка по модулю и оценки его сложности в базе ПЛИС. № 11.

Калинина Э. В., Ретинская И. В., Серова Д. А. Вероятностные и нечеткие модели оценки качества передачи информации в телекоммуникационных сетях. № 10.

Калитин Д. В. Проектирование параллельных декомпозиций управляющих автоматов в k -значных логиках. № 7.

Корячко В. П., Шибанов В. А., Ижванов Ю. Л., Шибанов А. П. Оценка времени передачи файла с учетом старения информации. № 10.

Макошенко Д. В. Назначение переменных на регистры с помощью древовидного параметрического алгоритма раскраски графа. № 6.

Медведев Н. В., Троицкий И. И., Квасов П. М. Аналитическая модель телекоммуникационной сети нового поколения с приоритетным обслуживанием информационных сообщений. № 1.

Морев Н. В. Сравнение алгоритмов планирования распределения задач для однородных распределенных вычислительных систем. № 5.

Огнев В. А., Иванов С. Р. Повышение помехоустойчивости информационных коммуникаций между навигационными спутниками и приемовычислителями при использовании сигналов E10S и L1C. № 11.

Панько С. П., Нагорный А. В. Особенности многостанционного доступа в сеть с разделяемым ресурсом. № 12.

Пономарев В. А., Богоявленская О. Ю., Богоявленский Ю. А. Конфигурируемая модульная система мониторинга по-

ведения транспортного протокола на уровне ядра операционной системы. № 1.

Саак А. Э. Локально-оптимальный синтез расписаний для GRID-технологий. № 12.

Савкин В. Б. Имитационное моделирование механизмов справедливого распределения коммуникационных ресурсов компьютерных сетей между пользователями и приложениями. № 5.

Сотов Л. С. Комбинаторная модель функционального формирования разбиений бинарного множества. № 10.

Умбиталиев А. А., Шипилов Н. Н., Фахми Ш. С. Сложно-функциональный блок транскодирования видеоконтента. № 12.

Ховансков С. А., Норкин О. Р. Алгоритм оптимизации мультиагентной системы распределенных вычислений и его исследование с помощью программной модели. № 11.

Шахов В. Г., Коваленко О. Н. Модель распределения пропускной способности каналов мультисервисной сети между потоками с различными параметрами качества обслуживания. № 3.

Штейнберг Б. Я. Блочко-аффинные размещения данных в параллельной памяти. № 6.

ИНТЕЛЛЕКТУАЛЬНЫЕ СИСТЕМЫ

Балакирев Н. Е., Добрышина Е. В. Концептуальная модель и структура системы обработки текстовой информации. № 2.

Бельков С. А., Гольдштейн С. Л. Представление материала текстовых и гипертекстовых источников сетью паттернов. № 1.

Бородащенко А. Ю., Глотов Д. В., Бочков С. М. Алгоритм контент-анализа новостного потока RSS-каналов. № 9.

Глова В. И., Катасёв А. С., Корнилов Г. С. Преднастройка и оптимизация параметров нечеткой нейронной сети при формировании баз знаний экспертных систем. № 5.

Голомазов Д. Д. Выделение терминов из коллекции текстов с заданным тематическим делением. № 2.

Горбатов В. А., Добрик И. И. Теоретико-графовое экспертное оценивание альтернатив. № 11.

Карпов А. А., Цирульник Л. И., Железны М. Разработка компьютерной системы «говорящая голова» для аудиовизуального синтеза русской речи по тексту. № 8.

Комарцова Л. Г., Кадников Д. С., Ковалев И. В. Особенности построения гибридных интеллектуальных систем обработки информации. № 5.

Леденева Т. М., Сергиенко М. А. Организация структуры нечеткой базы правил. № 6.

Норенков И. П. Генетические алгоритмы поиска решений в онтологических базах знаний. № 9.

Норенков И. П. Интеллектуальные технологии на базе онтологий. № 1.

Норенков И. П., Уваров М. Ю. Поиск решений по прецедентам на основе кластеризации онтологий. № 8.

Попов С. И., Попов Ю. И., Шалыто А. А. Задача о муравьеде и муравьях. № 8.

Титов А. С. Построение деревьев визуализации с помощью методов гравитационной кластеризации. № 1.

Туманов В. Е. Применение искусственной нейронной сети для прогноза реакционной способности молекул в радикальных реакциях. № 5.

Чаньшев О. Г. Метод автоматической кластеризации текстов на основе анализа пересечений кластеров доминант. № 11.

Черний А. В., Тузовский А. Ф. Semantic Web масштаба организации. № 6.

Шапкин П. А. Модели и методы разработки веб-приложений на основе онтологии предметной области. № 2.

МОДЕЛИРОВАНИЕ И ОПТИМИЗАЦИЯ

- Абдуллаев Н. Т., Алиев Т. А., Дышин О. А. Алгоритм вычисления статистической оценки обобщенных размерностей множеств. № 4.
- Гальченко В. Я., Якимов А. Н., Остапушенко Д. Л. Поиск глобального оптимума функций с использованием гибрида мультиагентной роевой оптимизации с эволюционным формированием состава популяции. № 10.
- Гизатуллин З. М., Чермошенцев С. Ф. Моделирование электромагнитных помех в неэкранированной витой паре при внешнем гармоническом электромагнитном воздействии. № 6.
- Зелепухина В. А. Разработка Интернет-ориентированных виртуальных лабораторий математического моделирования посредством разделения вычислительных и визуализационных задач. № 10.
- Карпенко А. П., Мухлисулина Д. Т., Овчинников В. А. Нейросетевая аппроксимация функции предпочтений лица, принимающего решения, в задаче многокритериальной оптимизации. № 10.
- Карпенко А. П., Селиверстов Е. Ю. Глобальная оптимизация методом роя частиц. Обзор. № 2.
- Клосс Ю. Ю., Мартынов Д. В., Черемисин Ф. Г. Разработка методов компьютерного моделирования и анализа микронасоса Кнудсена. № 10.
- Ковель И. В. Вероятностный метод построения окрестности статистического глобального оптимума NPC-задач. № 2.
- Лиманова Н. И., Мамзин Е. А. Дискретная математическая модель детерминированного клеточного автомата и ее программная реализация. № 2.
- Май В. П., Мукомел Д. В. Моделирование поверхностного водостока с реализацией параллельных вычислений на многопроцессорной системе. № 6.
- Месягутов М. А. Задача двухмерной ортогональной упаковки: поиск нижней границы на базе решения одномерной продолженной упаковки. № 6.
- Мухачева Э. А., Валеев Р. С. Локальный поиск размещения товарных позиций на базе анализа их номенклатурной принадлежности. № 6.
- Мухачева Э. А., Валиахметова Ю. И., Хасанова Э. И., Телицкий С. В. Проектирование размещения ортогональных объектов на полигонах с препятствиями. № 10.
- Переварюха А. Ю. Новый метод компьютерного моделирования режима переходного хаоса. № 2.
- Соболев С. С., Сотов Л. С., Харин В. Н. Алгоритм работы и модель функционального генератора перестановок. № 4.
- Тарнавский Г. А. Облачные вычисления: технология «Data Files Cruise» организации информационных потоков на портале SciShop.ru. № 6.

СИСТЕМЫ АВТОМАТИЗИРОВАННОГО ПРОЕКТИРОВАНИЯ

- Бабак Л. И., Дорофеев С. Ю., Песков М. А., Черкашин М. В., Шеерман Ф. И., Барышников А. С., Кошевой С. Е. Интеллектуальная система автоматизированного проектирования СВЧ устройств INDESYS. № 2.
- Волокоморов С. В., Карпенко А. П., Мартынюк В. А. Анализ кинематики параллельных механизмов средствами системы автоматизированного проектирования САПР. № 11.
- Жук Д. М., Маничев В. Б., Ильинский А. О. Методы и алгоритмы решения дифференциально-алгебраических уравнений для моделирования систем и объектов во временной области. № 7, 8.

- Кузьмин А. А., Норенков И. П. Автоматизация проектирования мнемосхем электрических сетей на диспетчерских щитах управления. № 11.
- Перепелкин Е. А., Поддубный В. И., Варкентин А., Ган М. Применение прикладного пакета CAMeL-Veiv для моделирования управляемого движения колесного трактора. № 7.
- Разумовский А. И., Сизова Л. Н. Проектирование и трассировка печатных плат с использованием программного комплекса «Графика-ТР». № 2.
- Черемисинова Л. Д. Оценка энергопотребления КМОП-схем на логическом уровне. № 8.

ОРГАНИЗАЦИОННО-ИНФОРМАЦИОННЫЕ СИСТЕМЫ

- Богданов Ю. М., Карнаух Д. И. Информационное обеспечение комплексных процессов государственного управления. № 11.
- Карнаух Д. И. Разработка единой технологической платформы как ключевой инфраструктурный элемент построения электронного правительства. № 8.

ГЕОИНФОРМАЦИОННЫЕ СИСТЕМЫ

- Андреев А. К. Компьютерная поддержка принятия решений при угрозе цунами. № 11.
- Антонченков А. А. Задачи трехмерной визуализации в ГИС и программный прототип для их решения. № 3.
- Бобков В. А., Морозов М. А. Векторизация фотоизображений городских сцен для 3D-реконструкции. № 3.
- Богданов М. Р., Богданова А. В., Габидуллин Ю. З. Беспроводные сенсорные сети в экологии. № 7.
- Бушмелева К. И., Плюснин И. И., Увайсов С. У. Модель мобильного устройства дистанционного зондирования магистрального газопровода. № 3.
- Гальяно Ф. Р. Алгоритм классификации участков поверхности Земли на основе сингулярного разложения матриц. № 12.
- Косников Ю. Н. Геометрическое моделирование и отображение местности по топографическим данным. № 7.
- Немтинов В. А., Горелов А. А., Немтинов К. В., Кулин Д. В. Пространственно-временная модель объекта культурно-исторического наследия. № 7.
- Струченков В. И., Козлов А. Н. Кусочно-линейная аппроксимация плоских кривых при наличии ограничений. № 12.
- Филиппова А. С., Кузнецов В. Ю., Басимов Э. Р. Моделирование препроцессора в задачах оптимального размещения сенсоров в области мониторинга. № 11.

ИНФОРМАЦИОННО-ИЗМЕРИТЕЛЬНЫЕ СИСТЕМЫ

- Бушмелева К. И., Яценко Е. А. Специализированный пакет программ для калибровки измерительных приборов. № 10.
- Владимирский Э. И., Тагиев Ф. К. Синергетический подход формирования интегральных размерностей в интеллектуальных информационно-измерительных системах. № 6.
- Колесникова С. И., Лаходынов В. С., Цой Ю. Р. Исследование качества распознавания состояний стохастической системы. № 6.
- Курсков С. Ю., Ершова Н. Ю., Мурсалимов О. А. Организация удаленного доступа к микроконтроллерам MCS-96 на базе распределенной информационно-измерительной и управляющей системы поддержки научно-образовательного процесса. № 10.

БЕЗОПАСНОСТЬ ИНФОРМАЦИИ И УПРАВЛЕНИЕ РИСКАМИ

- Алгулиев Р. М., Имамвердиев Я. Н., Деракшанде С. А. Пути повышения точности методов оценки рисков информационной безопасности. № 12.
- Амербаев В. М., Максименко А. В. Модулярные рюкзачные преобразования в информационных технологиях. № 5.
- Аникин И. В. Метод количественной оценки уровня ущерба от реализации угроз на корпоративную информационную сеть. № 1.
- Васинев Д. А., Жусов Д. Л., Цынгаев М. С., Ватулин И. Ю. Предложения по повышению защищенности Ethernet-сетей от атак канального уровня. № 9.
- Генералов Д. Н. Противодействие нарушению целостности модулей специализированного программного обеспечения бортового устройства. № 12.
- Девянин П. Н. Обзор семейства ДП-моделей безопасности логического управления доступом и информационными потоками в компьютерных системах. № 5.
- Дрюченко М. А., Сирота А. А. Математическое и программное обеспечение для решения задачи стегоанализа на основе статистических и нейросетевых алгоритмов обработки информации. № 4.
- Кустов Г. А., Николаева М. А., Зотова О. Ф., Шарапов Р. А. Алгоритмы выбора превентивных мер снижения рисков. № 4.
- Лёвин В. Ю. Повышение криптостойкости протокола цифровой подписи на эллиптических кривых. № 5.
- Массель А. Г. Методологический подход к организации интеллектуальной поддержки исследований проблемы энергетической безопасности. № 9.
- Суханов А. В. Количественная оценка свойства защищенности информационных систем. № 1.
- Тарасюк М. В., Исаченко Ю. С. Подавление модуляции длин дейтаграмм в сетях шифрованной связи. № 1.
- Типикин А. П., Глазков А. С. Метод и функциональная организация контроля обращений и закрытия доступа к секторам файлов при хищении накопителя информации. № 5.

ПРОГРАММНАЯ ИНЖЕНЕРИЯ

- Грибова В. В., Черкезишвили Н. Н. Развитие онтологического подхода для автоматизации разработки пользовательских интерфейсов с динамическими данными. № 10.
- Зуев А. С., Петров Ю. И. Модернизация контекстных меню браузера, дерева папок и рабочего стола операционной системы Windows. № 10.
- Зуев А. С., Петров Ю. И. Описание модернизации дерева папок проводника Windows Explorer. № 1.
- Петропавловский М. В., Полевщиков Д. А. Особенности создания транслятора для языка генерации документов по шаблону в формате WordProcessingML в информационной системе государственной аккредитации. № 5.
- Сафронов В. В., Федорец О. Н. Метод построения эффективных моделей разработки программного обеспечения. № 1.
- Силаков Д. В. Информационно-аналитическая система для разработки и использования базового стандарта операционной системы Linux (LSB). № 5.
- Суховерхов С. Е., Штейнберг О. Б. Автоматическое распараллеливание рекуррентных циклов с проверкой устойчивости. № 1.

БАЗЫ ДАННЫХ

- Киселева Н. Н., Дударев В. А. База данных «Информационные ресурсы неорганической химии и материаловедения». № 12.
- Кудрявцев К. Я. Спектральный метод поиска ключевых слов в полнотекстовых базах данных. № 4.
- Полищук Ю. В., Черных Т. А. Построение хранилищ данных, ориентированных на работу с квазиструктурированным информационным наполнением документов. № 4.
- Токарчук А. М. Разработка паттерна ActiveModel для программирования интернет-приложения методом объектно-реляционного отражения. № 9.

КОМПЬЮТЕРНАЯ ГРАФИКА, ОБРАБОТКА ИЗОБРАЖЕНИЙ

- Барский А. Б. Трехмерная экранизация компьютерных объектов «живого» моделирования. № 9.
- Вахромцев Е. М. Применение методов «живого» моделирования для производственного мониторинга. № 7.
- Копцов А. А., Пронкин Ю. Н. Об одном новом методе обнаружения объектов. № 3.
- Коханенко И. К. Нейросетевой и морфологический компьютерный анализ изображений с использованием фракталов. № 9.
- Кривцов О. А., Кориков А. М. Методы отслеживания головы человека на видеокдрах: классификация и выбор. № 9.
- Кривцов О. А., Кориков А. М. Методы отслеживания головы человека на видеокдрах: состояние проблемы. № 7.
- Медведева Е. В. Нелинейная фильтрация статистически связанных видеопоследовательностей цифровых полутонных изображений. № 3.

НЕЙРОСЕТИ И НЕЙРОИНФОРМАТИКА

- Глушенко О. И., Горбатов С. А. Нейросетевая модель поддержки принятия решений в процессах муниципального бюджетирования. № 4.
- Жернаков С. В. Восстановление потерянной информации с датчиков ГТД в условиях бортовой реализации с использованием технологии нейронных сетей. № 4.
- Жернаков С. В. Идентификация характеристик ГТД на основе технологии нейронных сетей. № 3.
- Крючин О. В. Нейросетевой симулятор с автоматическим подбором активационных функций. № 7.
- Осипов В. Ю. Рекуррентная нейронная сеть с управляемыми синапсами. № 7.

WEB-ТЕХНОЛОГИИ

- Тарнавский Г. А., Чесноков С. С. Компьютерное моделирование в Интернете: краткий обзор Web-ресурсов. № 5.

ИНФОРМАЦИОННЫЕ ТЕХНОЛОГИИ В ОБРАЗОВАНИИ

- Герасимова И. Б. Системный анализ процессов передачи и усвоения знаний на основе когнитивного моделирования. № 9.
- Зак Ю. А. Математические методы обработки результатов тестирования квалификации и знаний специалистов и абитуриентов. № 9.
- Карданова Е. Ю., Карданов Р. С. Построение интервальных оценок статистик согласия при анализе результатов массового тестирования в рамках моделей Г. Раша. № 9.

Львович Я. Е., Королев Е. Н. Преимущества объектно-ориентированного подхода для проектирования систем обучения в интеллектуальных системах принятия решений. № 4.

Пантелеев Е. Р., Карпов Я. Э. Разработка и интерпретация решения расчетных задач в среде Web-обучения. № 4.

Сафронов В. В. Метод ветвей и границ в прямой и обратной задачах оптимального распределения времени на изучение цикла дисциплин. № 4.

ИНФОРМАЦИОННЫЕ ТЕХНОЛОГИИ В ЭКОНОМИКЕ И УПРАВЛЕНИИ

Бронштейн Е. М., Муслимова Г. Р. Формирование оптимальных портфелей, состоящих из инвестиционных проектов, с учетом групповых выплат. № 6.

Гюльмамедов Р. Г. Выбор стратегии информатизации инновационного региона. Когнитивный подход. № 12.

Димитриенко Ю. И., Димитриенко О. Ю. Кластерно-континуальное моделирование в экономике на основе методов механики многомерных сплошных сред. № 8.

Димитриенко Ю. И., Димитриенко О. Ю. Модель деформируемых кластеров для анализа динамических данных в экономике. № 9.

Казаков П. В. Автоматизация синтеза оптимальных инвестиционных портфелей на основе кластерного генетического алгоритма. № 12.

Кравченко Т. К. Разработка системы ключевых показателей для оценки эффективности процессов службы информационных технологий. № 8.

Новицкий В. О., Карпов В. И. Методология исследования и моделирования сложных систем управления для предприятий и компаний зернового сектора АПК. № 9.

Чертовской В. Д. Анализ процесса согласованного автоматизированного планирования в иерархической системе управления производством. № 6.

Чучуева И. А. Модель экстраполяции временных рядов по выборке максимального подобию. № 12.

Штрик А. А. Анализ воздействия кризиса 2009 года на состояние и перспективы развития ИТ-рынка России. № 8.

ИНФОРМАЦИОННЫЕ ТЕХНОЛОГИИ В МЕДИЦИНЕ И БИОЛОГИИ

Гоменюк С. М., Емельянов А. О., Карпенко А. П., Чернецов С. А. Обзор методов и систем прогнозирования оптимальных доз инсулина для больных сахарным диабетом I типа. № 3.

Зыкин С. В., Полуянов А. Н., Чернышев А. К., Ревзин А. И. Технология подготовки и анализа данных для построения медицинских оценочных шкал. № 12.

Кузнецов А. А. Информационные технологии донозологической диагностики. № 8.

Кузнецов А. А. Системный анализ и обработка электрокардиографической информации. № 5.

Май В. П., Мельман С. В. Система объемной визуализации объектов компьютерной томографии. № 5.

Малыхин В. М., Меркушева А. В. Нейронные сети с вейвлет-преобразованием для классификации биосигналов. № 12.

Митрушкин Е. И., Робский С. Ю. Оценка вклада информационной системы в эффективность медицинского учреждения. № 8.

Острейковский В. А., Дрожжин Е. В., Федоров Д. А. О формализации предметной области интеллектуальной информационной системы поддержки принятия решений врачом. № 3.

Петрухин А. В., Золотарев А. В. Методика автоматизации начальных этапов процесса проектирования биомеханических систем. № 5.

Токмачев М. С., Рязанцев П. П. Формирование персонализированной базы данных и программная реализация исследования здоровья населения региона. № 3.

ПРИКЛАДНЫЕ ИНФОРМАЦИОННЫЕ СИСТЕМЫ

Девятьяров Д. А. Нечеткие дескрипторы молекул в задаче «структура–свойство». № 3.

Домников П. А. Распараллеливание задач конечно-элементного моделирования трехмерных магнитотеллурических полей. № 3.

Замятин А. В. Распределенные вычисления в задачах автоматизированной интерпретации аэрокосмических изображений. № 6.

Михеева Т. И., Сапрыкин О. Н. Паттерны проектирования подсистемы анализа состояния пространственно-координированных объектов интеллектуальной транспортной системы. № 7.

Силина А. Ю., Гермашев И. В., Васильева В. Д., Дербишер В. Е. Комплексная оценка научной деятельности. № 3.

Углова Е. С., Моисеева Л. С. Математическая модель и информационная система поддержки принятия решений в задачах коррозионной защиты. № 2.

Чичко А. Н., Сачек О. А., Лихоузов С. Г. Методы автоматизации обработки изображений микроструктур перлитных сталей. № 7.

Яковлев А. А., Яковлева Е. В. Автоматизация синтеза и выбора технических решений преобразователей энергии. № 11.

Ясницкий Л. Н., Петров А. М., Сичинава З. И. Технологии построения детектора лжи на основе аппарата искусственных нейронных сетей. № 11.

КОДИРОВАНИЕ И ОБРАБОТКА СИГНАЛОВ

Агиевич С. Н. Применение локальных сплайнов для оценивания параметров распределений. № 8.

Дворников С. В., Дворников С. С. Обнаружение сигналов с высоким различием динамики их амплитуд. № 2.

Дворников С. В., Дворников С. С., Коноплев М. А. Алгоритм распознавания сигналов радиосвязи на основе симметрических матриц. № 9.

Кухаренко Б. Г. Байесовская фильтрация в технологии спектрального анализа на основе быстрого преобразования Фурье. № 8.

Назаров Л. Е. Алгоритмы посимвольного приема сигналов. № 2.

Савченко В. В., Акатьев Д. Ю. Обучение звуковому строю языка глухонемых и слабослышащих на основе информационной теории восприятия речи. № 2.

КОНТРОЛЬ И ДИАГНОСТИКА СИСТЕМ

Дюндиков Е. Т., Качкин А. А. Технология динамической интеграции и представления разнородных данных для анализа и оценки состояния многопараметрических объектов. № 2.

ПРОГРАММНЫЕ ПРОДУКТЫ И СИСТЕМЫ

Гетманов В. Г., Орлов С. Е. Программный комплекс для спектрально-временного анализа нестационарных колебательных сигналов. № 10.

ДИСКУССИОННЫЙ КЛУБ

Вяткин В. Б. Введение в синергетическую теорию информации. № 12.

Гуревич И. М. Информация как универсальная неоднородность. № 4.

Нариньяни А. С. Между эволюцией и сверхвысокими технологиями: новый человек ближайшего будущего. № 1.

ОБМЕН ОПЫТОМ

Кузнецов А. В., Суриков Е. Н. Проблемы внедрения ERP-систем при разработке и мелкосерийном производстве высокотехнологичной продукции. № 7.

ИНФОРМАЦИЯ

Государственный университет Высшая школа экономики подписал соглашение с IEEE CS. № 8.

Рожкова М. В. Мои заметки о Mobile World Congress 2010. Тезисно: значимые события и тренды. № 4.

Тарнавский Г. А. Первый международный конгресс по «облачным» технологиям — Cloud Computing Congress CCC-2010. № 10.

Указатель приложений к журналу «Информационные технологии», опубликованных в 2010 г.

Барский А. Б. Оптимизационные задачи в основе пакета параллельных прикладных программ и системы информационного обслуживания Центра GRID-технологий. № 10.

Барский А. Б., Дмитриев А. А., Барская О. А. Медицинские информационно-справочные системы на логических нейронных сетях. № 1.

Волокоморов С. В., Каганов Ю. Т., Карпенко А. П. Моделирование и оптимизация некоторых параллельных механизмов. № 5.

Кухаренко Б. Г. Анализ независимых компонент и скрытая марковская модель для определения доминантных компонент многомерных временных рядов. № 11.

Львович Я. Е., Белецкая С. Ю. Адаптивные методы оптимального проектирования. № 7.

Макаров И. М., Лохин В. М., Манько С. В., Романов М. П., Кадочников М. В. Технологии обработки знаний в задачах управления автономными мехатронно-модульными реконфигурируемыми роботами. № 8.

Нариньяни А. С., Гофман И. Д., Липатов А. А., Инишев Д. А. Интеллектуальная технология неопределенного календарно-ресурсного планирования и управления проектами *Time-EX*[®]. № 2.

Позин Б. А. Ввод в действие информационных систем и сопровождение их программного обеспечения. № 4.

Свиридов А. П. Статистическая теория обучения. № 9.

Шилов В. В. Говорящие головы. Мифы и реальность в истории механических генераторов речи. № 12.

Шилов В. В. От Гомера до Чапека. Краткая история механических автоматов. № 3.

Штрик. А. А. Состояние и перспективы формирования информационного общества в России до 2015 года. № 6.

CONTENTS

Gurevich I. M. *Information Methods of Research of Physical Systems: the Review of the First Achievements* 2

In article the overview of the works in which are using informational research techniques of physical systems is given. There are shown the fundamental works of the given direction of researches. The findings of investigation gained by the author, including the results published in the journal "Information technology" are enumerated. The overview of the works of foreign scientists is given. It is shown, that the works of the author and foreign scientists (American, Canadian, European, Chinese ...) confirm primacy of informational laws. Informational laws (informatics laws) have general, universal character, operate in all possible Universes, even in the Universes with different physical laws. The overview shows, that the priority of using by informational methods of physical effects belongs to the author though the last findings of research of foreign scientists are very interesting and important.

Keywords: the information, informational methods, the informational approach, informational performances, physical systems, informational and physical laws of the nature, the Universe, findings of investigations

Alguliev R. M., Imamverdiyev Yu. N., Derakhshandeh S. A. *Ways of Increasing the Accuracy of Information Security Risk Assessment Methods* 6

The comparative analysis of the existing approaches to information security risk assessment is presented. Traditional risk assessment methods are mainly based on qualitative measurements which are highly subjective and do not possess necessary accuracy. Uncertainty is necessary and sufficient condition for risks in decision making and adequate risk assessment methods have to consider the nature of uncertainty. Probabilistic, fuzzy logic and expert assessment approaches to modeling of various types of uncertainty are analyzed in the paper. Main advan-

tages and disadvantages of information security risk assessment approaches are structured and the important and actual research directions in this area are identified.

Keywords: information security, threat, vulnerability, risk assessment, fuzzy logic

Generalov D. N. Counteraction to Infringement of Integrity Modules of the Specialised Software of the Onboard Device. 12

In work the order of the basic operations of a choice of system counteraction to infringement of integrity modules of the specialized software of the onboard device is considered. The developed algorithm of protection against not authorized access to the moment of start, session end is resulted, at possible liquidation of protection the specialized software on the basis of distribution of data on contents.

Keywords: onboard device, the specialized software, information protection, infringement of integrity data

Saak A. E. Local-Optimal Synthesis of Schedules for GRID-Technologies 16

The solution of optimal resources allocation in Grid-systems and multiprocessor computer systems problem is given. User's tasks are construed by rectangles with its sides are equal to declared processors' resources and processing time. It is suggested a criterion function for functioning optimization i.e. minimization of area of comprehensive resort rectangle of additive graphic of original resort rectangles family — users requests. Paired synthesis of resort rectangles is considered in details. General tasks case brings to paired optimization by as tandeming as parallel algorithms.

Keywords: Grid-system, multiprocessor computer system, supervision, local schedule, optimal schedule, principle of minimum comprehensive resort rectangle area

Umbitaliev A. A., Shipilov N. N., Fahmi Sh. S. IP-Core Transcoder Videocontent 21

The new approach to distribution compression a videocontent with simultaneous use of two alternative compression technologies of moving images is considered: 1) on the basis of three-dimensional discrete cosine transformations (this technology not standardized for today) and 2) on the basis of two-dimensional discrete cosine transformations of the digital television signal realised in standards MPEG-2, MPEG-4 and H.264. Within the limits of the specified approach the method of fast transcoding of the main videodata stream is compressed on technology three-dimensional discrete cosine of transformation, in compression a stream of standard format MPEG-2 is offered.

Keywords: DCT, IP-core, transcoder

Panko S. P., Nagorny A. V. Features of Multistation Access in a Network with a Shared Resource 24

This article is about the influence of collisions on productivity of telecommunication networks with multistation access of workstations of the users to a shared resource. The technique for increase of productivity of such networks is offered.

Keywords: ethernet, multistation access, shared resource, packet network

Abbasova T. S. Increase of Efficiency of Operation of High-Speed Cable System by Means of Virtual Technologies . 28

Methods of increase of efficiency are considered at operation of cable system for a data-processing centre; the design procedure of the basic tested parameters of a network for a cable "twisted pair" categories 6, 6a in view of features of operation is resulted at work on high frequencies for achievement of throughput 10 Gb/s; presents an analytical evaluation of the effectiveness of the use of virtual technologies and statistical evaluation of the cost of ownership of the virtual structure.

Keywords: computing power, electromagnetic compatibility, virtual mode

Struchenkov V. I., Kozlov A. N., Egunov A. S. The Piece-Linear Approximation of Flat Curves by Special Restricts . . 32

In the computer aided design lines of linear structures, problems arise approximation of planar curves given discrete sequence elements of a certain type (straight lines, arcs or parabolas of the second degree, as well as klotoid). There are restrictions on the parameters of the elements. The number of elements is unknown. The article deals with the simplest of tasks-element approximation, in which the elements are straight lines, in the presence of a number of limitations. The problem is solved using dynamic programming.

Keywords: approximations, limitations, dynamic programming

Galjano F. R. An Earth's Surface Classification Algorithm Based on Matrix Singular Value Decomposition. 35

The paper is concerned with the application of classification algorithm, based on matrix singular value decomposition, to the task of analysis of earth's surface regions using remote sensing data. A modification of SVD-classification algorithm, that increase accuracy by taking into account weights of properties and classes is proposed. A comparison of remote sensing data's classification using classical and improved algorithm is given.

Keywords: remote sensing, classification, singular value decomposition

Kazakov P. V. The Automation of the Synthesis of the Optimal Investment Portfolio Based on the Cluster Extension of a Genetic Algorithm 38

The manner for synthesis of the optimal investment portfolio is considered. This manner based on the cluster modification of a genetic algorithm. Its using allow effectively to solve a such problem as multimodal and allow for a yield and a risk of an investment planning. The results of practical using a cluster based genetic algorithm for hunting of the optimal investment portfolio with a factual investment data are presented.

Keywords: investment planning, evolutionary computations, cluster based genetic algorithm, multimodal optimization

Chuchueva I. A. *The Time Series Extrapolation Model Based on Maximum Likeness Set.* 43

In paper the time series extrapolation model based on maximum likeness set is presented. There is the description of the extrapolation model and the model identification method. Test of the model efficiency has been implemented on the electricity consumption and prices time series.

Keywords: forecast, extrapolation, electricity consumption, electricity price, time series

Gulmammadov R. H. *Choice of Strategy of Informatization of Innovative Region. Cognitive Approach* 48

In the conditions of transition to knowledge-based economy information becomes one of leading factors of innovative development of regions. The article reviews problem of a choice of strategy of informatization of innovative region. Difficulties of the solution the problem caused by its ill structured character, limiting possibilities of using of traditional methods of economic-mathematical modelling and decision-making are marked. The methodology of cognitive support of problem, opening new possibilities for overcoming of these difficulties is offered. Tools of cognitive support can be useful to managers of ICT-projects and the administrators making the decision.

Keywords: knowledge-based economy, innovative region, strategy of informatization, cognitive approach

Malykhin V. M., Merkusheva A. V. *Neural Networks Including Wavelet Transformation for Biosignal Classification* . . 51

The methods for biomedical signal (BS) classification are considered that provide non-invasive type of diagnostics of several organs. The fundamentals of this method are artificial neural network (NN) implementing wavelet transformation and principles of patterns recognition. The NN structure, algorithm for its learning and elements of applications in cardiology are considered.

Keywords: biological signals, classification, pattern recognition, wavelet transformation, neural network, learning

Zykin S. V., Poluyanov A. N., Chernyshev A. K., Revzin A. I. *Technology of Preparation and the Analysis of the Data for Construction of Medical Estimated* 57

In work the technology of preparation and the analysis of the data on an example of construction of a scale of an estimation of a condition of patients is considered. The iterative approach with the correctness control to formation of multidimensional data presentation from initial relational representation is offered. The technology of carrying out of the analysis of the data and creation of estimated scales is considered.

Keywords: intermodels mapping, hypercube, medicine scales

Kiselyova N. N., Dudarev V. A. *"Information Resources on Inorganic Chemistry and Materials Science" Database* . 63

IRIC database (DB) — Information Resources on Inorganic Chemistry database containing data on computer information resources in the field of inorganic chemistry and materials science is developed. IRIC DB is accessible from the Internet global network (<http://iric.imet-db.ru>). DB includes information on organizations-owners (name, country, post and electronic addresses, phones, etc.), keywords which are used for data search, and references as well. Currently DB contains data on more than 100 information systems. IRIC DB is intended for navigation of the specialists in the sea of information resources in the above-stated subject domains. IRIC DB has two language versions — Russian and English.

Keywords: database, information resources in the field of inorganic chemistry and materials science, Internet

Vyatkin V. B. *Introduction in Synergistic Information Theory* 67

They are considered features of reflection of discrete systems through the parts and their relations with environment with help synergetic approach to definition of information quantity. New information functions are received and their interrelation is shown with traditional information measures. Structured conservation law of chaos and order, information law of reflection and conservation law of information are established.

Keywords: information, negentropy, entropy, reflection, system, part, structure, chaos, order

Адрес редакции:

107076, Москва, Стромьинский пер., 4

Телефон редакции журнала (499) 269-5510

E-mail: it@novtex.ru

Дизайнер Т.Н. Погорелова. Технический редактор Е. В. Конова.

Корректор Е. В. Комиссарова.

Сдано в набор 07.10.2010. Подписано в печать 18.11.2010. Формат 60×88 1/8. Бумага офсетная. Печать офсетная.

Усл. печ. л. 9,8. Уч.-изд. л. 11,50. Заказ 947. Цена договорная.

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

Свидетельство о регистрации ПИ № 77-15565 от 02 июня 2003 г.

Отпечатано в ООО "Подольская Периодика"

142110, Московская обл., г. Подольск, ул. Кирова, 15